

INTELIGENCIA ARTIFICIAL Y SECTOR PÚBLICO

Retos, límites y medios

EDUARDO GAMERO CASADO
Director

FRANCISCO L. PÉREZ GUERRERO
Coordinador



monografías

ALTA CALIDAD EN
INVESTIGACIÓN
JURÍDICA



**INTELIGENCIA ARTIFICIAL
Y SECTOR PÚBLICO**
Retos, límites y medios

COMITÉ CIENTÍFICO DE LA EDITORIAL TIRANT LO BLANCH

MARÍA JOSÉ AÑÓN ROIG

*Catedrática de Filosofía del Derecho
de la Universidad de Valencia*

ANA CAÑIZARES LASO

*Catedrática de Derecho Civil de
la Universidad de Málaga*

JORGE A. CERDIO HERRÁN

*Catedrático de Teoría y Filosofía de Derecho.
Instituto Tecnológico Autónomo de México*

JOSÉ RAMÓN COSSÍO DÍAZ

*Ministro en retiro de la Suprema
Corte de Justicia de la Nación y
miembro de El Colegio Nacional*

MARÍA LUISA CUERDA ARNAU

*Catedrática de Derecho Penal de la
Universidad Jaume I de Castellón*

CARMEN DOMÍNGUEZ HIDALGO

*Catedrática de Derecho Civil de la
Pontificia Universidad Católica de Chile*

EDUARDO FERRER MAC-GREGOR POISOT

*Juez de la Corte Interamericana
de Derechos Humanos*

Investigador del Instituto de

Investigaciones Jurídicas de la UNAM

OWEN FISS

*Catedrático emérito de Teoría del Derecho
de la Universidad de Yale (EEUU)*

JOSÉ ANTONIO GARCÍA-CRUCES GONZÁLEZ

Catedrático de Derecho Mercantil de la UNED

JOSÉ LUIS GONZÁLEZ CUSSAC

*Catedrático de Derecho Penal de
la Universidad de Valencia*

LUIS LÓPEZ GUERRA

*Catedrático de Derecho Constitucional
de la Universidad Carlos III de Madrid*

ÁNGEL M. LÓPEZ Y LÓPEZ

*Catedrático de Derecho Civil de
la Universidad de Sevilla*

RUTH ZIMMERLING

*Catedrática de Ciencia Política de la
Universidad de Mainz (Alemania)*

MARTA LORENTE SARIÑENA

*Catedrática de Historia del Derecho de
la Universidad Autónoma de Madrid*

JAVIER DE LUCAS MARTÍN

*Catedrático de Filosofía del Derecho y
Filosofía Política de la Universidad de Valencia*

VÍCTOR MORENO CATENA

*Catedrático de Derecho Procesal de la
Universidad Carlos III de Madrid*

FRANCISCO MUÑOZ CONDE

*Catedrático de Derecho Penal de la
Universidad Pablo de Olavide de Sevilla*

ANGELIKA NUSSBERGER

*Catedrática de Derecho Constitucional
e Internacional en la Universidad
de Colonia (Alemania)*

Miembro de la Comisión de Venecia

HÉCTOR OLASOLO ALONSO

*Catedrático de Derecho Internacional de la
Universidad del Rosario (Colombia) y
Presidente del Instituto Ibero-Americano
de La Haya (Holanda)*

LUCIANO PAREJO ALFONSO

*Catedrático de Derecho Administrativo
de la Universidad Carlos III de Madrid*

CONSUELO RAMÓN CHORNET

*Catedrática de Derecho Internacional
Público y Relaciones Internacionales
de la Universidad de Valencia*

TOMÁS SALA FRANCO

*Catedrático de Derecho del Trabajo y de la
Seguridad Social de la Universidad de Valencia*

IGNACIO SANCHO GARGALLO

*Magistrado de la Sala Primera (Civil)
del Tribunal Supremo de España*

ELISA SPECKMANN GUERRA

*Directora del Instituto de Investigaciones
Históricas de la UNAM*

Fueron miembros de este Comité:

Emilio Beltrán Sánchez, Rosario Valpuesta Fernández y Tomás S. Vives Antón

Procedimiento de selección de originales, ver página web:

www.tirant.net/index.php/editorial/procedimiento-de-seleccion-de-originales

INTELIGENCIA ARTIFICIAL Y SECTOR PÚBLICO

Retos, límites y medios

EDUARDO GAMERO CASADO

Director

FRANCISCO L. PÉREZ GUERRERO

Coordinador



UNIVERSIDAD
**PABLO
OLAVIDE**
SEVILLA



Junta de Andalucía



Unión Europea
Fondo Europeo
de Desarrollo Regional



Andalucía
se mueve con Europa

Programa Operativo FEDER 2014-2020

Proyecto de investigación UPO-1381574,
“Inteligencia artificial y Derecho Administrativo”

Red DAIA

Derecho Administrativo e Inteligencia Artificial

<http://reddaia.org/>

tirant lo blanch

Valencia, 2023

Copyright © 2023

Todos los derechos reservados. Ni la totalidad ni parte de este libro puede reproducirse o transmitirse por ningún procedimiento electrónico o mecánico, incluyendo fotocopia, grabación magnética, o cualquier almacenamiento de información y sistema de recuperación sin permiso escrito de los autores y del editor.

En caso de erratas y actualizaciones, la Editorial Tirant lo Blanch publicará la pertinente corrección en la página web www.tirant.com.

© Eduardo Gamero Casado (Dir.)
Francisco L. Pérez Guerrero (Coord.)

© TIRANT LO BLANCH
EDITA: TIRANT LO BLANCH
C/ Artes Gráficas, 14 - 46010 - Valencia
TELFs.: 96/361 00 48 - 50
FAX: 96/369 41 51
Email: tlb@tirant.com
www.tirant.com
Librería virtual: www.tirant.es
DEPÓSITO LEGAL: V-1554-2023
ISBN: 978-84-1169-063-8
MAQUETA: Dissert Ediciones

Si tiene alguna queja o sugerencia, envíenos un mail a: atencioncliente@tirant.com. En caso de no ser atendida su sugerencia, por favor, lea en www.tirant.net/index.php/empresa/politicas-de-empresa nuestro procedimiento de quejas.

Responsabilidad Social Corporativa: http://www.tirant.net/Docs/RSC_Tirant.pdf

Autores

EDUARDO GAMERO CASADO
FRANCISCO L. PÉREZ GUERRERO
FRANCISCO GONZÁLEZ CABANES
NORBERTO DÍAZ DÍAZ
CARLOS TORRECILLA-SALINAS
LUCA TANGI
PETER ULRICH
MARINA MANZONI
SVEN SCHADE
EVA MARTÍNEZ-RODRÍGUEZ
FRANCESCO PIGNATELLI
ANTONIO DAVID BERNING PRIETO
ISAAC MARTÍN DELGADO
JULI PONCE SOLÉ
RUBÉN MARTÍNEZ GUTIÉRREZ
LORENZO COTINO HUESO
JULIÁN VALERO TORRIJOS
AGUSTÍ CERRILLO I MARTÍNEZ
ISABEL GALLEGO CÓRCOLES
RAMON GALINDO CALDÉS
JUAN CARLOS HERNÁNDEZ PEÑA
MARÍA PÉREZ NARANJO
M. ASCENSIÓN MORO CORDERO
ORIO L MIR PUIGPELAT
EMILIA GÓMEZ
ISABELLE HUPONT TORRES
IGNACIO SÁNCHEZ
DAVID FERNÁNDEZ LLORCA
ALEJANDRO HUERGO LORA
IGNACIO ALAMILLO DOMINGO

Índice

Abreviaturas y acrónimos.....	25
Prefacio	31
EDUARDO GAMERO CASADO, FRANCISCO L. PÉREZ GUERRERO	

PRIMERA PARTE: LA INTELIGENCIA ARTIFICIAL Y SUS APLICACIONES EN EL SECTOR PÚBLICO

Capítulo I

¿Qué es la Inteligencia Artificial?
FRANCISCO GONZÁLEZ CABANES
NORBERTO DÍAZ DÍAZ

1.- INTRODUCCIÓN A LA INTELIGENCIA ARTIFICIAL	38
1.1.- Conceptos Previos y Aclaraciones.....	38
1.2.- Historia.....	43
1.3.- Casos de Uso de la Inteligencia Artificial	45
2.- EL PROCESO DE EXTRACCIÓN DE CONOCIMIENTO: KDD	47
2.1.- Mecanismos de Aprendizaje	49
2.2.- Generalización del Conocimiento	50
3.- ALGORITMOS CLÁSICOS DE APRENDIZAJE.....	51
3.1.- Aprendizaje Supervisado	51
3.1.1.- Árboles de Decisión	51
3.2.- Aprendizaje No Supervisado	54
3.2.1.- Clustering.....	54
3.2.2.- Biclustering.....	56
3.3.- Redes Neuronales.....	57
4.- LA REVOLUCIÓN DIGITAL Y DEL DATO	61
4.1.- La Importancia de Los Datos	63
5.- ALGORITMOS MODERNOS DE APRENDIZAJE: DEEP LEARNING.....	64
5.1.- Redes Neuronales Profundas: Deep Learning	64
5.2.- Diferencias entre Machine Learning y Deep Learning.....	64
5.3.- Transversalidad del Deep Learning.....	67
5.4.- Transfer Learning.....	68
6.- EL PRESENTE Y FUTURO DE LA INTELIGENCIA ARTIFICIAL...	69
7.- CONCLUSIONES	70
8.- BIBLIOGRAFÍA CITADA.....	71

Capítulo II

¿Para qué sirve la Inteligencia Artificial en el sector público? Casos de uso y perspectivas de aplicación

CARLOS TORRECILLA-SALINAS, LUCA TANGI, PETER ULRICH, MARINA MANZONI, SVEN SCHADE, EVA MARTÍNEZ-RODRÍGUEZ, FRANCESCO PIGNATELLI

1.- ¿CÓMO DEFINIR LA INTELIGENCIA ARTIFICIAL?	74
2.- ¿POR QUÉ ES IMPORTANTE LA INTELIGENCIA ARTIFICIAL PARA EL SECTOR PÚBLICO?	76
3.- ¿CÓMO SE UTILIZA ACTUALMENTE LA INTELIGENCIA ARTIFICIAL EN EL SECTOR PÚBLICO?	79
4.- CONCLUSIONES Y RECOMENDACIONES	88
BIBLIOGRAFÍA CITADA	90

SEGUNDA PARTE: RÉGIMEN JURÍDICO

Sección primera: Algoritmos y Derecho

Capítulo III

La naturaleza jurídica de los algoritmos

ANTONIO DAVID BERNING PRIETO

1.- INTRODUCCIÓN.....	96
2.- CONCEPTUALIZACIÓN DE LOS ALGORITMOS Y DISTINCIÓN DE FIGURAS AFINES.....	97
3.- POSICIONES DOCTRINALES EXISTENTES	101
4.- LA VERDADERA PROBLEMÁTICA QUE SUBYACE: LA NECESIDAD DE DOTAR DE GARANTÍAS JURÍDICAS SUFICIENTES AL USO DE ALGORITMOS POR PARTE DE LA ADMINISTRACIÓN	104
4.1.- La única solución posible: dotar a los algoritmos de un régimen jurídico adecuado	104
4.3.- Del control de las decisiones adoptadas por los poderes públicos mediante el uso de algoritmos predictivos.....	113
4.4.- De los diversos sesgos a que pueden dar lugar la actividad administrativa al hacer uso de los algoritmos	115
4.4.1.- Posibles sesgos derivados del empleo de algoritmos en el procedimiento administrativo.....	116
4.4.2.- Posibles sesgos derivados del empleo de algoritmos en la prestación de servicios públicos.....	119

4.4.3.- Posibles sesgos en usos algorítmicos relacionados con el ejercicio de las potestades de investigación o comprobación de las administraciones públicas	121
4.5.- De la necesidad de prever, en las licitaciones públicas, la cesión en favor de la Administración de los derechos de propiedad intelectual o industrial inherentes a los algoritmos que se desarrollen o se utilicen	124
5.- CONCLUSIONES	127
BIBLIOGRAFÍA CITADA	129

Capítulo IV

La aplicación del principio de transparencia a la actividad administrativa algorítmica

ISAAC MARTÍN DELGADO

1.- A MODO DE INTRODUCCIÓN	132
1.1.- El gran reto para el jurista	132
1.2.- Algunas preguntas previas	136
2.- CONCEPTO Y CONTEXTO	140
3.- PRETEXTO	146
4.- LA TRANSPARENCIA DE LA ACTIVIDAD ADMINISTRATIVA ALGORÍTMICA	153
4.1.- Premisa	153
4.2.- Un ejercicio práctico: regular la transparencia de la actividad administrativa algorítmica	154
4.3.- Publicidad activa y derecho de acceso en relación con el uso de IA en el sector público	162
5.- A MODO CONCLUSIÓN: ALGUNAS GARANTÍAS NECESARIAS	172
BIBLIOGRAFÍA CITADA	188

Capítulo V

Seres humanos e inteligencia artificial: discrecionalidad artificial, reserva de humanidad y supervisión humana

JULI PONCE SOLÉ

1.- INTRODUCCIÓN: OBJETO DE ESTE ANÁLISIS	196
2.- HUMANOS, EMPATÍA, INTELIGENCIA ARTIFICIAL, DISCRECIONALIDAD	199
3.- RESERVA DE HUMANIDAD Y “DISCRECIONALIDAD ARTIFICIAL”	206
3.1.- Reserva de humanidad general	207

3.2.- Reserva de humanidad referida al ejercicio de la discrecionalidad	209
4.- SUPERVISIÓN HUMANA DE LA IA: <i>THE HUMAN IN THE LOOP</i>	213
5.- CONCLUSIONES Y TEMAS PARA CONSIDERAR EN EL FUTURO	218
BIBLIOGRAFÍA CITADA	224

Capítulo VI

Responsabilidad administrativa por el uso de la Inteligencia Artificial

RUBÉN MARTÍNEZ GUTIÉRREZ

1.- PLANTEAMIENTO GENERAL.....	228
2.- ASPECTOS CONCEPTUALES DE LA RESPONSABILIDAD ADMINISTRATIVA POR EL USO DE LA IA	229
2.1.- Concepto de Inteligencia Artificial a los efectos del presente estudio.....	229
2.2.- Concepto legal de responsabilidad administrativa	232
3.- BREVE REFERENCIA A LAS ESPECIALIDADES DEL PROCEDIMIENTO PARA EXIGIR RESPONSABILIDAD ADMINISTRATIVA	235
3.1.- Iniciación del procedimiento.....	235
3.1.1.- Inicio de oficio. Una modalidad de inicio con buen encaje en supuestos de daño por IA.....	235
3.1.2.- Inicio a solicitud de las personas interesadas	236
3.2.- Instrucción del procedimiento	237
3.2.1.- Los informes y dictámenes en los procedimientos de responsabilidad patrimonial	237
3.2.2.- Las especialidades del trámite de audiencia en materia de responsabilidad patrimonial	239
3.3.- Terminación del procedimiento	239
3.3.1.- Los órganos competentes para resolver procedimientos de responsabilidad patrimonial	239
3.3.2.- La resolución de los procedimientos de responsabilidad patrimonial. Especialidades	240
4.- EL PARÁMETRO DE CALIDAD Y LOS ELEMENTOS PARA LA EXIGIBILIDAD DE LA RESPONSABILIDAD ADMINISTRATIVA POR IA	241
4.1.- Elementos para la exigibilidad de la responsabilidad administrativa.....	241
4.2.- Parámetro de la calidad en el uso de tecnologías de IA en el sector público	244
5.- SUPUESTOS DE POSIBLE APLICACIÓN DE LA RESPONSABILIDAD ADMINISTRATIVA POR INTELIGENCIA ARTIFICIAL.....	247

6.- CONCLUSIONES	249
BIBLIOGRAFÍA CITADA	251

Sección Segunda: Los derechos y la legislación administrativa como límites a la IA

Capítulo VII

Discriminación, sesgos e igualdad de la inteligencia artificial en el sector público LORENZO COTINO HUESO

1.- SESGOS Y DISCRIMINACIÓN ALGORÍTMICA ¿DE QUÉ HABLAMOS? Y ALGUNOS EJEMPLOS	260
1.1.- Sesgos, errores y discriminación algorítmica y la conveniencia de abordarlos conjuntamente.....	260
1.2.- Algunos ejemplos de discriminación algorítmica pública	262
2.- ¿POR QUÉ SE GENERAN LA DISCRIMINACIÓN Y LOS SESGOS ALGORÍTMICOS?	269
3.- ¿RESPECTO DE QUÉ IA PÚBLICA ES EXIGIBLE LA NO DISCRIMINACIÓN?	272
3.1.- La igualdad es exigible respecto de los sistemas de IA públicos, sus usos y sus efectos, aunque resulte difícil su categorización	272
3.2.- El grado de legalidad exigible para diseño, implantación y uso del sistema de IA público	276
4.- ¿CUÁNDO EL USO PÚBLICO DE UN SISTEMA DE IA TIENE IMPACTO Y SUFICIENTE RELEVANCIA JURÍDICA Y PUEDE CONSTITUIR UNA DISCRIMINACIÓN ALGORÍTMICA?	277
4.1.- Los criterios y variables para determinar el impacto, nivel de riesgo y relevancia jurídica de los sistemas algorítmicos públicos	277
4.2.- ¿Cuándo el sesgo y el error algorítmico es relevante para constituir una discriminación? Datos estadísticos en discriminaciones indirectas y la necesidad de “recalibrarlos” a la baja para la IA pública.....	283
5.- DISCRIMINACIONES ALGORÍTMICAS POR ERROR, DIRECTAS, INDIRECTAS, INTERESECCIONALES Y ACCIONES POSITIVAS ALGORÍTMICAS	285
5.1.- La discriminación o sesgo algoritmo será en la mayoría de supuestos “discriminación por error”	286
5.2.- Discriminaciones algorítmicas directas, interseccionales y por asociación.....	286
5.3.- La complejidad de las discriminaciones algorítmicas indirectas ..	287
5.4.- Acciones positivas algorítmicas, deber de diferenciación y modulaciones y compensaciones para la minimización de la discriminación.....	289

6.- ¿CUÁNDO UN SISTEMA ALGORÍTMICO PÚBLICO QUE NO TIENE EN CUENTA CIRCUNSTANCIAS ESPECIALMENTE SOSPECHOSAS ES DISCRIMINATORIO?	291
6.1.- La igualdad “en” la ley o norma que habilite el uso del sistema de IA y el control laxo de la posibilidad de establecer diferencias de tratos razonables y no arbitrarios.....	291
6.2.- Control de discriminación y arbitrariedad y exigencias de motivación en los casos de aplicación algorítmica de la ley. Decisiones automatizadas discrecionales y supuestos no especialmente sospechosos	294
7.- ¿CUÁNDO UN SISTEMA ALGORÍTMICO PÚBLICO QUE SÍ TIENE EN CUENTA CIRCUNSTANCIAS ESPECIALMENTE SOSPECHOSAS ES DISCRIMINATORIO?	298
7.1.- Las garantías frente a las discriminaciones algorítmicas especialmente prohibidas se superponen por lo general a las garantías de los tratamientos de datos especialmente protegidos	298
7.2.- En la mayoría de casos será necesaria una ley de calidad y con garantías que habilite a que un sistema de IA pueda diferenciar en razón de datos especialmente sospechosos de discriminación. Y no la hay	299
7.3.- Un canon de control “mucho más” estricto, e incluso excepcionalmente estricto	302
7.4.- La necesidad de regular legalmente el uso de datos especialmente protegidos para poder evaluar y minimizar la posible discriminación.....	303
8.- ZONAS GRISES. GARANTÍAS ANTE EL USO POR EL SISTEMA DE IA PÚBLICO DE “PROXIES” Y GRUPOS ALGORÍTMICOS.....	304
8.1.- Tratamientos diferentes a partir de “proxies” o datos afines o indirectos de grupos especialmente protegidos.....	304
8.2.- Creación dinámica de grupos y colectivos humanos y la privacidad y no discriminación colectiva y de grupo.....	305
9.-¿CÓMO PROBAR LA DISCRIMINACIÓN ALGORÍTMICA PÚBLICA? INVERSIÓN GENERAL DE LA CARGA DE LA PRUEBA Y LA NECESIDAD DE LA TRANSPARENCIA Y EXPLICABILIDAD.....	307
9.1.- Los indicios que debe aportar la persona afectada por el sistema de IA público. Inversión general de la carga de la prueba y sus excepciones.....	308
9.2.- Lo que debe probar el proveedor o el usuario del sistema de IA público	311
10.- MÁS VALE PREVENIR QUE DISCRIMINAR. ¿CUÁLES SON LOS REMEDIOS Y MEJORES PRÁCTICAS PARA EVITAR LA DISCRIMINACIÓN ALGORÍTMICA PÚBLICA?	313

11.- LA (CRECIENTE) OBLIGACIÓN DE DISEÑO Y ESTUDIOS DE IMPACTO FRENTE A LA DISCRIMINACIÓN ALGORÍTMICA DE LOS SISTEMAS DE IA PÚBLICOS	320
11.1.- Los deberes de responsabilidad proactiva, privacidad en el diseño o estudio de impacto ya son exigibles en general a los usos públicos de la IA.....	320
11.2.- Los usos públicos de IA de alto riesgo habrán de cumplir, además, los deberes del RIA.....	323
11.3.- La extensión de estudios de impacto y la obligatoriedad de medidas para la no discriminación de los sistemas de IA públicos	325
11.4.- La necesidad de una visión social y colectiva y de integrar garantías de calidad de datos y de participación	329
12.- Y SI LOS REMEDIOS NO VALEN, ¿QUÉ PASA CUANDO SE DETECTA UNA DISCRIMINACIÓN ALGORÍTMICA PÚBLICA?	330
12.1.- Obligaciones de mitigación, rectificación, corrección e incluso cese de uso del sistema de IA público.....	330
12.2.- Nulidad, anulabilidad e inderogabilidad de los sistemas de IA público y sus efectos	332
12.3.- Los efectos de la discriminación algorítmica en el contexto de la contratación administrativa	334
12.4.- Responsabilidad por el uso de IA público.....	335
13.- ¿QUIÉNES PUEDEN REACCIONAR ANTE UNA DISCRIMINACIÓN ALGORÍTMICA Y QUÉ AUTORIDADES DEBEN ACTUAR?	336
BIBLIOGRAFÍA CITADA	338

Capítulo VIII

Las singularidades del tratamiento de datos de carácter personal en entornos de inteligencia artificial en el sector público

JULIÁN VALERO TORRIJOS

1.- EL USO DE DATOS PERSONALES POR LAS ADMINISTRACIONES PÚBLICAS Y LA INTELIGENCIA ARTIFICIAL.....	354
2.- EL CONTEXTO TECNOLÓGICO Y REGULATORIO.....	357
2.1.- La singularidad de la tecnología y su proyección sobre el objeto de estudio: el planteamiento del RGPD.....	357
2.2.- La regulación sobre las garantías y medidas de seguridad que deben adoptarse con relación al uso de medios electrónicos en la actividad administrativa.....	362
3.- LAS BASES JURÍDICAS PARA EL TRATAMIENTO DE LOS DATOS DE CARÁCTER PERSONAL.....	364
3.1.- Las limitaciones del consentimiento como base jurídica para el tratamiento de datos personales en el sector público.....	364

3.2.- Las singularidades de los datos de categoría especial: el caso de los datos biométricos	366
3.3.- El sentido y alcance de la habilitación legal para el tratamiento de los datos personales sin consentimiento: una moneda de dos caras	368
4.- LAS CONDICIONES PARA LA LICITUD DEL TRATAMIENTO.....	372
4.1.- El alcance de los principios generales y su proyección en los tratamientos basados en IA por parte de las Administraciones Públicas	372
4.1.1.- La singular eficacia del principio de transparencia en la normativa sobre protección de datos personales.....	374
4.1.2.- El alcance de las limitaciones impuestas por el principio de minimización	376
4.1.3.- La prohibición de usos incompatibles y la limitación del tiempo de conservación de los datos.....	377
4.2.- El efectivo cumplimiento de las normas de seguridad: la relevancia del análisis de riesgos y la evaluación de impacto	379
4.3.- La externalización en la gestión de las herramientas basadas en el uso de IA: las condiciones jurídicas en el encargo del tratamiento de datos personales	384
5.- LA INCIDENCIA DEL INCUMPLIMIENTO DE LA REGULACIÓN EN MATERIA DE PROTECCIÓN DE DATOS PERSONALES SOBRE LA VALIDEZ DE LA ACTUACIÓN ADMINISTRATIVA.....	388
6.- REFLEXIÓN FINAL	391
BIBLIOGRAFÍA Y OTRAS REFERENCIAS CITADAS.....	392

Capítulo IX

Las garantías de régimen jurídico del sector público y del procedimiento administrativo común frente a la actividad automatizada y la inteligencia artificial

EDUARDO GAMERO-CASADO

1.- INTRODUCCIÓN.....	398
1.1.- Acotación material: sistemas de IA para soporte a la toma de decisiones	398
1.2.- Las especialidades de la IA en el sector público y su sujeción al Derecho administrativo general sin necesidad de disposiciones que lo establezcan expresamente.....	400
2.- ACTIVIDAD ADMINISTRATIVA AUTOMATIZADA: PREVISIONES EN LA LEY 40/2015, DE RÉGIMEN JURÍDICO DEL SECTOR PÚBLICO	403
2.1.- El concepto de “actuación administrativa automatizada” y los grados de automatización	404

2.2.- La aprobación de la implantación del sistema	407
2.3.- El órgano competente para implementar los sistemas de IA y atribuirle su actividad. Firma electrónica de decisiones automatizadas	409
2.4.- La supervisión, control de calidad y auditoría del sistema y de su código fuente.....	411
2.4.1.- Encuadre general. El deber legal de implantar estos controles de la IA.....	411
2.4.2.- Control de calidad y auditoría del sistema de información y de su código fuente	413
2.4.3.- Supervisión <i>humana</i> . La necesidad de <i>compliance</i> jurídico	416
2.4.4.- Incumplimiento del requisito de mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema: invalidez de la actuación automatizada por vulneración del art.48.2 LPAC	421
3.- INTELIGENCIA ARTIFICIAL Y PROCEDIMIENTO ADMINISTRATIVO: GARANTÍAS DEL PROCEDIMIENTO ADMINISTRATIVO COMÚN	423
3.1.- Encuadre: la pérdida de derechos ante la administración digital	423
3.2.- El derecho a subsanar la presentación de escritos	427
3.3.- El derecho de audiencia	428
3.4.- La motivación de los actos administrativos adoptados mediante sistemas de IA.....	430
3.4.1.- La motivación como requisito de validez de los actos dictados mediante IA.....	430
3.4.2.- Alcance y medio de la motivación: reglas especiales en relación con la IA	433
3.4.3.- La omisión de motivación genera indefensión y determina la invalidez de los actos dictados con sistemas de IA. Jurisprudencia nacional, comparada y del TJUE.....	439
3.4.4.- Consecuencia: prohibición de algoritmos de caja negra en la toma de decisiones por el sector público. ¿Posibilidad de articular en el futuro una manera distinta de motivar, cuando se trata de decisiones sustentadas en IA?	445
3.5.- Inteligencia artificial y discrecionalidad administrativa.....	448
4.- FINAL: LA INACEPTABLE TENSIÓN ENTRE PROGRESO TECNOLÓGICO Y GARANTÍAS DE LA CIUDADANÍA; LUCES Y SOMBRAS DE LA REGULACIÓN ESPECÍFICA DE LA IA EN EL SECTOR PÚBLICO.....	452
BIBLIOGRAFÍA CITADA	455

TERCERA PARTE:
MEDIOS PARA LA IMPLEMENTACIÓN DE LA INTELIGENCIA
ARTIFICIALE EN EL SECTOR PÚBLICO

Capítulo X

Reutilización de la información del sector público e inteligencia artificial

AGUSTÍ CERRILLO I MARTÍNEZ

1.- DATOS E INTELIGENCIA ARTIFICIAL	466
2.- EL IMPULSO EUROPEO DE LA REUTILIZACIÓN DE DATOS PÚBLICOS Y SU CONTRIBUCIÓN AL DESARROLLO DE LA INTELIGENCIA ARTIFICIAL.....	469
3.- LA REUTILIZACIÓN DE DATOS PARA EL DESARROLLO DE LA INTELIGENCIA ARTIFICIAL	475
3.1.- La reutilización de datos públicos.....	475
3.1.1.- La apertura de datos públicos.....	475
3.1.2.- La reutilización de determinadas categorías de datos públicos protegidos por la confidencialidad.....	483
3.2.- El intercambio de datos.....	489
3.2.1.- Los servicios de intermediación de datos	489
3.1.2.- La cesión altruista de datos	492
3.3.- La obtención de datos necesarios para el desempeño de las tareas realizadas en interés público	494
4.- REFLEXIONES FINALES: DISPONIBILIDAD Y CALIDAD DE LOS DATOS COMO FUNDAMENTO DEL DESARROLLO DE LA INTELIGENCIA ARTIFICIAL	498
BIBLIOGRAFÍA CITADA	500

Capítulo XI

La contratación de soluciones de inteligencia artificial

ISABEL GALLEGO CÓRCOLES

1.- LA CONTRATACIÓN PÚBLICA COMO IMPULSOR Y GARANTE DEL USO DE SOLUCIONES BASADAS EN INTELIGENCIA ARTIFICIAL	504
2.- CUESTIONES CLAVE EN EL DISEÑO Y PREPARACIÓN DEL PROCEDIMIENTO DE CONTRATACIÓN	511
2.1.- La conformación del equipo que impulsará la licitación	511
2.2.- La determinación de las necesidades a satisfacer y la delimitación del objeto del contrato: ¿cuándo adquirir una solución de inteligencia artificial?	513
2.2.1.- Planteamiento.....	513

2.2.2.- El conocimiento del mercado: la posibilidad de realizar consultas preliminares	515
2.2.3.- Evaluación preliminar de los riesgos.....	517
2.2.4.- ¿Existen datos de calidad?	518
2.3.- La elección del procedimiento de licitación.....	519
2.4.- Algunos aspectos a valorar en las ofertas.....	522
2.5.- La adecuada delimitación del equipo de trabajo: la necesaria incorporación de juristas	523
2.6.- La necesidad de prever posibles incidencias en la ejecución	524
3.- ASPECTOS RELEVANTES DEL CONTENIDO DEL CONTRATO ..	527
3.1.- La necesidad de cumplir las exigencias de transparencia, motivación y libre competencia: ¿debe exigirse la cesión de la titularidad del código fuente?.....	527
3.1.1.- Planteamiento.....	527
3.1.2.- El eventual acceso al código fuente en las Guías sobre contratación de soluciones de inteligencia artificial	530
3.1.3.- El principio de competencia y entrega del código fuente..	538
3.1.4.- La regulación de la eventual cesión de derechos de propiedad intelectual en la LCSP	544
3.2.- Mecanismos para garantizar la supervisión, auditoria y control de los sistemas algorítmicos.....	547
3.2.1.- Planteamiento.....	547
3.2.2.- Supervisión humana de los algoritmos.....	548
3.2.3.- Control y auditoría del sistema.....	552
3.3.- Cláusulas relativas a la protección de datos.....	559
4.- CONCLUSIONES	562
BIBLIOGRAFÍA CITADA	564

Capítulo XII

Empleo público e inteligencia artificial

RAMON GALINDO CALDÉS

1.- TRES GRADOS, DIFERENTES AFECTACIONES.....	570
1.1.- El impacto de la automatización en el empleo público: destrucción y redefinición de puestos y tareas	570
1.2.- El uso de datos en la toma de decisiones.....	573
1.3.- La sustitución del empleado en la toma de decisiones.....	574
1.3.1.- La sustitución de las decisiones por un sistema de IA	574
1.3.2.- ¿Hay una reserva de intervención de los empleados públicos?	575
2.- ¿HACIA UNA NUEVA SEGMENTACIÓN DE LOS EMPLEADOS PÚBLICOS?.....	578
2.1.- Un impacto a diferentes niveles	579

2.1.1.- Redefinición de puestos y tareas	579
2.1.2.- Cambios necesarios en los empleados actuales	580
2.1.3.- Nuevos empleados públicos.....	581
2.2.- La contratación pública: ¿personal especializado o recurso al sector privado?	582
2.3.- ¿Dos clases de empleados? Si no es posible transformar a todos los empleados.... ¿Lo hacemos con una parte?.....	584
3.- CUATRO PRESUPUESTOS NECESARIOS PARA LA ADAPTACIÓN DEL EMPLEO PÚBLICO	586
3.1.- Planificación	587
3.2.- Selección	589
3.3.- Transformación	590
3.4.- Los datos en la gestión de los recursos humanos	591
BIBLIOGRAFÍA CITADA	593

CUARTA PARTE: ORGANIZACIÓN Y GOBERNANZA

Capítulo XIII

Organización y gobernanza de la inteligencia artificial: marco general

JUAN CARLOS HERNÁNDEZ PEÑA

1.- EL MARCO DE LA GOBERNANZA DE LA IA: UN TRINOMIO NORMATIVO BAJO UN ENFOQUE BASADO EN RIESGOS Y DE-RECHOS.....	600
2.- ALGUNOS PRINCIPIOS QUE CONDICIONAN LA ACTUACIÓN DE LOS PODERES PÚBLICOS Y EL MODELO DE GOBERNAN- ZA.....	603
2.1.- Principio de neutralidad tecnológica.....	604
2.2.- Principio de innovación	605
2.3.- Principio de precaución	607
2.4.- Principio de competencia por diseño	608
4.- ESTRUCTURAS DE GOBERNANZA DE LA IA	612
4.1.- El Comité Europeo de Inteligencia Artificial. Composición, competencias y funciones.....	613
4.2.- Autoridades nacionales. Designación y competencias	615
4.3.- Organismos notificados y organismos de evaluación de la con- formidad.....	624
BIBLIOGRAFÍA CITADA	627

*Capítulo XIV***Organización y gobernanza de la inteligencia artificial en la Junta de Andalucía:
la Estrategia Andaluza de IA y la Agencia Digital de Andalucía**

MARÍA PÉREZ NARANJO

1.- LA AGENCIA DIGITAL DE ANDALUCÍA	632
1.1.- Contextualización	632
1.2.- Un nuevo modelo de gestión TIC	633
1.3.- Plan inicial de actuación.....	634
1.3.1.- Líneas y objetivos estratégicos	635
1.3.2.- Siete planes estratégicos.....	637
2.- ESTRATEGIA DE INTELIGENCIA ARTIFICIAL DE ANDALUCÍA .	639
2.1.- La inteligencia artificial en Andalucía	642
2.2.- Misión y visión de la Estrategia	644
2.3.- Comité de expertos.....	644
2.4.- Metodología.....	646
2.5.- Modelo de gobernanza	648

*Capítulo XV***Organización y gobernanza de la Inteligencia Artificial en las entidades locales**

M. ASCENSIÓN MORO CORDERO

1.- INTRODUCCIÓN.....	652
2.- EL RETO DE LA TRANSFORMACIÓN DIGITAL DE LA ADMINISTRACIÓN LOCAL ¿UN PROCESO AD INFINITUM?	656
3.- LA NECESARIA GOBERNANZA MULTINIVEL Y MULTIDIMENSIONAL	662
3.1.- Un modelo de gestión inteligente.....	666
3.2.- Un modelo de gobernanza de los datos.....	669
3.3.- Una metodología	676
4.- POSIBLES USOS DE LA INTELIGENCIA ARTIFICIAL Y DESAFÍOS	677
5.- CONCLUSIONES	680
BIBLIOGRAFÍA CITADA	681

QUINTA PARTE:
DERECHO EUROPEO Y PROPUESTAS REGULATORIAS DE
FUTURO

Capítulo XVI

**Algoritmos, inteligencia artificial y procedimiento administrativo: principios
comunes en el Derecho de la Unión Europea**

ORIOl MIR PUIGPELAT

1.- INTRODUCCIÓN.....	686
2.- EL PAPEL CENTRAL DEL PROCEDIMIENTO ADMINISTRATIVO EN LA ADMINISTRACIÓN ANALÓGICA Y DIGITAL.....	688
2.1.- Actos administrativos	688
2.2.- Normas reglamentarias	691
3.- DISTINCIONES RELEVANTES RELATIVAS AL USO DE ALGO- RITMOS EN EL MARCO DEL PROCEDIMIENTO ADMINISTRA- TIVO.....	692
3.1.- Distinción en función de la fase del procedimiento en que se emplea el algoritmo y de su grado de incidencia sobre la deci- sión final.....	692
3.2.- Distinción en función del tipo de algoritmo empleado.....	694
3.3.- Distinción en función del carácter favorable o desfavorable de la decisión automatizada	696
3.4.- Distinción en función del carácter discrecional o reglado de la decisión automatizada	697
4.- ADAPTACIONES PROCEDIMENTALES NECESARIAS EN CASO DE UTILIZACIÓN DE ALGORITMOS	698
4.1.- La necesidad de informar adecuadamente a los interesados y al público sobre los sistemas automatizados empleados.....	699
4.2.- El establecimiento de un principio de supervisión humana	701
4.3.- La necesidad de realizar evaluaciones de impacto antes y des- pués de automatizar la toma de decisiones administrativas.....	704
5.- ANEXO: EL USO DE IA POR PARTE DE LA ADMINISTRACIÓN DE LA UNIÓN: UN ESTUDIO DE CAMPO	709
5.1.- Finalidad del estudio y casos examinados.....	709
5.1.1.- Caso 1 (DG-Agri/ESA)	712
5.1.2.- Caso 2 (EFSA)	713
5.1.3.- Caso 3 (EUIPO).....	715
5.1.4.- Caso 4 (eu-LISA).....	716
5.2.- Algunas conclusiones que cabe extraer del estudio de campo realizado.....	720
BIBLIOGRAFÍA CITADA	722

*Capítulo XVII***Hacia un marco regulatorio holístico de la Inteligencia Artificial fiable en la Unión Europea: una perspectiva científico-técnica**

EMILIA GÓMEZ, ISABELLE HUPONT TORRES, IGNACIO SÁNCHEZ, DAVID FERNÁNDEZ LLORCA

1.- EL ENFOQUE EUROPEO PARA LA INTELIGENCIA ARTIFICIAL	728
2.- CARACTERÍSTICAS DE LOS SISTEMAS DE INTELIGENCIA ARTIFICIAL.....	730
3.- LA LEY DE INTELIGENCIA ARTIFICIAL (AI ACT)	732
3.1.- Definición de IA	732
3.2.- Niveles de riesgo y sus requisitos	733
4.- LA DIRECTIVA SOBRE RESPONSABILIDAD EN MATERIA DE INTELIGENCIA ARTIFICIAL (<i>AI LIABILITY DIRECTIVE</i>)	736
5.- LA REVISIÓN DE LA DIRECTIVA DE RESPONSABILIDAD POR PRODUCTOS DEFECTUOSOS	738
6.- CONCLUSIONES	739
BIBLIOGRAFÍA CITADA	740

*Capítulo XVIII***Hacia la regulación europea de la inteligencia artificial**

ALEJANDRO HUERGO LORA

1.- INTRODUCCIÓN.....	743
2.- DE QUÉ ESTAMOS HABLANDO	745
3.- LA IMPORTANCIA DEL CONTEXTO.....	749
4.- SESGOS, ERRORES Y LÍMITES DE LAS PREDICCIONES ALGORÍTMICAS	755
5.- EL ENCAJE JURÍDICO DE LAS PREDICCIONES ALGORÍTMICAS: ALGUNAS TENDENCIAS DEL REGLAMENTO EUROPEO DE INTELIGENCIA ARTIFICIAL	759

*Capítulo XIX***Inteligencia artificial y reconocimiento biométrico en el procedimiento administrativo**

IGNACIO ALAMILLO DOMINGO

1.- INTRODUCCIÓN.....	764
2.- PRECISIONES TERMINOLÓGICAS PREVIAS.....	765
2.1.- Precisiones desde la perspectiva tecnológica	765
2.2.- Precisiones desde la perspectiva jurídica	769

3.- ALGUNOS CASOS DE USO DEL RECONOCIMIENTO BIOMÉTRICO BASADO EN INTELIGENCIA ARTIFICIAL EN EL PROCEDIMIENTO ADMINISTRATIVO	775
3.1.- Reconocimiento biométrico en la asistencia por canal telefónico	776
3.2.- Reconocimiento biométrico en la asistencia por videoconferencia.....	778
3.3.- La captura electrónica de firma manuscrita	779
4.- PROPUESTAS PARA EL USO DEL RECONOCIMIENTO BIOMÉTRICO BASADO EN INTELIGENCIA ARTIFICIAL EN EL PROCEDIMIENTO ADMINISTRATIVO	780
BIBLIOGRAFÍA CITADA	784

Abreviaturas y acrónimos

AA:	<i>Actualidad Administrativa</i>
AAPP:	Administraciones públicas
ADA:	Agencia Digital de Andalucía
AEAD:	Agencia Estatal de Administración Digital
AEPD:	Agencia Española de Protección de Datos
AESIA:	Agencia Española de Supervisión de la Inteligencia Artificial
AGE:	Administración General del Estado
AI-HLEG:	Grupo de Alto Nivel sobre Inteligencia Artificial de la Comisión Europea
art./arts.:	artículo/artículos
BOE:	Boletín Oficial del Estado
CCAA:	Comunidades Autónomas
CDFUE:	Carta de los Derechos Fundamentales de la Unión Europea
CE:	Constitución Española de 27 de diciembre de 1978
CENDOJ:	Centro de Documentación Judicial
CEPD:	Comité Europeo de Protección de Datos
cfr.:	confróntese
D:	Decreto
DA:	Disposición adicional
DAIA:	Derecho Administrativo e Inteligencia Artificial (Red).
DDI:	<i>Derecho Digital e Innovación</i> (revista)
DEHU:	Dirección Electrónica Habilitada Única
DF:	Disposición final
DL:	<i>Deep learning</i>
DNI:	Documento Nacional de Identidad
DOUE:	Diario Oficial de la Unión Europea
DPD (o DPO):	Delegado de protección de datos

DT:	Disposición transitoria
EA:	Estatuto de Autonomía
EDPS:	Supervisor Europeo de Protección de Datos
EELL:	Entidades locales
EIT:	Electrónico, Informático y Telemático
ENI:	Real Decreto 4/2010, de 8 de enero, por el que se aprueba el Esquema Nacional de Interoperabilidad
ENIA:	Estrategia Nacional de Inteligencia Artificial
ENS:	Real Decreto 311/2022, de 3 de mayo, por el que se regula el Esquema Nacional de Seguridad
ERDAL:	<i>European Review of Digital Administration & Law</i>
FNMT:	Fábrica Nacional de y Timbre-Real Casa de
f.j./ff.jj.:	fundamento jurídico/fundamentos jurídicos
IA:	inteligencia artificial
IDP:	<i>Revista de Internet, Derecho y Política</i>
IoT:	<i>Internet of things</i> (o Internet de las cosas)
LAECSP:	Ley 11/2007, de 22 de junio, de Acceso Electrónico de los Ciudadanos a los Servicios Públicos
LEC:	Ley 1/2000, de 7 de enero, de Enjuiciamiento Civil
LFE:	Ley 59/2003, de 19 de diciembre, de Firma Electrónica
LG:	Ley 50/1997, de 27 de noviembre, del Gobierno
LGT:	Ley 58/2003, de 17 de diciembre, General Tributaria
LGTel:	Ley 9/2014, de 9 de mayo, General de Telecomunicaciones
LJCA:	Ley 29/1998, de 13 de julio, Reguladora de la Jurisdicción Contencioso-administrativa
LOFAGE:	Ley 6/1997, de 14 de abril, de Organización y Funcionamiento de la Administración General del Estado
LOPD:	Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales
LOPJ:	Ley Orgánica 6/1985, de 1 de julio, del Poder Judicial

LOTG:	Ley Orgánica 2/1979, de 3 de octubre, del Tribunal Constitucional
LPAC:	Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas
LRBRL:	Ley 7/1985, de 2 de abril, Reguladora de las Bases del Régimen Local.
LRISP:	Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público
LRJPAC:	Ley 30/1992, de 26 de noviembre, del Régimen Jurídico de las Administraciones Públicas y del Procedimiento Administrativo Común
LRJSP:	Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público
LSSI:	Ley 34/2002, de 11 de julio, de servicios de la sociedad de la información y de comercio electrónico
LTBG:	Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno
LUTICAJ:	Ley 18/2011, de 5 de julio, reguladora del uso de las tecnologías de la información y la comunicación en la Administración de Justicia
ML:	<i>Machine learning</i>
n.º:	número (de revista, etc.)
NTI:	Norma técnica de interoperabilidad.
<i>op.cit.</i> :	obra citada
p./pp.:	página/páginas
RAAP:	<i>Revista Andaluza de Administración Pública. Administración de Andalucía</i>
RADNT:	<i>Revista Aranzadi de Derecho y Nuevas Tecnologías</i>
RAFESP:	Real Decreto 203/2021, de 30 de marzo, por el que se aprueba el Reglamento de actuación y de funcionamiento electrónico del sector público
RArAP:	<i>Revista Aragonesa de Administración Pública</i>
RAP:	<i>Revista de Administración Pública</i>
RCDP:	<i>Revista Catalana de Dret Públic</i>

REDETI:	<i>Revista del Derecho de las Telecomunicaciones e Infraestructuras en Red</i>
RD:	Real Decreto
RDL:	Real Decreto-ley
RDLeg:	Real Decreto Legislativo
REDA:	<i>Revista Española de Derecho Administrativo</i>
RGAP:	<i>Revista Galega de Administración Pública</i>
RGD:	Reglamento de Gobernanza de Datos. Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos y por el que se modifica el Reglamento (UE) 2018/1724
RGDA:	<i>Revista General de Derecho Administrativo</i>
RGPD:	Reglamento General de Protección de Datos; Reglamento (UE) 2016/679, del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE
ReIDAS:	Reglamento (UE) núm. 910/2014, del Parlamento Europeo y del Consejo, de 23 de julio de 2014, relativo a la identificación electrónica y los servicios de confianza para las transacciones electrónicas en el mercado interior y por la que se deroga la Directiva 1999/93/CE
RIA:	Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial).
RN:	Red neuronal
RVAP:	<i>Revista Vasca de Administración Pública</i>
SAN (Sala):	Sentencia de la Audiencia Nacional
SAP-Ciudad:	Sentencia de Audiencia Provincial-Ciudad sede
SEDIA:	Secretaría de Estado de Digitalización e Inteligencia Artificial

SJCCA:	Sentencia de Juzgado Central de lo Contencioso-administrativo
SJCA-Ciudad:	Sentencia de Juzgado de lo Contencioso-administrativo-Ciudad sede
STC:	Sentencia del Tribunal Constitucional
STEDH:	Sentencia del Tribunal Europeo de los Derechos Humanos
STJUE:	Sentencia del Tribunal de Justicia de la Unión Europea
STS (Sala):	Sentencia del Tribunal Supremo (Cardinal de la Sala sentenciadora)
ss.:	siguientes (en citas de obras; p.ej., pp.57 ss.)
STSJ-Ciudad:	Sentencia de Tribunal Superior de Justicia-Ciudad sede de la Sala sentenciadora
TFUE:	Tratado de Funcionamiento de la Unión Europea, de 25 de marzo de 1957, ratificado por Instrumento de 13 de diciembre de 2007.
TICs:	Tecnologías de Información y Comunicaciones
TJUE:	Tribunal de Justicia de la Unión Europea
Tol.:	Referencia marginal del repertorio <i>Tirant on line</i>
TREBEP:	Real Decreto Legislativo 5/2015, de 30 de octubre, por el que se aprueba el Texto Refundido del Estatuto Básico del Empleado Público
TUE:	Tratado de la Unión Europea firmado en Maastricht el 7 de febrero de 1992
UE:	Unión Europea
<i>Vid.</i> :	véase

Prefacio

**EDUARDO GAMERO CASADO,
FRANCISCO L. PÉREZ GUERRERO**

El principal objetivo de esta monografía es analizar los retos, límites y medios que entraña la utilización de la inteligencia artificial en el sector público, especialmente cuando se usa como soporte a la toma de decisiones.

La IA es una realidad cotidiana que afecta a todos los ámbitos. Supone un enorme potencial de mejora de los servicios públicos y de progreso económico. Su desarrollo es una apuesta estratégica en todos los países avanzados: como magnitudes significativas, cabe resaltar que la UNESCO estima que supondrá inversiones por importe de cuatro billones de euros durante esta década; que en ese mismo período la Comisión Europea estima en 10.000 millones de euros al año la inversión en esta tecnología solo en los países de la Unión; o que, según los informes disponibles, ya en 2022 la tercera parte del sector público español utilizaba sistemas de IA.

Sin embargo, no existen normas que regulen específicamente su uso, lo cual genera una gran inseguridad, tanto para la Administración, como para las empresas que están acometiendo el desarrollo de las soluciones. Lo cual no significa que la IA se encuentre al margen del Derecho, sino que es necesario inferir su marco jurídico del conjunto del Ordenamiento jurídico, como sucede con cualquier otra actividad de nueva aparición en el escenario socio económico, y, concretamente, en el ámbito tecnológico.

En ese contexto, el propósito de esta monografía es doble. En primer lugar, se aspira a compilar el conocimiento existente, sistematizando las aportaciones que han tenido lugar hasta hoy, ensamblándolas en una obra que pretende abordar todo el espectro de cuestiones jurídicas que entraña el uso de la IA en el sector público. En segundo lugar, se quiere abordar el estudio de cuestiones no tratadas hasta ahora y alumbrar aportaciones novedosas a la luz de la madurez que vamos alcanzando en esta materia.

La obra se estructura en cinco grandes partes. En la Primera de ellas se explica qué es la IA, es decir, en qué consisten las tecnologías que se agrupan bajo esa etiqueta. Aunque no exista univocidad científica en la cuestión, se trata de la premisa sobre la que debe descansar todo el análisis posterior, pues para pronunciarnos acerca del régimen jurídico de una determinada actividad –especialmente, cuando no tiene una normativa o regulación específica–, lo primero que debemos es delimitarla, conocer en qué consiste esa actividad, para poder proyectar así sobre ella el conjunto del Ordenamiento jurídico. Esta Primera parte se compone de dos relevantes aportaciones, elaboradas por especialistas en la materia: un capítulo de perfil tecnológico, en el que ingenieros de sistemas explican qué es la IA; y un capítulo elaborado por investigadores del JRC, en el que se ofrece un amplio muestrario de utilidades y casos de uso de sistemas de IA.

Una vez encuadrado y acotado de este modo el ámbito material sobre el que trabajamos, la Segunda parte analiza el marco jurídico general de la IA, articulándose a su vez internamente en diferentes secciones. De este modo, se van abordando los diferentes aspectos que preocupan a las Ciencias Jurídicas en la cuestión: tanto la naturaleza jurídica de los algoritmos, como los derechos fundamentales implicados en el uso de la IA (igualdad y no discriminación, protección de datos, garantías del procedimiento y a no padecer indefensión), la transparencia en la actuación de los poderes públicos... Además de estas cuestiones, sobre las que ya se disponía de una cierta literatura que aquí se sistematiza y actualiza con sugerentes novedades, también se ofrecen temas que, hasta donde se nos alcanza, se tratan por primera vez, como la responsabilidad patrimonial por el uso de la IA en el sector público.

En la Tercera parte de la obra se analizan los medios de que debe servirse el sector público para implantar sistemas de IA. Contiene capítulos de temática innovadora en esta cuestión: la conexión entre los sistemas de IA y la reutilización de la información del sector público; las particulares determinaciones a seguir en la contratación pública de desarrollos de IA; o las implicaciones de la IA en el empleo público.

La Cuarta parte también se relaciona con los medios, pero en este caso la temática específica es la organización y gobernanza de la IA. Además

de un capítulo consagrado al encuadre general de la cuestión, se ofrecen aportaciones y experiencias específicas en el ámbito autonómico y local.

Por último, la Quinta parte emprende un análisis de Derecho comparado –especialmente el europeo– y las propuestas regulatorias de futuro, con especial atención al Reglamento de IA que en estos momentos elabora la Unión Europea. Constituye una novedad de enorme impacto a escala mundial, y en esta obra se analiza desde diferentes perspectivas que integran tanto la explicación del enfoque y alcance de la normativa por parte de quienes han estado muy activos en su gestación, como propuestas y comentarios desde el ámbito académico.

La metodología desarrollada para elaborar esta obra ha consistido en la celebración de unas (intensas) jornadas en Sevilla, en el marco del Proyecto de Investigación UPO-1381574, “Inteligencia artificial y Derecho Administrativo”, en el que además de los coordinadores de esta obra, se integran también la Prof.^a Rocío Navarro González y el Prof. Antonio Berning Prieto. Con ocasión de las jornadas los autores de los diferentes capítulos fueron exponiendo una versión provisional de sus trabajos y sometiéndolos a fructíferos coloquios, lo que ha permitido enriquecer las versiones finalmente aportadas a la obra.

Tanto para la celebración de las jornadas, como para la estructuración y publicación de esta obra, ha resultado decisivo el concurso de diferentes actores.

En primer lugar, de la Red Derecho Administrativo e Inteligencia Artificial, en la que nos integramos especialistas que llevamos colaborando más de veinte años en el estudio del Derecho público digital. En el seno de la red se gestó esta iniciativa y su respaldo ha sido crucial para llevarla a buen puerto. Debemos mencionar especialmente a Agustí Cerrillo Martínez, Lorenzo Cotino Hueso, Isaac Martín Delgado, Juli Ponce Solé, Julián Valero Torrijos y Clara Velasco Rico. Nuestra gratitud para con su labor es inmensa. Gurumelos. Las actividades de la Red tendrán continuidad en el marco de otras iniciativas, como el proyecto “Transición digital de las Administraciones públicas e inteligencia artificial: claves, riesgos y propuestas de política regulatoria (DIGITAPIA)”, concedido por el Ministerio en el marco de la convocatoria de Proyectos estratégicos orientados a la transición ecológica y a la transición digital, del Plan Estatal de Investigación

Científica, Técnica y de Innovación para el período 2021-2023", cuyos IPs son Agustí Cerrillo y Clara Velasco.

Debemos agradecer igualmente su apoyo a quienes contribuyeron con medios humanos y económicos a la celebración de las jornadas: la Universidad Pablo de Olavide, concedente del proyecto de investigación en el marco de una convocatoria competitiva autonómica; la Fundación Cajazol, que acogió las jornadas en sus excelentes instalaciones, y el Instituto de Estudios Cajazol, que facilitó toda la logística para su celebración: inscripciones, retransmisión en *streaming*, etc.; la Secretaría General de Administración Pública de la Junta de Andalucía, contraparte del proyecto de investigación en el que se integran estas jornadas; y Montero-Aramburu Abogados, con su respaldo económico a la celebración de la actividad.

Esperamos que esta monografía cumpla la función que se le pretende: servir de referencia para todos los operadores jurídicos en cuya labor presente interés la aplicación de la IA en el sector público, ya sea como objeto de estudio científico o debido a los problemas de orden práctico que plantea su uso. De hecho, la aproximación a las diferentes cuestiones tiene carácter intemporal, puesto que se trata, como hemos visto, de reglas y principios generales del Ordenamiento jurídico.

Esperamos, por tanto, que, a pesar de la volatilidad de las tecnologías, los contenidos y conclusiones de esta obra sean perdurables y útiles para cuantos acudan a ella.

Sevilla, diciembre de 2022.

PRIMERA PARTE:
LA INTELIGENCIA ARTIFICIAL Y
SUS APLICACIONES EN EL SECTOR
PÚBLICO

Capítulo I

¿Qué es la Inteligencia Artificial?

FRANCISCO GONZÁLEZ CABANES¹

*Miembro del Grupo de Investigación DATAi
Universidad Pablo de Olavide, de Sevilla*

NORBERTO DÍAZ DÍAZ²

*Profesor Titular de Lenguajes y Sistemas Informáticos
Universidad Pablo de Olavide, de Sevilla*

Resumen: A lo largo de la historia se ha investigado y hecho avances de forma recurrente con el objeto de imitar las habilidades humanas y optimizar procesos automatizándolos y disminuyendo la intervención humana. Muchos de estos avances están en relación con la Inteligencia Artificial, algunos lo son, y otros muchos, no.

En este primer capítulo haremos una aproximación a la Inteligencia Artificial, viendo qué es y qué no es, cómo los algoritmos de aprendizaje automático extraen conocimiento de nuestros datos en un proceso que se conoce integralmente como Knowledge Discovery Data (KDD), los tipos de aprendizaje, los algoritmos de aprendizaje clásicos y modernos y cómo la revolución del dato auspiciada por la evolución tecnológica ha influido en todo ello, y lo desarrollaremos todo con un enfoque cronológico y de evolución histórica, haciendo aproximaciones lo suficientemente técnicas para conocer y entender las técnicas que se aplican, pero sin entrar en profundidad.

Abstract: Throughout history, research and significant advances have been made on a recurring basis with the aim of imitating human skills and optimizing processes by automating them and decreasing human intervention. Many of these advances are in relation to artificial intelligence, some are, and many others are not.

In this first chapter we will make an approach to Artificial Intelligence, seeing that it is and is not, how machine learning algorithms extract knowledge from our data in a process that is integrally known as Knowledge Discovery Data (KDD), the learning types, the classical and modern algorithms and how the data revolution driven by the leverage of technological evolution has influenced all this. We will develop everything with a chronological approach and historical evolution, making sufficiently technical approaches to know and understand the techniques that are applied, but without going deep.

¹ ORCID: <https://orcid.org/0000-0001-8100-8781>

² ORCID: <https://orcid.org/0000-0003-0359-6092>

Palabras-clave: inteligencia artificial, aprendizaje automático, aprendizaje profundo, big data, redes neuronales, redes neuronales profundas, neurona.

Key-words: artificial intelligence, machine learning, deep learning, big data, neural networks, deep neural networks, neuron.

SUMARIO.- 1.- INTRODUCCIÓN A LA INTELIGENCIA ARTIFICIAL-1.1.- Conceptos Previos y Aclaraciones-1.2.- Historia-1.3.- Casos de Uso de la Inteligencia Artificial-2.- EL PROCESO DE EXTRACCIÓN DE CONOCIMIENTO: KDD-2.1.- Mecanismos de Aprendizaje. 2.2.- Generalización del Conocimiento-3.- ALGORITMOS CLÁSICOS DE APRENDIZAJE-3.1.- Aprendizaje Supervisado - 3.1.1. Árboles de Decisión 3.2.- Aprendizaje No Supervisado- 3.2.1- Clustering 3.2.2- Biclustering-3.3.- Redes Neuronales-4.- LA REVOLUCIÓN DIGITAL Y DEL DATO.-4.1.- La Importancia de los Datos-5.- ALGORITMOS MODERNOS DE APRENDIZAJE: DEEP LEARNING-5.1.- Redes Neuronales Profundas: Deep Learning-5.2.- Diferencias entre Machine Learning y Deep Learning-5.3-Transversalidad del Deep Learning-5.4.- Transfer Learning.-6.- EL PRESENTE Y FUTURO DE LA INTELIGENCIA ARTIFICIAL.-7.- CONCLUSIONES.-8.- BIBLIOGRAFÍA CITADA

1.- INTRODUCCIÓN A LA INTELIGENCIA ARTIFICIAL

A lo largo de la historia se ha investigado y hecho avances de forma recurrente con el objeto de imitar las habilidades humanas y optimizar procesos automatizándolos y disminuyendo la intervención humana. Muchos de estos avances están en relación con la Inteligencia Artificial, en adelante IA, algunos lo son, y otros muchos, no.

Comenzaremos este libro introduciendo la IA, su definición directa e indirecta, es decir, aclarando posibles errores de concepto adquiridos de forma involuntaria, también veremos el proceso que se conoce como Knowledge Discovery Data (KDD) mediante el cual los algoritmos de aprendizaje automático extraen conocimiento de nuestros datos, cómo se dieron las condiciones necesarias para provocar la Revolución del Dato, y lo desarrollaremos todo de forma cronológica y poniendo el foco en el dato, al que consideramos actor principal de este capítulo y haciendo aproximaciones para conocer y entender las técnicas que se aplican, pero sin entrar en profundidad técnica.

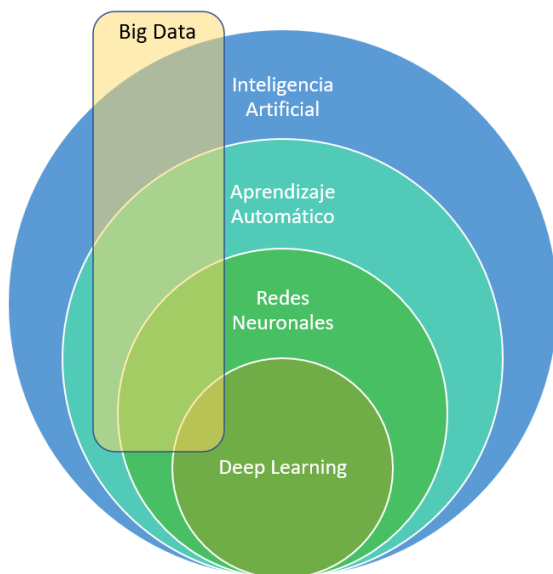
1.1.- *Conceptos Previos y Aclaraciones*

La IA vive un momento dulce de esplendor en todas sus fases, desde la generación y almacenamiento del dato, hasta la implementación

y puesta en producción de la solución, y la optimización y mantenimiento en producción de esta.

Este esplendor transversal queda desvirtuado por la espectacularidad de los avances que se están produciendo en algunos campos concretos dentro de la IA, como es el de las redes neuronales profundas. En este ámbito se están usando modelos con un gran número de parámetros que, por el momento, entenderemos como la forma en que las redes neuronales almacenan la información que aprenden. Estos avances en particular están generando expectación y difusión mediática por medios no especializados, lo cual está llevando a generar confusión en torno a una serie de conceptos básicos. A continuación, procederemos a aclarar dichos conceptos de forma escueta, empezando por una ilustración que plasma muy bien la relación entre todos ellos:

Fig. 1. Relación entre los Conceptos Clave.



Inteligencia Artificial

La IA es la disciplina que intenta replicar y desarrollar la inteligencia natural, de humanos y seres vivos, a través de computadoras o máquinas.

Cuando se centra de forma específica en la habilidad de ver, se le llama Computer Vision o Visión Artificial, cuando se centra en la habilidad del lenguaje con un enfoque hacia el texto, se llama Procesamiento del Lenguaje Natural (PLN), cuando se enfoca más al discurso o al lenguaje hablado, se llama Speech y cuando trata de emular los movimientos y desplazamiento se habla de robótica, si bien es cierto que la IA no se restringe a estos 4 campos.

Atendiendo a la definición de IA, muchas de las tareas asociadas a la computación o la informática están consideradas tareas inteligentes resultas por máquinas y por tanto IA, pero hoy en día son tan comunes y están tan integradas en nuestro día a día, que no las consideramos IA esperando que esta resuelva problemas más avanzados y complejos.

Un ejemplo de ello sería por ejemplo el filtro de spam de nuestro correo electrónico, el cual de forma automática decide qué correo es no deseado y cuál debe ser recibido en nuestra bandeja de entrada, pero también sería IA un problema matemático resuelto por una computadora a partir de unos datos concretos, como calcular el área de un polígono regular a partir de la longitud de sus lados.

Machine Learning

El Machine Learning, en adelante ML, es el área de estudio que busca emular en las máquinas *la capacidad de aprender* que tenemos humanos y seres vivos. Supone la automatización de la capacidad de aprender y es por tanto una disciplina dentro de la IA puesto que esta emula cualquier tipo de inteligencia natural, pero el ML, en concreto, emula la capacidad de aprender.

De una manera más técnica podríamos decir que es el conjunto de algoritmos que son capaces de aprender a partir de conjuntos de datos.

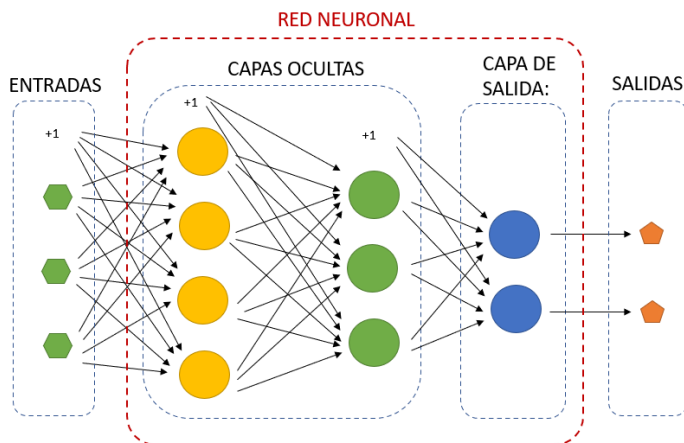
Llegados este punto y atendiendo a las definiciones de AI y ML, todas las aplicaciones del ML serán implícitamente aplicaciones de la IA, pero no ocurrirá lo contrario. El ML se centra en la creación de sistemas que aprenden en función de los datos que se le aportan, por lo tanto, cualquier sistema que realice algún tipo de automatización, que imite la inteligencia natural de alguna forma, será IA, pero sólo será ML si el sistema aprende de forma autónoma en base a la experiencia que adquiere procesando los datos que se le aportan.

Redes Neuronales

Las Redes Neuronales son una de esas familias de algoritmos dentro de la disciplina del ML. La peculiaridad del nombre se debe a ciertas analogías con las neuronas y el cerebro humano, si bien es cierto que estas analogías son más bien conceptuales que físicas.

Una representación abstracta típica de una red neuronal sería la que se muestra a continuación, donde tenemos unos datos de entrada X , unos datos de salida Y que han sido procesados por la red neuronal, cuyo conocimiento queda almacenado en unos parámetros W también llamados “pesos”. La red neuronal se compone de una capa de salida y un conjunto de capas ocultas, que pueden ir desde una única capa, a un alto valor:

Fig. 2. Representación Abstracta de una Red Neuronal.



Más adelante en el apartado 3.1.3.- Redes Neuronales daremos un mayor detalle sobre las características y el funcionamiento de las redes neuronales.

Deep Learning

El término Deep Learning o Aprendizaje Profundo, DL en adelante, hace referencia a las redes neuronales con una gran cantidad de capas ocultas. Si bien esta aproximación permite la generación de conocimiento en nuevos ámbitos, estos requieren una gran cantidad de recursos computacionales, así como de un gran volumen de datos de entrenamiento.

Big Data

Big Data es un concepto que hace referencia a conjuntos de datos tan grandes y complejos que precisan de aplicaciones informáticas no tradicionales para su procesamiento de forma adecuada.

En 2001, Doug Laney, introdujo el concepto de las 3 V's³ como características diferenciadores y que definían el Big Data:

- Volumen, en referencia al tratamiento de grandes volúmenes de datos.
- Variedad, en referencia a la heterogeneidad de dichos datos, en cuanto a tipos, orígenes, etc.
- Velocidad, como consecuencia de la rapidez con la que se generan dichos datos y con la que deben ser procesados.

Esta veloz producción de cantidades ingentes de datos heterogéneos y consumidos a gran velocidad, han generado nuevas necesidades de análisis y como consecuencia de ello, nuevas herramientas para procesarlos. Por todo ello, se ha pasado a usar también el término Big

³ Douglas, Laney (6 de febrero de 2001), “3D Data Management: Controlling Data Volume, Velocity and Variety”. Gartner. <https://studylib.net/doc/8647594/3d-data-management—controlling-data-volume—velocity—an...> (última visita: 30/08/22)

Data para denominar al conjunto de tecnologías creadas para almacenar, analizar y gestionar dichos volúmenes de datos.

En el apartado 4.- La Revolución Digital y del Dato volveremos a hablar sobre Big Data y en general, durante el desarrollo de este capítulo iremos retomando y profundizando sobre estos cinco conceptos: Inteligencia Artificial, Machine Learning, Redes Neuronales, Deep Learning y Big Data.

1.2.- Historia

Podríamos remontarnos a 1637 cuando el filósofo y matemático René Descartes se preguntaba si las máquinas podrían llegar a pensar, o incluso a la *Illiada*, de Homero, que data aproximadamente del s. VIII a.C. donde ya aparecían dos mujeres que parecían ser robots que pensaban, pero vamos a centrarnos en los hitos y periodos más destacados que han generado puntos de inflexión destacados en el desarrollo y uso de la IA.

Es interesante destacar lo que se conoce como los inviernos de la IA, que son periodos de tiempo en los que como resultado de un furor desmedido, debido a promesas poco realistas por parte de los desarrolladores, expectativas altas de los usuarios finales y una amplia publicidad en los medios, tras unos avances limitados en la materia, se produce una retirada de la inversión y financiación, dando lugar a periodos con avances muy escasos.

Historia de la IA



1950

- **1950: "Test de Turing"**, Alan Turing se plantea si puede pensar una máquina y publica su famoso test, el "Test de Turing" dentro de su trabajo "Computer machine and intelligence"[1], siendo uno de los primeros escritos serios que planteaba la posibilidad de una IA.
- **1956:** Conferencia de Dartmouth. En dicha conferencia, el profesor John McCarthy emplea por primera vez el término "**Inteligencia Artificial**" y se hicieron previsiones triunfalistas a diez años. Este punto se considera el inicio de la era de la Inteligencia Artificial.
- **1958:** Frank Rosenblatt diseña el **Perceptrón**, la primera red neuronal artificial.



Las previsiones que se hicieron sobre la IA no se cumplen y los avances limitados. Los grandes proyectos pierden financiación.

1970

- **1974 a 1980:** da lugar a lo que se conoce como "el **Primer Invierno de la Inteligencia Artificial**", auspiciado por detractores como Hubert Dreyfus que en 1972 publica el libro "Lo que las máquinas no pueden hacer". Aún así, la investigación dió ciertos frutos.
- En **1979**, la máquina BKG 9.8 derrota al campeón de backgammon y Ernst Dickmanns ya trabajaba en un modelo de **Visión Artificial**.



Se recupera el entusiasmo por la IA y por tanto la financiación destinada a la misma.

1980

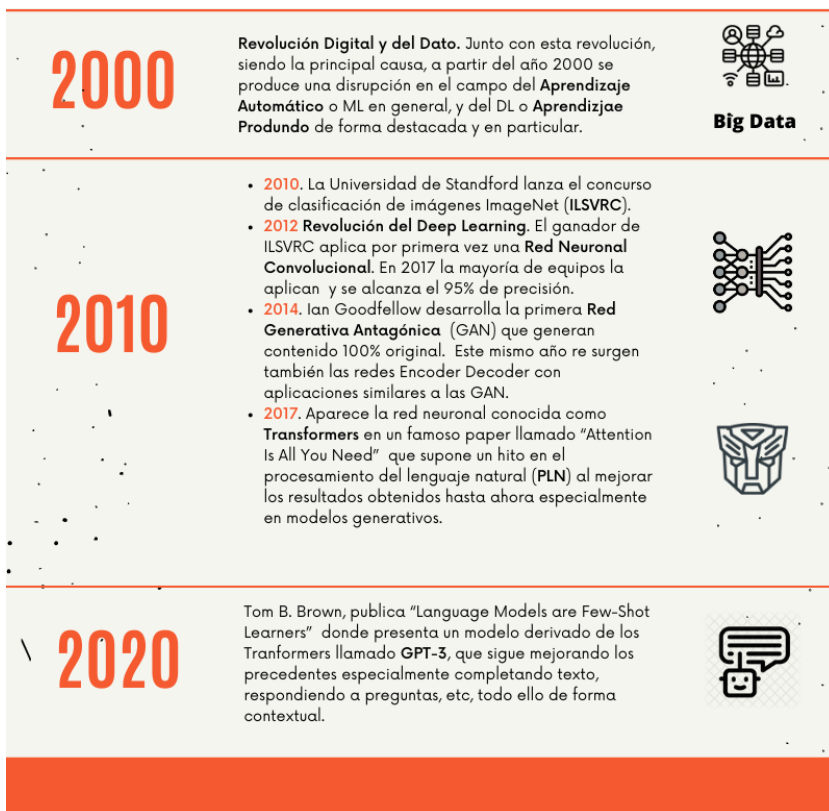
- **1982:** Dickmanns, primeros pasos en la **conducción autónoma**.
- **1987 a 1993:** **Segundo Invierno de la IA**, alcanzando su punto más bajo en 1990 y finalizando allá por 1993. Aunque hubo fracasos destacados, estuvo muy marcado por el del proyecto americano-japonés de las computadoras de 5ª generación, basadas en IA, cancelado en 1990. A pesar de ello se producen avances importantes en investigación, especialmente en **Visión Artificial**, disciplina que pasa a ser la punta de lanza de la IA.
- **1989:** Yanne Lecun desarrolla **Lenet**, la primera **Red Neuronal Convocucional (CNN)**, la cual muestra interesantes posibilidades para el procesamiento de imágenes.



1990

- **1993** Nuevo despegue de la IA
- **1997:** Deepblue vence a Kasparov.
- **1998:** Yanne Lecun aplica Lenet al reconocimiento de caracteres manuscritos, obteniendo una gran precisión. Lo refleja en su publicación "Gradient-Based Learning Applied to Document Recognition"





Una vez revisada la evolución histórica de la IA, incidiremos en que este capítulo se desarrollará en la medida de lo posible siguiendo un orden cronológico de los hechos.

1.3.- Casos de Uso de la Inteligencia Artificial

Veamos algunos usos de la IA, lo cual nos mostrara de paso lo integrada que está en nuestro día a día:

- Sistemas de recomendación en las compras online y en los sistemas de publicidad. Estos sistemas aprenden de nuestras búsquedas y consultas para recomendarnos productos que puedan ser de nuestro interés.

- Autocompletado de los sistemas de búsqueda web. El sistema aprende de las búsquedas que hacen los usuarios para predecir con bastante precisión lo que aún no has terminado de escribir. Estos sistemas se han incorporado ya a gestores de correo electrónico, como Gmail.
- Asistentes personales, como los que tenemos en nuestros smartphones, como el famoso Siri de Apple.
- Sistemas de traducción automática de textos, habiendo ya numerosas soluciones y de gran calidad, como DeepL⁴ o Reverso⁵
- Smarthomes, con dispositivos IoT (Internet of Things) que toman continuamente medidas para aprender de nuestros comportamientos, optimizar nuestros consumos y ofrecer consejos de uso, por ejemplo.
- Fabricación inteligente. Sistemas en gran parte apoyados en la Visión Artificial para por ejemplo usar robots que optimicen los recorridos en planta, hagan el “picking” de materiales de forma autónoma, o algo más clásico que lleva años en uso, que validen los productos que cumplen los estándares requeridos en la misma línea de producción.
- Mantenimiento predictivo. Sistemas que en base a históricos de datos recopilados, como medidas, realicen predicciones del mantenimiento necesario de la maquinaria para evitar o disminuir la tasa de fallo, alargando la vida útil de estas.
- Detección de fraude bancario. Analizando patrones de ingresos y gastos se puede predecir un fraude. Por ejemplo, si el gasto de una tarjeta de crédito aumenta anormalmente, entendiendo como tal, con un valor muy superior al patrón habitual, se puede predecir que ha sido sustraída y bloquearla automáticamente.

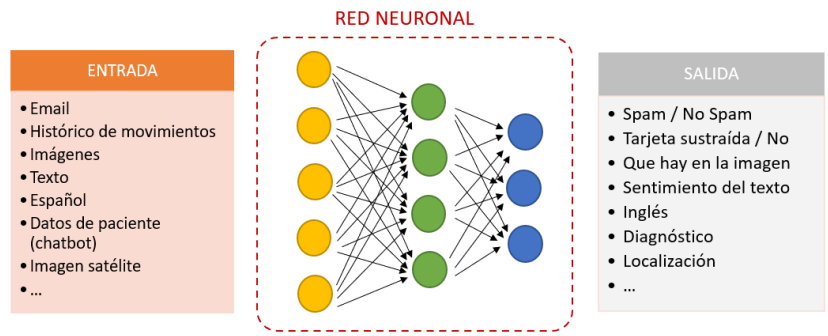
Llegados a este punto es importante señalar que, a día de hoy, en la inmensa mayoría de los casos, cuando hablamos de IA nos estamos refiriendo realmente a ML, lo mismo que ocurre con los casos de uso

⁴ <https://www.deepl.com/translator> (última visita: 29/08/2022)

⁵ <https://www.reverso.net/traduccion-texto> (última visita: 29/08/2022)

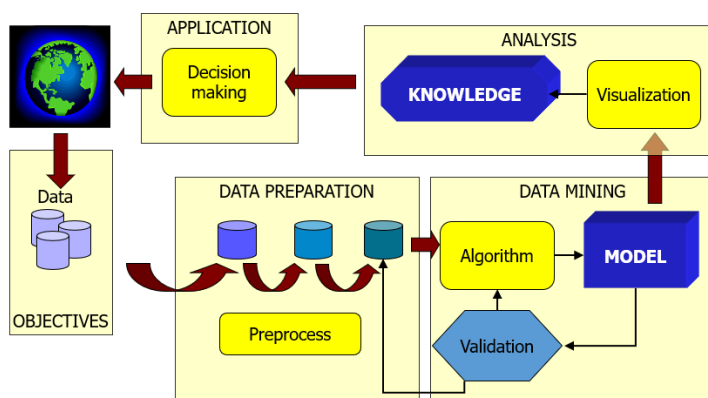
anteriormente mencionados, que son llevados a la práctica mayoritariamente por algoritmos de ML, y de forma más concreta, por redes neuronales.

Fig. 3. Casos de uso de IA a través de Redes Neuronales.



2.- EL PROCESO DE EXTRACCIÓN DE CONOCIMIENTO: KDD

El proceso global de búsqueda de nuevo conocimiento a partir de un conjunto de datos se denomina KDD (Knowledge Discovery in Data bases). Como se muestra en la figura 4, este proceso comprende diversas etapas, que van desde la obtención de los datos hasta la aplicación del conocimiento adquirido en la toma de decisiones.

Fig. 4. Proceso KDD⁶

Entre esas etapas, se encuentra la que puede considerarse como el núcleo del proceso KDD y que se denomina Minería de Datos o Data Mining (DM), etapa durante la cual se aplica el algoritmo de aprendizaje automático encargado de extraer el conocimiento inherente a los datos, además de ser refinado y validado. Es por tanto en la fase de Minería de Datos en la que interviene la IA, ML de forma más concreta, pudiéndose llevar a cabo esta extracción de conocimiento tanto por algoritmos clásicos como por redes neuronales.

No obstante, el éxito de esta fase se ve influida en gran medida por la calidad de los datos que llegan para su análisis desde la fase previa, en la que se llevan a cabo unos procesos conocidos como ETL (Extract, Transform and Load) en los que los datos se someten a un proceso de revisión, limpieza y transformación, para que la información sea más fácilmente extraída, y posteriormente se cargan en el sistema de aprendizaje.

Habitualmente el esfuerzo y tiempo dedicado a estas labores de preprocesamiento de los datos superan con creces a los que se dedican al análisis de datos propiamente dichos (Data Mining).

⁶ Norberto Díaz-Díaz (2012): “Similitud Funcional de Genes basada en Conocimiento Biológico”. Tesis Doctoral.

2.1.- Mecanismos de Aprendizaje

Dentro de este proceso KDD de descubrimiento del conocimiento en los datos, los algoritmos de ML se pueden clasificar con respecto al nivel y tipo de supervisión humana hasta de 4 formas o paradigmas de aprendizaje⁷:

- Aprendizaje Supervisado: en el cual los datos de entrenamiento que aportas al algoritmo incluyen la solución deseada para que este pueda aprender. Se dice que están “etiquetados”.

Para el ejemplo del filtro de spam en el correo electrónico, entrenando el sistema a partir de un conjunto de emails etiquetados con spam y no spam, el sistema podría predecir qué tipo de correo sería uno recién recibido.

- Aprendizaje no Supervisado: los datos de entrenamiento no están etiquetados, el sistema aprende de los datos sin conocer las respuestas, sólo a partir de los datos, y es capaz de generar algún tipo de salida.
- Aprendizaje Semisupervisado: se da cuando sólo una parte de los datos de entrenamiento están etiquetados. Los algoritmos que pueden ser entrenados de esta forma suelen ser combinaciones de supervisados y no supervisados, y es el propio algoritmo el que etiqueta los datos no etiquetados. Un ejemplo muy ilustrativo sería el servicio Google Fotos, que automáticamente reconoce en determinadas fotos a la persona A, en otras a la persona B, sin saber quién es cada una, lógicamente. Esta sería la parte no supervisada del algoritmo, pero luego necesita que el usuario le diga quién es la persona A y quién es la persona B, sólo una vez cada persona y las etiquetará en cada foto que aparezcan, siendo esta la parte supervisada.
- Aprendizaje reforzado: tipo de aprendizaje en el que el sistema es un simulador o “agente” y aprende en base a ensayo-error. Tras cada ensayo llega una recompensa o una penalización, y el sistema aprende generando una estrategia o “política” que refuerza las acciones que le han llevado a la recompensa, definiendo

⁷ GÉRON, AURÉLIEN (2019): *Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow*, O'reilly, USA, pgn. 8

que acciones debe escoger el agente en una situación dada. Este aprendizaje se potencia con supercomputadoras que aceleran el aprendizaje con varias simulaciones en paralelo, de forma que en poco tiempo el sistema adquiere el aprendizaje de plazos mucho más largos. Ahora mismo es el aprendizaje que tiene menos aplicación real, aunque tenga muchas posibilidades de desarrollo. Veamos algunos casos de uso:

- Ampliando el caso del filtro de spam para el correo, una vez entrenado el sistema como supervisado, en la bandeja de spam nos pregunta si los correos son realmente spam o no lo son, para generar esa política que potencie las decisiones que ha tomado de forma correcta.
- Tiene muchas aplicaciones a la robótica, por ejemplo, para el control de calidad en líneas de producción. Los descartes, piezas defectuosas, son reforzados mediante validación del defecto por parte de operarios y el sistema aprende de esta forma potenciando los criterios que le llevaron a considerarlo como tal, defectuoso.
- Deepblue, el ordenador ajedrecista que venció a Gary Kasparov en 1996 o AlphaGo, que venció en 2015 a Ke Jie, campeón de Go, fueron entrenados con aprendizaje reforzado jugando cientos de partidas simultáneas y recibiendo recompensas o penalizaciones según procediera, con cada movimiento.
- La conducción autónoma es otro ámbito en el que el aprendizaje reforzado tiene recorrido.

Al proceso concreto del aprendizaje, también se le conoce como “entrenamiento”, término que, en adelante, usaremos indistintamente.

2.2.- Generalización del Conocimiento

Al fenómeno de evaluar una red neuronal con unos datos diferentes a los de entrenamiento o aprendizaje, en los que pueden existir patrones diferentes, obteniendo unos resultados satisfactorios, se conoce como generalización del conocimiento por parte de la red neuronal.

Es decir, esta generalización se da cuando la red da respuestas cercanas a la realidad frente a patrones que no han sido mostrados a la red en la fase de entrenamiento, y cuanto más precisa es la respuesta, se dice que la red generaliza mejor. Además, para que una red neuronal sea capaz de generalizar es necesario contar con un conjunto de datos de entrenamiento suficientemente representativo de la globalidad del problema. Si la red no generaliza es indicador de que los datos no son suficientes.

3.- ALGORITMOS CLÁSICOS DE APRENDIZAJE

Una vez visto los tipos de aprendizaje con los que podemos entrenar un sistema de ML, es decir, las formas de la que puede aprender, también los podemos clasificar a su vez en clásicos y modernos, y ahora nos centraremos en los algoritmos clásicos de aprendizaje.

Los algoritmos clásicos se consideran los diseñados con anterioridad al año 2000 y se caracterizan por su *interpretabilidad*, siendo posible entender las decisiones que toman y, en consecuencia, por poder manejar la ética de dichas decisiones, algo muy interesante y necesario en una parte de los campos de aplicación, como pueden ser la sanidad y el ámbito científico en general.

Demos un paseo por los principales algoritmos clásicos de aprendizaje centrándonos en los más conocidos:

3.1.- Aprendizaje Supervisado

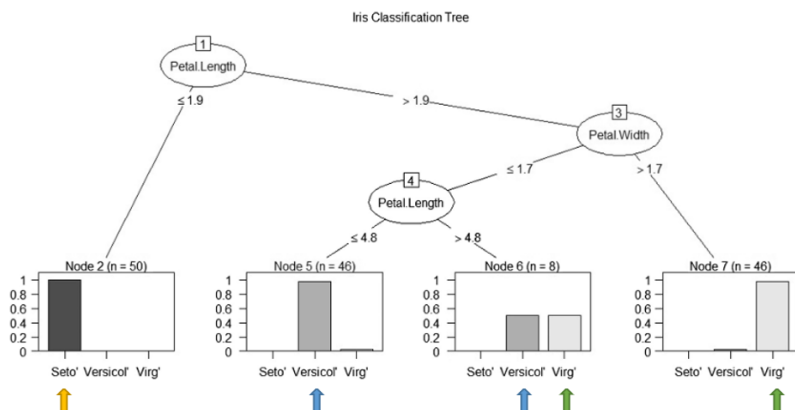
3.1.1.- Árboles de Decisión⁸

Es un modelo en el que las clases se predicen concatenando decisiones binarias, obteniendo un árbol de reglas de decisión muy fácilmente interpretable. Un problema ampliamente conocido y con gran capacidad docente para mostrar este algoritmo es el de clasificación de las flores recogidas en el conjunto de datos o dataset llamado

⁸ L. Breiman, J. H. Freidmen, R. A. Olsen and C. J. Stone (1984): *Classification and Regression Trees*, Wadsworth.

“Iris”, en el que se recogen las medidas largo y ancho, de los pétalos y sépalos de diferentes tipos de flores (setosa, versicolor y virgínica).

Fig. 5. Árbol de Decisión aplicado al conjunto de datos “Iris”⁹



Podemos apreciar como el algoritmo va dividiendo en dos grupos y paso a paso el conjunto de datos según sus características (longitud de pétalo, ancho de pétalo, en este caso) hasta obtener una clasificación del tipo de flor con una alta probabilidad.

En el proceso de aprendizaje, el algoritmo calcula los umbrales de las reglas binarias de clasificación que se aprecian en la figura anterior y que, para este algoritmo en particular, serían los parámetros que almacena. Se puede observar fácilmente la transparencia de los resultados que arroja el algoritmo, la sencillez de la interpretación de las decisiones que toma (los umbrales que mencionamos) y las variables que influyen en dichas decisiones (el ancho y largo de los pétalos). Este problema de clasificación con el algoritmo de Árboles de Decisión alcanza una precisión cercana al 100% en las predicciones.

Una versión más avanzada de los Árboles de Decisión son los *Random Forest*¹⁰ o Bosques de Árboles, que consisten en entrenar varios

⁹ Rokach, L. (2016). “Decision forest: Twenty years of research. Information Fusion”, 27, 111-125.

¹⁰ Tim Kam Ho (1995): “Random decision forests”, Proceedings of 3rd International Conference on Document Analysis and Recognition, 1995, pp. 278-282 vol.1.

Árboles de Decisión con diferentes subconjuntos del conjunto de datos de entrenamiento, y ante una nueva predicción, esta es procesada con cada uno de los árboles anteriores y el sistema elige la predicción que tenga un mayor número de votos, la predicción que se haya repetido más.

Los Árboles de Decisión tienen sus ventajas y desventajas que no los hacen adecuado para todo tipo de problemas, pero esto ocurre con cualquier algoritmo, de hecho, en gran parte el conocimiento de los expertos en esta materia, los Científicos de Datos o Data Scientist, está en conocer el mayor número de algoritmos posibles y saber escoger el apropiado para cada tipo de problema.

Son muchos los algoritmos que aprenden por supervisión, de hecho, son claramente mayoritarios en el ecosistema de algoritmos existentes. Nombremos algunos más de esos algoritmos:

- Regresiones. Las regresiones son los algoritmos que modelan la relación entre variables utilizando una función de error que minimizan en un proceso iterativo aproximándose a los valores de entrenamiento:
 - Regresión Lineal. Cuando los valores a aproximar son continuos y la aproximación se realiza mediante una curva de primer orden o lineal. En el caso de una sola variable de entrada, la curva de aproximación o regresión sería una recta.
 - Regresión Logística. Cuando los valores a aproximar son discretos, y por tanto el problema es de clasificación.
- Algoritmos basados en vecindad: K-Nearest Neighbors¹¹. Es un algoritmo que predice en base al concepto de similitud o cercanía, y que tiene la peculiaridad de que, a pesar de ser supervisado, siempre utiliza el conjunto de datos de entrada para predecir, es decir, que los datos de entrenamiento no son prescindibles una vez entrenado el algoritmo.

¹¹ S. Cheng, F. Lu, P. Peng, S. Wu, “Short-term traffic forecasting: an adaptive ST-KNN model that considers spatial heterogeneity, *Comput. Environ. Urban Syst.* 71 (2018) 186–198.

- Clasificadores Bayesianos¹². Es un algoritmo basado en teoría de probabilidades y que obtiene resultados estadísticos.
- Reglas de Asociación. Se utilizan para descubrir hechos que ocurren en común dentro de un conjunto que datos, por ejemplo, alguien que vea El Padrino y Uno de los Nuestros, ambas películas relacionadas con la mafia, posiblemente vea la serie Los Soprano.
- Support Vector Machines (SVMs)¹³. Este algoritmo busca encontrar un hiperplano que separe 2 clases diferentes de forma óptima, y esto lo consigue con el margen más amplio entre las 2 clases.
- Redes Neuronales. Ya las hemos tratado en el apartado 1.1.- Conceptos previos y Aclaraciones y lo ampliaremos más adelante.

3.2.- *Aprendizaje No Supervisado*

Recordemos que en este tipo de aprendizaje el algoritmo extrae información relevante a partir los datos de entrada, pero sólo de los datos de entrada. No tiene unas salidas de las que aprender durante el entrenamiento, es decir, los datos no están etiquetados.

Las grandes áreas donde se sitúa el aprendizaje no supervisado son:

3.2.1.- *Clustering*

El aprendizaje basado en clustering se realiza en base a algoritmos que de forma automática buscan similitudes entre los datos de entrada y crean grupos o clústeres en función de esas similitudes para poder extraer información relevante a partir de estos conjuntos. De entre los grandes tipos de clustering existentes destacamos dos:

¹² T. Liu, S. Shi, X. Gu, “Naive Bayes classifier based driving habit prediction scheme for VANET stable clustering”, *Mob. Netw. Appl.* 25 (5) (2020) 1708–1714.

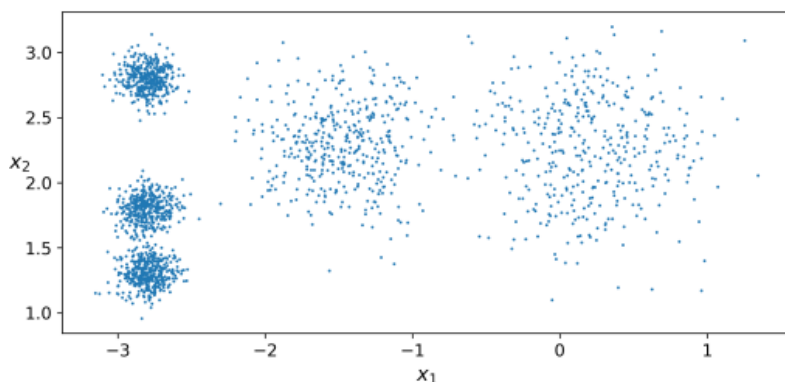
¹³ M. Hearst, S. Dumais, E. Osuna, J. Platt, and B. Scholkopf, “Support vector machines,” *IEEE Intelligent Systems and their Applications*, vol. 13, no. 4, pp. 18–28, 1998.

Clustering Particional

El más popular de los algoritmos dentro de este grupo es *K-Means*. El objetivo que persigue es encontrar grupos o clústeres en los datos, para ello, calcula la “similitud” de los datos entre sí que viene a ser un indicador de lo parecidas que son sus características, y en base a ello le asigna un clúster a cada punto.

En un conjunto de datos como el que se muestra a continuación, se representan las 2 características que definen a sus puntos, x_1 y x_2 , siendo fácil apreciar que los datos presentan una relación tal que tienden a agruparse en 5 grupos ($K=5$). K-Means, en un proceso iterativo, dividiría el espacio en 5 zonas, asignando una zona a cada punto.

Fig. 6. Conjunto de datos con varios grupos.



Una característica importante del clustering es que cada punto pertenece a un único cluster, y nunca a ninguno.

K-Means es un algoritmo muy potente, que ofrece un alto rendimiento y no tiene por qué consumir muchos recursos, pero tienes sus limitaciones, hay conjuntos de datos con los que no es efectivo, pero K-Means es sólo uno de los muchos algoritmos de clustering que existen.¹⁴

¹⁴ Xu, R., & Wunsch, D. (2005). “Survey of clustering algorithms.” IEEE Transactions on Neural Networks, 16(3), 645-678.

Clustering Jerárquico: Análisis de Clúster Jerárquico (HCA)

El clustering jerárquico (Hierarchical Clustering)¹⁵ se basa en descomponer jerárquicamente el conjunto de datos de entrada, siendo la solución representada por un dendograma. Si se parte de todos los individuos individualmente y se va agrupando por similitud de forma sucesiva, al método se le llama “método jerárquico aglomerativo”, si por el contrario, se parte de un clúster de todos los individuos, y se separan sucesivamente, se dice que es “método jerárquico divisor”.

Un dendograma es el gráfico que muestra como se van haciendo las divisiones o las agrupaciones, etapa a etapa.

Aplicaciones concretas del clustering

Una vez estudiado en qué consiste el clustering entenderemos que las aplicaciones de este, lógicamente, están relacionadas con la agrupación de instancias de datos según sus características. Por tanto, aplicaciones típicas del clustering son:

- Análisis exploratorio de datos. Cuando se analiza inicialmente un conjunto de datos, es útil saber si hay instancias de datos similares.
- Segmentación de clientes.
- Agrupación de textos por su temática.
- Localización de zonas con alta tasa de criminalidad.
- Detección de anomalías, con amplia aplicación a la detección de fraude.

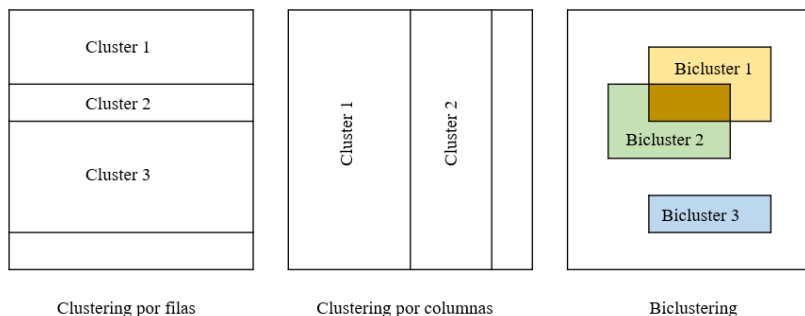
3.2.2.- Biclustering

Las técnicas de biclustering consisten, al igual que las de clustering, en agrupar los datos en conjuntos llamados biclústeres en este caso, pero, como podemos apreciar en la Fig.14, de una forma más flexible. En el caso del biclustering, las características de los datos son tratadas

¹⁵ A. K. Jain and R. C. Dubes (1998): “Algorithms for clustering data.” Prentice-Hall, Advanced Reference Series.

bidireccionalmente agrupando simultáneamente por filas y columnas, la similitud se puede producir en un subconjunto de las variables o características y se permite además el solape, es decir, que un punto puede pertenecer a varios clústers simultáneamente:

Fig. 7. Comparativa Clustering vs Biclustering



Como se aprecia, cada biclúster determina una submatriz de la matriz de datos original con algunas propiedades comunes, pudiendo tener cada biclúster espacios de características diferentes:

Los algoritmos de biclustering tienen amplia aplicación en la bioinformática.

3.3.- *Redes Neuronales*

Puesto que este capítulo trata de desarrollarse en concordancia con la evolución histórica de la IA como ya se ha comentado, haremos una breve introducción a las redes neuronales en el contexto del periodo que va de 1958 a 1986, para pasar a ampliarlo más adelante en el apartado 5.- Algoritmos modernos de aprendizaje: Redes Neuronales y Deep Learning tras la Revolución Digital y del Dato que describimos en el apartado 4.- La Revolución Digital y del Dato, momento en el cual se produce una eclosión en el desarrollo y uso de este tipo de algoritmos.

No sólo por la efectividad de sus resultados, sino también por su nombre, la representación abstracta de su topología y ciertas analogías con el cerebro humano, las redes neuronales siempre han generado cierta curiosidad. En todo caso, agrupan a una gran variedad de algoritmos de aprendizaje automático y dicha generalidad del concepto

hace que no las podamos clasificar exclusivamente dentro de alguno de los tipos de aprendizaje, existiendo redes que pueden ser entrenadas mediante mecanismos de aprendizaje supervisado, no supervisado e incluso semisupervisado.

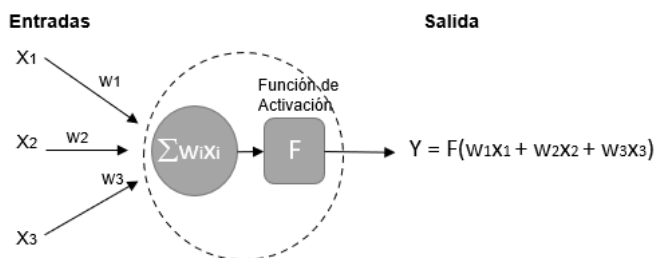
Las redes neuronales, también llamadas redes neuronales artificiales (RNA), son un modelo computacional de aprendizaje automático que procesa la información a través de un conjunto de unidades llamadas neuronas, o neuronas artificiales (Fig.17), que están conectadas entre sí y organizadas por capas, formando una red. Los datos de entrada atraviesan la red neuronal, donde son procesados mediante operaciones matemáticas, generando una salida.

Pasemos a explicar algunos detalles y conceptos que caracterizan a las redes neuronales:

Perceptron – Neurona Artificial

El perceptrón es una de las arquitecturas de red neuronal más simples, que como ya vimos fue inventada por Frank Rosenblatt en 1958, y simplificando mucho, podríamos decir que se corresponde con una neurona artificial.

Fig. 8. Neurona Artificial



Pesos – Parámetros

Los pesos o parámetros, términos sinónimos representados en la Fig.8 por “w”, expresan la importancia de la entrada respectiva en la salida a través de la suma ponderada de entradas y pesos. Estos parámetros se calculan en el proceso de aprendizaje de la red.

Activación

La función de activación de una neurona, “F” en la Fig.8, es la función que se aplica sobre la suma ponderada de los pesos con las entradas.

Se llama así porque la neurona clásica o perceptrón generaba una salida si tras aplicar dicha función “F” a las entradas, la salida “Y” superaba un determinado umbral y por tanto se activaba.

Perceptrón Multicapa

Es una red neuronal compuesta de varias capas formadas cada una de ellas por neuronas. Constan de una capa de entrada, una o varias capas ocultas, y una capa de salida. El Perceptrón Multicapa o Multi-layer Perceptron (MLP), es como se llaman a las redes neuronales de dos (capa oculta y capa de salida) o más capas, y que son feedforward y fully connected.

- Feedforward. Es la arquitectura de las redes neuronales en las que una capa sólo alimenta a la siguiente, y por tanto no existen retroalimentaciones.
- Fullyconnected. También conocidas como capas densas, son las capas de neuronas en las que cada neurona se conecta con todas las de la capa anterior.

Con estos conceptos ya podemos entender la representación abstracta de la red neuronal de tipo Perceptrón Multicapa:

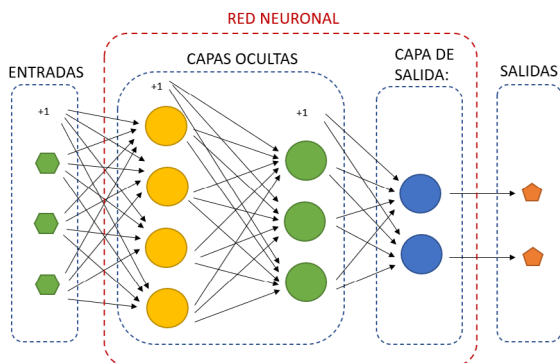


Fig. 9. Red Neuronal del tipo Perceptrón Multicapa.

El Perceptrón Multicapa utiliza el aprendizaje supervisado y el backpropagation, que veremos a continuación, como método de entrenamiento.

Fases de Entrenamiento y Prueba

Existen dos fases por las que debe pasar toda red neuronal: la fase de aprendizaje o entrenamiento (train, en inglés) y la fase de prueba (test, en inglés).

En la fase de entrenamiento, se usa un conjunto de datos para determinar los pesos o parámetros, que definen la red neuronal. Una vez entrenado el modelo, se pasa a la fase de prueba, en la que se procesan los datos de prueba, datos que no han sido procesados por la red en la fase de entrenamiento, analizándose de esta manera el desempeño de la misma.

Entrenamiento Backpropagation¹⁶

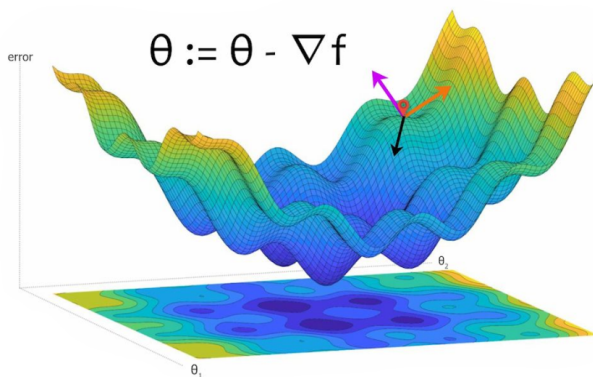
Profundizando algo más en la fase de entrenamiento, una red neuronal aprende no sólo calculando sus propios pesos, sino actualizándolos. Dicha actualización se realiza buscando unos pesos óptimos que aproximen en la mayor medida los resultados que la red genera a los reales o exactos, y para llevar a cabo esa optimización, se utiliza lo que se denomina una “función de coste” que valore esa diferencia de valores calculados vs reales, *convirtiéndose así el proceso de entrenamiento en uno de optimización de la función de coste*. Así, por ejemplo, las fases del entrenamiento para un aprendizaje supervisado serían:

1. Forward Pass: los pesos de la red se inicializan aleatoriamente y se procesa la entrada capa a capa hasta calcular a la salida, en un proceso hacia adelante.
2. Elegimos una función de coste o error mediante la cual comparamos el resultado predicho por la red con el real esperado.

¹⁶ Wythoff, B. J. (1993). “Backpropagation neural networks: A tutorial.” *Chemo-metrics and Intelligent Laboratory Systems*, 18(2), 115-155.

3. Iniciamos el proceso de optimización a través de la modificación de los valores de los parámetros de forma que disminuya el error. Y esta es la clave del *Algoritmo de Backpropagation*, mediante el cual en un recorrido hacia atrás se identifica la “responsabilidad” de cada neurona en el error, y los errores son propagados hacia atrás para ir ajustando los parámetros conforme a esa “responsabilidad”.
4. En un último paso se aplica el algoritmo de optimización mediante el cual se calculan los nuevos pesos. Hay varios algoritmos, pero de ellos quizá el más conocido sería el *Descenso del Gradiente*, que básicamente modifica los pesos en la dirección en que el error disminuye con máxima pendiente, de forma que en cada iteración se va buscando el mínimo de la función de error. En la siguiente figura se representa la curva de la función de error y como el punto de cálculo se modifica en la dirección de máxima disminución de la pendiente

Fig. 10. Descenso del Gradiente



4.- LA REVOLUCIÓN DIGITAL Y DEL DATO

En torno al año 2000, se producen una serie de fenómenos y circunstancias que facilitarán una disrupción en el desarrollo de la IA, a través fundamentalmente de las RN. Dicho fenómeno es el que se conoce como La Revolución Digital y del Dato.

Las causas que la facilitaron son, por un lado, que se produce una evolución del hardware tanto hacia una mejor capacidad de computación como de almacenamiento, motivadas por:

- La Ley de Moore¹⁷ sigue cumpliendo su previsión, y la capacidad de procesamiento sigue duplicándose cada dos años.
- El mercado del gaming ha hecho evolucionar las GPUs (Unidades de Procesamiento Gráfico) a ritmo exponencial y de forma casual y posteriormente, se comprobó que estas GPUs optimizaban el procesamiento de las redes neuronales, motivo por el cual también se desarrollaron procesadores similares pero específicos para el procesamiento de redes neuronales, como son las TPUs.
- Con todo ello, lo que antes tardaba semanas en entrenarse, ahora se entrena en días o incluso en horas.
- Gran despliegue de los servicios en la nube (cloud) como soluciones económicas de almacenamiento y procesamiento con pago por uso, “variabilizando” un importante coste como el de infraestructura.
- Abaratamiento de la memoria RAM.

Gracias a todo esto, hoy podemos almacenar y procesar cantidades de datos impensables hace unos años, a unos costes muy inferiores y fácilmente controlables.

Por otro lado, la creciente digitalización en términos generales de la sociedad, bien sean los servicios, las empresas, nuestros hábitos de consumo, de relacionarnos, los dispositivos conectados (IoT) y un largo etc, ha desencadenado en una generación masiva de datos y en el reconocimiento del dato como un activo estratégico por las empresas. En las empresas, además, se comienza a utilizar la analítica avanzada.

Y finalmente la investigación en algoritmos de aprendizaje automático, especialmente redes neuronales, no se detuvo y como consecuencia de ello se descubrieron mejores algoritmos y mejores técnicas para entrenarlas.

¹⁷ Moore, G. E. (1998). “Cramming More Components onto Integrated Circuits.” PROCEEDINGS OF THE IEEE, 86(1), 4.

Estas circunstancias han generado el ecosistema perfecto para la eclosión de la IA, y especialmente, de las redes neuronales profunda o Deep Learning (DL) como veremos a continuación.

Prácticamente todos los avances en IA sorprendentes y mediáticos que se han producido recientemente son logrados usando ML y DL, usando un hardware de última generación y cantidades ingentes de datos para entrenar a redes neuronales.

4.1.- La Importancia de Los Datos

Ya hemos visto que para que nuestros algoritmos de ML funcionen correctamente en todo el dominio de aplicación, es necesario tener un conjunto de datos de entrenamiento suficientemente representativo del dominio sobre el que versa el problema.

A menudo se oye hablar del sesgo de la IA. Si tenemos un conjunto de datos con unos sesgos implícitos y entrenamos un sistema de aprendizaje automático con dicho conjunto de datos, ese sesgo queda impregnado en el sistema mediante el aprendizaje adquirido por el sistema. Esta es una de las grandes críticas que recibe la IA, pero debemos ser conscientes que la IA no ha adquirido más sesgo que el que ya tenían los datos, y que los datos están influenciados por las personas que los han generado o preparado.

Si un algoritmo discrimina por raza o sexo, por ejemplo, el motivo será que los datos de entrenamiento de ese sistema tenían implícitos patrones discriminatorios. Por ejemplo: en 2014 el equipo de IA de Amazon desarrolló una aplicación para revisar currículums. Cuando el sistema estaba en funcionamiento se dieron cuenta que el algoritmo discriminaba a las mujeres para puestos técnicos. Sin embargo, al basarse en los datos de los últimos 10 años, el sistema aprendió que los hombres eran “preferibles”, ya que eran los hombres quienes más se presentaban a este tipo de puestos.

Hay técnicas para evitar el sesgo en los conjuntos de datos, como el balance de clases, es decir, que las proporciones entre las diferentes clases a predecir sean equilibradas.

5.- ALGORITMOS MODERNOS DE APRENDIZAJE: DEEP LEARNING

Llegado este momento histórico en el que ya se ha producido la revolución digital y del dato descrita anteriormente y nos encontramos en la era del Big Data, dedicaremos un espacio exclusivo a las Redes Neuronales Profundas y al uso que se hace de ellas en los tiempos recientes, posteriores al año 2012, pues desde entonces muchos de los avances que se han producido en IA son debidos a las redes neuronales profundas.

5.1.- *Redes Neuronales Profundas: Deep Learning*

El concepto de DL hace referencia a las redes neuronales de un gran número de capas. No existe un criterio claro en cuanto a partir de qué número de capas ocultas podemos considerar una red neuronal como profunda, y por tanto, DL, pero hay una opinión cada vez más extendida entre los expertos que afirma que cualquier red con más de 2 capas ocultas puede considerarse “profunda”¹⁸.

5.2.- *Diferencias entre Machine Learning y Deep Learning*

Más allá de las aclaraciones mostradas en el apartado 1.1.- Conceptos previos y Aclaraciones, entraremos ahora a un mayor detalle para diferenciar ambos subconjuntos de la IA, matizando que cuando nos refiramos a ML lo haremos realmente a los algoritmos clásicos de ML, anteriores al año 2000, y no incluimos por tanto las redes neuronales profundas.

La primera diferencia que podríamos comentar sería en cuanto a la *interpretabilidad*. Ya vimos, usando los árboles de decisión, como se muestra en el apartado 3.1.1.- Árboles de Decisión, como de interpretables pueden ser los algoritmos clásicos de ML. Todo lo contrario ocurre con las redes neuronales profundas, que son difícilmente interpretables hasta tal punto de considerarlas cajas negras (black box)

¹⁸ Coşkun, M., Yildirim, Ö., Uçar, A., & Demir, Y. (2017). “AN OVERVIEW OF POPULAR DEEP LEARNING METHODS”. *European Journal of Technique (EJT)*, 7(2), 165-176.

que generan salidas a partir de unas entradas, con un mecanismo en la toma de decisión muchas veces difícil de explicar aunque los resultados del entrenamiento sean altamente precisos.

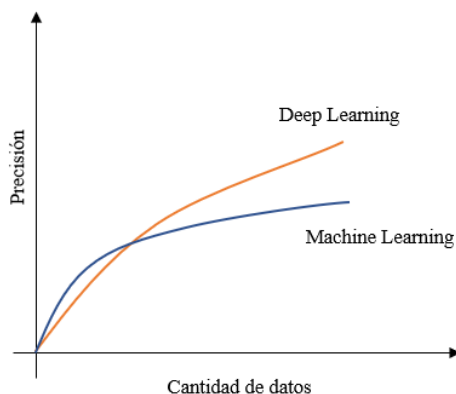
Fig. 11. Red Neuronal Profunda – Blackbox



El hecho de que las decisiones que toman no sean interpretables supone un impedimento para el uso de las RN Profundas en determinados ámbitos como son la sanidad y el ámbito científico, ámbitos en los que se deben justificar las decisiones que se toman o ámbitos en los que una persona, con nombres y apellidos, asume la responsabilidad sobre dichas decisiones.

Imaginemos una red neuronal que a partir de imágenes radiológicas es capaz de diagnosticar un cáncer de pulmón. Si no sabemos en base a qué o por qué motivos la red determina un positivo, es decir, que un paciente tiene cáncer, no podemos utilizar el diagnóstico determinado por la RN al menos como única base para ello. El médico podrá considerarlo un indicio, pero necesitará otras pruebas que lo corroboren.

Otra diferencia por destacar entre ML y DL sería el tamaño de los modelos y la cantidad de datos de entrenamiento. Se ha comprobado que a más datos le suministremos a una red neuronal, más neuronas y más capas, y por tanto, más parámetros tenga, las redes neuronales tienen mayor capacidad para aprender y alcanzan mejores resultados presentando un mejor rendimiento. Esto no ocurre de la misma forma con los modelos clásicos de ML, en los que el rendimiento mejora cuando aumentamos la cantidad de datos aportados para el entrenamiento, pero llegado a un punto, el rendimiento satura llegando a un máximo y deja de mejorar, como podemos apreciar en la Fig. 12. Por este motivo los algoritmos clásicos no se pueden aplicar a problemas complejos, que requieran una gran cantidad de datos.

Fig. 12. Cantidad de datos de entrenamiento

Esto es lo que sucede con el modelado del lenguaje humano, ámbito en el cual una red neuronal que ha supuesto una disrupción es la conocida como GPT-3¹⁹, y que tiene 175 mil millones de parámetros. Este problema sería inabarcable con modelos clásicos de ML con los que no se pueden alcanzar estas escalas o tamaños.

Si observamos con detenimiento la ilustración de la Fig.12 podemos ver que para problemas en los que no tengamos una gran cantidad de datos, será más apropiado usar un algoritmo clásico de ML pues dará mejores resultados que una red neuronal profunda. Sin embargo, ante una gran cantidad de datos vemos que el desempeño que dan las redes neuronales profundas es siempre mejor, y mejorando a mayor cantidad de datos.

Este es uno de los principales motivos del espectacular crecimiento de la IA a través de las redes neuronales profundas a partir del año 2000, que la disponibilidad de gran cantidad de datos que se generan de forma masiva en la era digital se traduce en mejoras de rendimiento. Había modelos que existían con anterioridad, como las

¹⁹ Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodi, D. (2020). "Language Models are Few-Shot Learners" (arXiv:2005.14165). arXiv.

Redes Neuronales Convolucionales (CNN) o las Recurrentes (RNN), que consumían muchos recursos, pero su uso y resultados se vieron potenciados a mayores con las mejoras en capacidad de procesamiento, memoria y almacenamiento.

En el punto de desarrollo en el que nos encontramos ahora mismo, la adecuación de usar algoritmos clásicos o redes neuronales dependerá mucho del dominio de uso, por ejemplo, está más que contrastado que las redes neuronales tienen un excelente comportamiento en el procesamiento de imágenes y del lenguaje natural, pero por ejemplo en series temporales, hay métodos clásicos y estadísticos, como ARIMA, que dan tan buenos resultados incluso mejores que las redes neuronales. Habrá casos en los que una técnica funcione mejor que la otra, y casos en los que merecerá la pena probar ambas técnicas, y esto dependerá del dominio o ámbito del problema y del tamaño del conjunto de datos de entrenamiento del algoritmo, fundamentalmente.

No hay una regla objetiva para esta elección, aunque si no te estás moviendo en el rango de varios miles de datos, no tiene sentido trabajar con redes neuronales profundas.

Por último, el consumo de recursos será otro aspecto a tener en cuenta, pues en términos generales, una red neuronal profunda requerirá una mayor capacidad de procesamiento que un algoritmo clásico y aunque habrá problemas sencillos que se podrían resolver con una red neuronal, como sería el caso de una regresión lineal, no es lo óptimo pudiendo hacerse con un algoritmo clásico, con un menor consumo de recursos.

5.3.- Transversalidad del Deep Learning

Las RN en general y el DL en particular, son algoritmos con amplia aplicación en campos muy diversos, como ya hemos visto. Los algoritmos empleados en muchos de estos campos son esencialmente los mismos o muy similares, quizá con algunas adaptaciones al dominio que permitan, por ejemplo, usarlos con un determinado tipo de datos. Esto se traduce en que una mejora en dichos algoritmos repercute en una gran cantidad de aplicaciones y sectores.

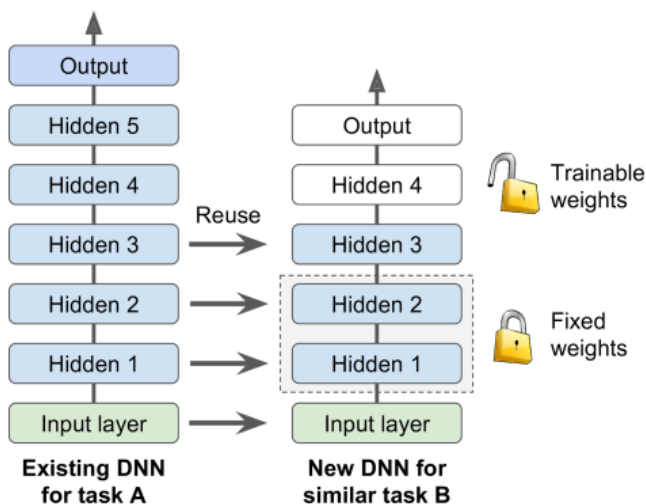
Por ejemplo, el hecho de que el DL sea muy apropiado para el procesamiento de imágenes hace que un avance en el tipo de RN utilizados

para ello, como pueden ser las convolucionales, repercute directamente en aplicaciones tan diversas como la conducción autónoma, la medicina (detección del cáncer a través de pruebas basadas en imágenes, como la radiología) o el control de calidad en fabricación (identificando piezas defectuosas en la misma línea de producción), entre otras.

5.4.- Transfer Learning

En términos generales, no es buena idea entrenar desde cero una RN profunda, por la gran cantidad de datos y recursos que precisa. En vez de ello, lo que se puede hacer es buscar una red que resuelva un problema parecido a lo que necesitamos y utilizar las primeras capas, que habrán aprendido patrones más genéricos. Así no sólo acelerará el entrenamiento, sino que además necesitará muchos menos datos para el entrenamiento.

Fig. 13. Reutilización de capas pre entrenadas²⁰



²⁰ GÉRON, AURÉLIEN (2019): Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow, O'reilly, USA, pgn. 340

Si el dominio es muy parecido, incluso se podrá utilizar la red completa y entrenada, y sólo habrá que hacerle un entrenamiento con un conjunto de datos relativamente pequeño en comparación a entrenarla desde cero.

Este sería el caso de tener una RN que ha sido entrenada con un gran conjunto de imágenes de una amplia variedad de clases en general (flores, animales, electrodomésticos, mobiliario urbano, vehículos, señales de tráfico...), y te surja la necesidad de identificar motos en determinadas imágenes. Los patrones que la red tiene aprendidos, como podrían ser líneas curvas, texturas, quizá hasta círculos, etc. junto con una segunda fase de entrenamiento sólo con imágenes de motos, lo que se conoce como “orientación a dominio”, probablemente darán resultados de alto desempeño.

Gracias al Transfer Learning no siempre será estrictamente necesario una gran cantidad de datos ni una gran capacidad de computación para desarrollar una solución de IA.

6.- EL PRESENTE Y FUTURO DE LA INTELIGENCIA ARTIFICIAL

Siguiendo el desarrollo cronológico de este capítulo, venimos y estamos en una época de esplendor de la IA, pero deberíamos plantearnos qué pasará en los próximos años.

En base a los hechos que hemos desarrollado, todo apunta a que se mantendrá esta época de esplendor, porque se mantienen las condiciones que la auspiciaron y que siguen siendo las propicias para el sostenimiento del mismo, pero desde un punto de vista más técnico, se vislumbra un cambio de paradigma en el desarrollo de los sistemas de IA.

Hasta la fecha se ha puesto la mayor parte de los esfuerzos en desarrollar modelos potentes, de gran tamaño con millones de parámetros, modelos poco interpretables como hemos comentado, sin embargo, desde hace años se está produciendo un giro para poner el foco en los datos y en la calidad de estos. Ya en 2001 se pudo comprobar que el factor más influyente en la precisión de un sistema es

la cantidad de datos disponibles por encima del modelo elegido²¹, sin embargo, según palabras textuales de Andrew NG²²:

“Un análisis de una muestra de publicaciones recientes reveló que el 99% de las publicaciones estaban basadas en el modelo y sólo un 1% lo estaban en el dato”

En una de sus famosas conferencias en la Universidad de Stanford el propio Andrew NG bautizó este cambio de paradigma como “from Model-Centric to Data-Centric Approach”.

Por ello, el próximo enfoque para seguir mejorando el desempeño de los modelos va alineado con disponer de datos de una mayor calidad que permitan extraer conocimiento de ellos más fácilmente, teniendo en cuenta que, gracias al Transfer Learning, podemos obtener resultados similares a los de entrenar una RN desde cero y con una gran cantidad de datos, simplemente con una cantidad relativamente pequeña de estos, pero de calidad y centrados en el dominio de nuestra aplicación.

7.- CONCLUSIONES

La Inteligencia Artificial viene y vive una época de esplendor que está teniendo un gran impacto en el mundo, hasta el punto de existir soluciones que se entrenan y aprende de forma autónoma sin ser necesarias programarlas “ad hoc” o explícitamente para una aplicación concreta, como sería el caso de las redes neuronales profundas que están procesando textos realizando resúmenes, traduciendo en tiempo real, etc. Lo hacen sin entender de reglas gramaticales, pero con capacidad de aprender y extraer conocimiento si las entrenas con cientos de libros.

²¹ Banko, M. and Brill, E. (2001) , “Scaling to Very Very Large Corpora for Natural Language Disambiguation”, ACM DL DIGITAL LIBRARY.

²² Andrew Ng (1976), Profesor Asociado al Dpto de Ciencias de la Computación, Universidad de Standford. Director del Laboratorio de IA, Standford. Fundador de Coursera.

Todo esto ha sido posible gracias a las mejoras en hardware, en el desarrollo de algoritmos y técnicas de entrenamiento, y a la disponibilidad de gran cantidad de datos.

A pesar de ello, la IA aún debe superar barreras como son la interpretabilidad de los algoritmos de DL, lo cual permitirá que se usen en cualquier ámbito una vez las personas responsables de tomar decisiones lo puedan hacer con el respaldo de poder justificarlas.

El progreso de la IA sin duda pasará por la puesta en el foco en los datos, en su calidad y en la importancia que estos tienen, porque no debemos olvidar, que usemos el algoritmo que usemos, el problema lo resolverá a partir del conocimiento que extraiga de los datos que le aportemos.

8.- BIBLIOGRAFÍA CITADA

- BANKO M., BRILL, E. (2001), *Scaling to Very Very Large Corpora for Natural Language Disambiguation*, ACM DL DIGITAL LIBRARY.
- BROWN, T. B., MANN, B., RYDER, N., SUBBIAH, M., KAPLAN, J., DHARIWAL, P., NEELAKANTAN, A., SHYAM, P., SASTRY, G., ASKELL, A., AGARWAL, S., HERBERT-VOSS, A., KRUEGER, G., HENIGHAN, T., CHILD, R., RAMESH, A., ZIEGLER, D. M., WU, J., WINTER, C., AMODEI, D. (2020). "Language Models are Few-Shot Learners" (arXiv:2005.14165). arXiv.
- COŞKUN, M., YILDIRIM, Ö., UÇAR, A., DEMIR, Y. (2017). "AN OVERVIEW OF POPULAR DEEP LEARNING METHODS". *European Journal of Technique (EJT)*, 7(2), 165-176.
- DOUGLAS, L. (2001), "3D Data Management: Controlling Data Volume, Velocity and Variety". Gartner.
- GÉRON, AURÉLIEN (2019): *Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow*, O'reilly, USA, pgn. 8, 340.
- JAIN K. and DUBES R. C. (1998): *Algorithms for clustering data*. Prentice-Hall, Advanced Reference Series.
- BREIMAN L., FREIDMEN J. H., OLSEN R. A., STONE C. J., (1984): *Classification and Regression Trees*, Wadsworth.
- HEARST M., DUMAIS S., OSUNA E., PLATT J., SCHOLKOPF B., "Support vector machines," *IEEE Intelligent Systems and their Applications*, vol. 13, no. 4, pp. 18-28, 1998.
- MOORE, G. E. (1998). "Cramming More Components onto Integrated Circuits." *PROCEEDINGS OF THE IEEE*, 86(1), 4.

- DÍAZ-DÍAZ, N. (2012): “Similitud Funcional de Genes basada en Conocimiento Biológico”. Tesis Doctoral.
- ROKACH, L. (2016). “Decision forest: Twenty years of research”. *Information Fusion*, 27, 111-125.
- CHENG S., LU F., PENG P., WU S., “Short-term traffic forecasting: an adaptive ST-KNN model that considers spatial heterogeneity”, *Comput. Environ. Urban Syst.* 71 (2018) 186–198.
- TIM KAM HO (1995): “Random decision forests”, Proceedings of 3rd International Conference on Document Analysis and Recognition, 1995, pp. 278-282 vol.1.
- LIU T., SHI S., GU X., “Naive Bayes classifier based driving habit prediction scheme for VANET stable clustering”, *Mob. Netw. Appl.* 25 (5) (2020) 1708–1714.
- XU, R., WUNSCH, D. (2005). “Survey of clustering algorithms.” *IEEE Transactions on Neural Networks*, 16(3), 645-678.
- WYTHOFF, B. J. (1993). “Backpropagation neural networks: A tutorial.” *Chemometrics and Intelligent Laboratory Systems*, 18(2), 115-155.

Capítulo II

¿Para qué sirve la Inteligencia Artificial en el sector público? Casos de uso y perspectivas de aplicación

CARLOS TORRECILLA-SALINAS, LUCA TANGI,
PETER ULRICH, MARINA MANZONI, SVEN SCHADE,
EVA MARTÍNEZ-RODRÍGUEZ, FRANCESCO PIGNATELLI
*Unidad de Economía Digital. Centro Común de Investigación
Comisión Europea*

Resumen: La transformación digital es uno de los fenómenos clave de nuestro tiempo. Dentro de las tecnologías en las que se basa dicha transformación, destaca el conjunto de las que denominamos Inteligencia Artificial por su potencial transformador. El despliegue de las mismas no es ajeno al sector público, que se enfrenta a retos específicos que pueden cuestionar su propia legitimidad pero que deben ser afrontados para la obtención de los beneficios asociados al uso de esas tecnologías. El presente artículo presenta, a alto nivel, algunas de las especificidades del uso de la Inteligencia Artificial en el sector público así como un resumen de las conclusiones de los estudios que el Centro Común de Investigación de la Comisión Europea ha realizado tanto en relación al uso de la Inteligencia Artificial en la administración pública europea como para la definición de una hoja de ruta para el despliegue exitoso de proyectos de implantación de soluciones basadas en Inteligencia Artificial.

Abstract: Digital transformation is one of the key phenomena of our time. Within the technologies that drive such transformation, the set of them that we call Artificial Intelligence emerge due to its transformative potential. The public sector cannot escape from the uptake of AI and it is facing specific challenges that can even question government legitimacy. Such challenges have to be successfully faced to grasp the benefits associated to the use of AI. This paper presents, in a high level, some of the specificities of the use of AI within the public sector, together with a summary of the conclusions of the reports performed by the Joint Research Centre of the European Commission, both in relation to the use of AI within the European public administration together with the definition of a roadmap to successfully deploy AI-based solutions.

Palabras-clave: Inteligencia Artificial, sector público, Unión Europea, transformación digital

Key-words: Artificial Intelligence, public sector, European Union, digital transformation

SUMARIO.- 1. ¿Cómo definir la Inteligencia Artificial?.-2. ¿Por qué es importante la Inteligencia Artificial para el sector público?.-3. ¿Cómo se utiliza actualmente la Inteligencia Artificial en el sector público?.-4. Conclusiones y recomendaciones.

1.- ¿CÓMO DEFINIR LA INTELIGENCIA ARTIFICIAL?

Desde finales del siglo XX, las tecnologías digitales han ido incorporándose a nuestras organizaciones en sucesivas oleadas, primero a través de procesos de informatización (inclusión de tecnologías de la información y las comunicaciones en los procesos de las organizaciones, buscando en general una mejora de la eficiencia) y más adelante, a través de lo que se ha dado en llamar “transformación digital”, es decir, el uso de tecnologías digitales para modificar integralmente el funcionamiento de las organizaciones. Este proceso de adopción se ha acelerado notablemente, y en particular a raíz de la pandemia de COVID-19, en el que la adopción de herramientas digitales para el trabajo, la educación o el comercio se incrementó de forma significativa.

Dentro de este proceso general de transformación digital que describimos, hay un conjunto de tecnologías que destacan, por su carácter transformador. Es el conjunto de tecnologías que se conoce como Inteligencia Artificial. Podemos destacar tres razones fundamentales por las cuales las tecnologías de Inteligencia Artificial se distinguen de las demás tecnologías digitales en el marco de la transformación digital:

- Su capacidad disruptiva, ya que debido al elevado volumen de datos que pueden procesar y a su capacidad predictiva, pueden afectar de una forma radical a las organizaciones.
- Su capacidad de penetración, ya que muchas de las tecnologías de Inteligencia Artificial se pueden clasificar como tecnologías de propósito general, virtualmente aplicables a cualquier ámbito de la sociedad y la economía.
- Su rápida velocidad de adopción, ya que muchas de esas tecnologías se encuentran en niveles de madurez adecuados para su adopción por parte de organizaciones públicas y privadas (para más detalles sobre los niveles de madurez de las diversas tecnologías de Inteligencia Artificial ver MARTÍNEZ PLUMED (2022)).

Sin embargo, encontrar una definición única y consensuada de la Inteligencia Artificial es un proceso bastante complicado. El trabajo de SAMIOLI (2021) nos muestra la existencia de más de 60 defini-

ciones diferentes de ese conjunto de tecnologías desde los comienzos de la investigación en el área durante la década de los 50 del pasado siglo, propuestas por diversos organismos de distinta índole (institutos de investigación, organismos públicos y/o organizaciones de estandarización). Además, podemos detectar que el término Inteligencia Artificial es incluso utilizado de forma errónea, inclusive a veces como “reclamo comercial”, por parte de algunas organizaciones.

Sin embargo, esta diversidad de definiciones también nos muestra la dificultad de precisar el concepto de Inteligencia Artificial y nos pone de manifiesto el hecho de que quizá debamos asumir la co-existencia de múltiples y diversas definiciones, en función del fin para el cual se quiera utilizar la misma. No obstante, y debido a la emergencia de estas citadas tecnologías, el encontrar una definición de Inteligencia Artificial que pueda ser utilizada con fines regulatorios es cada vez más urgente.

En relación a este último objetivo, cabría destacar por tanto dos definiciones de Inteligencia Artificial que son verdaderamente relevantes a nivel europeo. La primera, elaborada por el Grupo Independiente de Expertos de Alto Nivel sobre Inteligencia Artificial (HLEG-AI, 2018) en sus “Guías Éticas para una IA fiable”, que mostramos a continuación:

“Los sistemas de inteligencia artificial (IA) son sistemas de software (y en algunos casos también de hardware) diseñados por seres humanos que, dado un objetivo complejo, actúan en la dimensión física o digital mediante la percepción de su entorno a través de la obtención de datos, la interpretación de los datos estructurados o no estructurados que recopilan, el razonamiento sobre el conocimiento o el procesamiento de la información derivados de esos datos, y decidiendo la acción o acciones óptimas que deben llevar a cabo para lograr el objetivo establecido”.

En esta definición encontramos ya algunos elementos clave relativos a la Inteligencia Artificial: están basados en objetivos, interactúan de forma independiente con su entorno y deciden de forma autónoma las acciones a realizar para la consecución de sus objetivos. Pueden ser sistemas solo basados en software pero también estar integrados en sistemas físicos.

Posteriormente al trabajo del citado grupo de expertos, la Comisión Europea propuso en 2021 una serie de normas armonizadas en materia

de Inteligencia Artificial, más conocidas como Ley de Inteligencia Artificial o AI Act (COMISIÓN EUROPEA, 2021). Este reglamento, uno de los primeros en el mundo que tiene por objetivo la regulación de la Inteligencia Artificial, incluye también una definición que determina su marco de aplicación. Tanto el reglamento como la definición que incorpora, se encuentran bajo discusión actualmente, sin embargo, dada la ambición y la novedad de la propuesta, es interesante conocerla incluso si aún no se ha adoptado formalmente por la Unión Europea. El reglamento define los sistemas de Inteligencia Artificial como:

“El software que se desarrolla empleando una o varias de las técnicas y estrategias que figuran en el anexo I y que puede, para un conjunto determinado de objetivos definidos por seres humanos, generar información de salida como contenidos, predicciones, recomendaciones o decisiones que influyan en los entornos con los que interactúa”.

El anexo I incluye diversas técnicas tales como estrategias de aprendizaje automático, supervisado o no, aprendizaje profundo, motores de inferencia y deducción, sistemas expertos y de razonamiento (simbólico), entre otros. En este caso, podemos comprobar como la propuesta de legislación se centra en los sistemas lógicos (y no físicos) e incluye una lista actualizable de técnicas que calificarían a los sistemas como Inteligencia Artificial. También podemos comprobar que elementos como definición de objetivos, autonomía, toma de decisiones o interacción con el entorno son comunes a ambas definiciones.

2.- ¿POR QUÉ ES IMPORTANTE LA INTELIGENCIA ARTIFICIAL PARA EL SECTOR PÚBLICO?

Tradicionalmente, el sector público ha jugado dos papeles principales en relación a la Inteligencia Artificial. Por un lado, el sector público actúa como regulador, estableciendo el marco legal y las reglas para el uso de la tecnología en la sociedad. Por otro, la administración pública tiene también el rol de “acelerador” proporcionando financiación y dando soporte a la adopción de la Inteligencia Artificial por parte de la sociedad. Existe, sin embargo, un tercer papel para el sector público que es el de desarrollador y usuario de soluciones de Inteligencia Artificial. Este último rol es el que recibe menos atención de los tres.

No obstante, el uso de la Inteligencia Artificial por parte de la administración pública presenta una serie de oportunidades evidentes, cómo pueden ser, entre otras, las siguientes:

- La mejora de los procesos internos de la administración, gracias a la capacidad de procesar grandes volúmenes de información rápidamente y de una forma eficaz.
- La mejora de los servicios públicos que el gobierno presta al ciudadano, incrementando la personalización, el soporte y ayuda, o la interacción con el mismo.
- La innovación en los servicios públicos, a través de la creación de nuevas formas de proporcionar atención a los ciudadanos.
- Una toma de decisiones mejorada, a través del análisis de más y más diversas fuentes de datos.

Estas oportunidades, aunque evidentes, deben ponerse en relación con los riesgos que el uso de este conjunto de tecnologías acarrea. Algunos de estos riesgos son comunes a la aplicación de la Inteligencia Artificial en el sector privado, cómo pueden ser aquellos asociados a la protección de datos de carácter personal o a garantizar que las decisiones que se toman no se encuentran sesgadas, ya sea por los datos utilizados o por los propios sistemas algorítmicos desarrollados. Sin embargo, existen riesgos que son específicos a la aplicación de la Inteligencia Artificial al sector público.

Como sabemos, la relación entre los ciudadanos y la administración es única y particular, ya que, a través de nuestros sistemas políticos y nuestras constituciones, la sociedad otorga a la administración una serie de potestades especiales en ámbitos especialmente importantes para la vida de las personas, como pueden ser la educación, la salud o el empleo. Esta relación especial que otorga una serie de poderes exclusivos o semi-exclusivos al sector público y que permite que los gobiernos provean de servicios a los ciudadanos a veces en régimen de monopolio, se sustenta en la legitimidad que nuestros sistemas democráticos otorgan a la administración, que tiene la obligación de tratar a todos los ciudadanos de la misma manera, partiendo de la base que su actuación será siempre equilibrada, justa, razonada y ajustada a derecho.

Sin embargo, la toma de decisiones administrativa basada en sistemas de Inteligencia Artificial que dé lugar a decisiones arbitrarias, no

fundamentadas y evidentemente injustas, sobre todo en ámbitos críticos para el ciudadano y en los que muchas veces el ciudadano no tiene capacidad de elección, puede afectar directamente a la legitimidad de los gobiernos y las administraciones públicas. El incidente ocurrido recientemente en Holanda, donde el resultado de un algoritmo determinó erróneamente que ciertas familias con pocos recursos debían devolver una prestación social presuntamente percibida indebidamente, es un ejemplo de este riesgo¹.

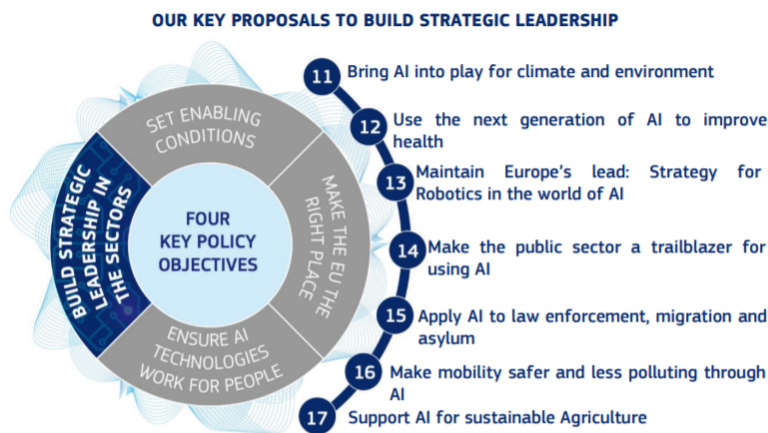
Hemos indicado ya las oportunidades y los riesgos que conlleva la Inteligencia Artificial para los gobiernos, pero también son múltiples las razones por las cuales el uso de la Inteligencia Artificial por parte del sector público puede acelerar su despliegue y uso en las sociedades europeas, con evidente beneficio económico. Por un lado, la compra pública representa el 14% del Producto Interior Bruto europeo (COMISIÓN EUROPEA, 2017), lo cual sería una palanca muy poderosa si tan solo una pequeña parte de ese volumen se utilizara para la adquisición y el desarrollo de soluciones de Inteligencia Artificial. Sin embargo, según (COMISIÓN EUROPEA, 2021), en 2018 solo el 3,5% del volumen dedicado a compra pública se dedicó a soluciones digitales innovadoras, lo que demuestra el enorme potencial aún sin explotar por parte del sector público europeo.

Debido a las razones anteriormente mencionadas y en relación con el uso de la Inteligencia Artificial en Europa, la Comisión Europea adoptó en 2018 (COMISIÓN EUROPEA, 2018) una comunicación denominada “*Inteligencia Artificial para Europa*” con el objetivo de garantizar que la Unión fuera competitiva en el ámbito de la Inteligencia Artificial, asegurando las inversiones, el marco regulatorio y las condiciones de contorno necesarias para la adopción de estas tecnologías (el denominado “*ecosistema de excelencia*”), de forma que se pueden aprovechar las oportunidades y beneficios que la Inteligencia Artificial ofrece, así como un enfoque centrado en las personas y basado en valores para minimizar los riesgos de las mismas (el denominado “*ecosistema de confianza*”).

¹ <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/#>

Como continuación a este documento, la Comisión adoptó un plan coordinado en Inteligencia Artificial en 2019, que fue revisado en 2021 (COMISIÓN EUROPEA, 2021) en el que se marcan algunas de las acciones que los Estados Miembros de la Unión Europea y la Comisión deben llevar a cabo, de cara a materializar los objetivos anteriormente mencionados. Entre las acciones establecidas, se destaca un grupo de ellas que tiene como objetivo construir un liderazgo estratégico europeo en sectores de alto impacto, entre los que se incluye el sector público, tal y como se muestra en la Figura 1:

Figura 1. Construcción de un liderazgo estratégico europeo en áreas de alto impacto.



Fuente: Plan Europeo coordinado en Inteligencia Artificial. 2021.

En el plan coordinado se indica la relevancia de los Fondos de Recuperación y Resiliencia (RRF) y otros programas europeos para acelerar el despliegue.

3.- ¿CÓMO SE UTILIZA ACTUALMENTE LA INTELIGENCIA ARTIFICIAL EN EL SECTOR PÚBLICO?

En la sección anterior, hemos presentado las razones por las cuales el sector público debe beneficiarse de la promesa de la Inteligencia

Artificial, qué beneficios puede traer y con qué elementos cuenta, así como los riesgos, generales y específicos, que el uso de este conjunto de tecnologías presenta para el sector público. Sin embargo, es legítimo preguntarse: *¿Se usan estas tecnologías actualmente en el sector público europeo?*. Y, si es así, *¿Dónde y cómo?*. Asimismo, y de cara a la adopción masiva de las mismas por parte de los gobiernos, es clave el contar con marcos, guías y hojas de ruta que ayuden a las administraciones a desplegar proyectos de adopción exitosos.

Para responder a esas preguntas, el Centro Común de Investigación de la Comisión Europea (JRC en sus siglas en inglés), a través de su proyecto “AI Watch”², lanzó dos estudios, que presentaremos en esta sección. El primero de ellos (TANGI, 2022) intenta hacer un análisis del panorama de uso de la inteligencia artificial en la administración pública europea, sin pretender ser una cartografía sistemática y detallada de todos y cada uno de los usos de la Inteligencia Artificial por los distintos gobiernos europeos. Adicionalmente a este catálogo, el informe del JRC incluye un análisis de las estrategias nacionales para la adopción de la Inteligencia Artificial por parte de los diferentes estados miembros de la Unión Europea, así como un análisis detallado de ciertos casos de uso que se consideran de interés para las administraciones públicas europeas.

Junto con el informe, el JRC ha publicado como “open data” la lista de casos de uso identificados³⁴, para su análisis y reutilización por parte de administraciones e investigadores. La lista de casos contiene actualmente 686, que han sido recopilados a través de repositorios nacionales, fuentes públicas y encuestas realizadas a los distintos estados miembros. Según nuestro conocimiento, esta lista de casos es la mayor publicada en Europa hasta la fecha.

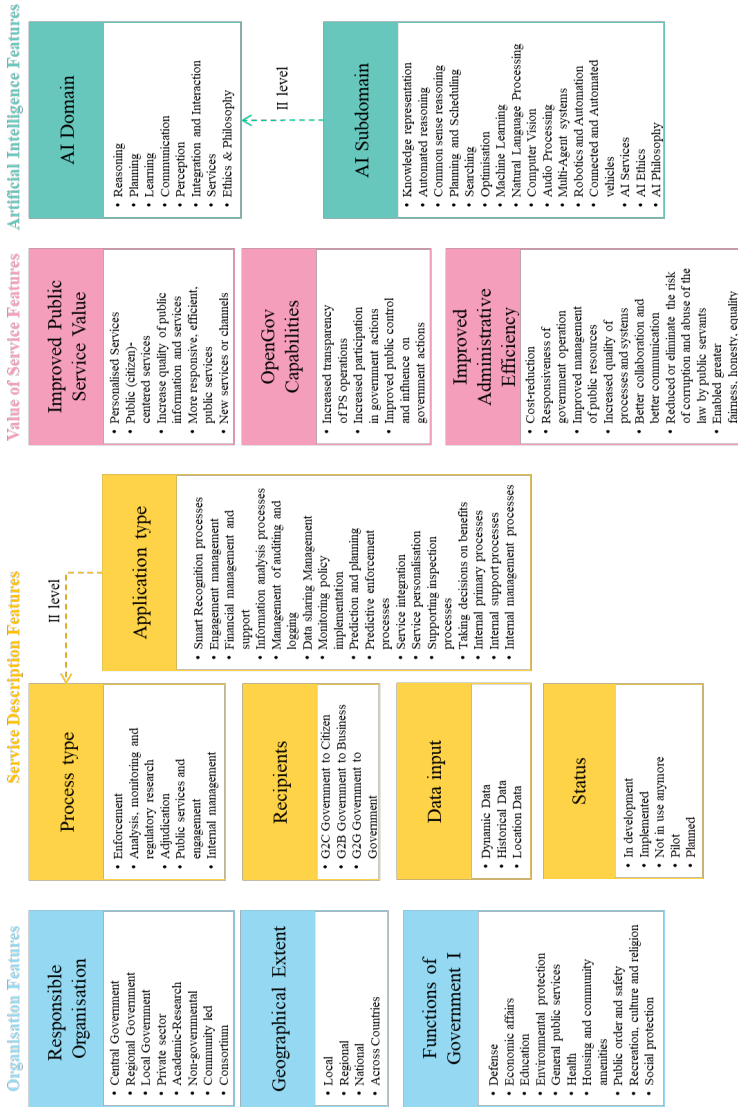
Para la clasificación de los casos de uso, se definió una taxonomía que permitiera categorizar los mismos en función de diferentes dimensiones, las cuales se presentan en la Figura 2:

² https://ai-watch.ec.europa.eu/index_en

³ <https://data.jrc.ec.europa.eu/dataset/7342ea15-fd4f-4184-9603-98bd87d8239a>

⁴ <https://ipsoeu.github.io/ips-explorer/>

Figura 2. Taxonomía para la clasificación de los casos de uso analizados por el JRC.



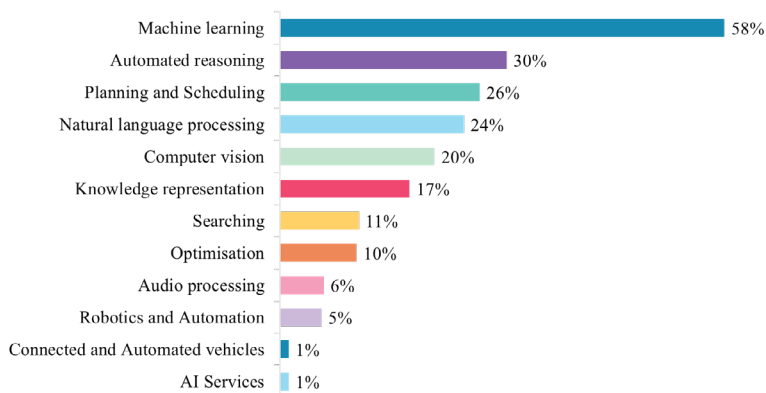
La citada taxonomía incluye los siguientes criterios:

- Criterios relativos a la organización encargada del proyecto, su ámbito geográfico y las funciones gubernamentales que desempeña.
- Criterios relativos al tipo de procedimiento administrativo que soporta, así como los destinatarios del mismo (ciudadanos, empresas o gobierno), así como el origen de los datos que se utilizan.
- Criterios relativos al valor de los servicios públicos a los que se da soporte.
- Criterios relacionados con la solución tecnológica escogida.

A lo largo de la presente sección, se van a presentar en más detalle algunas de las conclusiones obtenidas en el análisis de los casos de uso en función de los criterios anteriormente descritos.

La Figura 3 presenta la distribución de los casos de uso analizados en relación a las soluciones tecnológicas escogida:

Figura 3. Distribución de casos de uso en base a las soluciones tecnológicas utilizadas.



Fuente: JRC

Tal y como se puede ver en la Figura, la amplia mayoría de los casos analizados (58%) se basan en tecnología de *Machine Learning* o aprendizaje automático, un tipo de soluciones de Inteligen-

cia Artificial que permiten que los sistemas aprendan, predigan y se adapten a cambios sin requerir una programación detallada de las condiciones para la toma de decisiones. El aprendizaje automático es una de las aproximaciones de Inteligencia Artificial más clásicas y probadas y puede ser la razón detrás del elevado número de casos que usan esas tecnologías. Otro 30% de los casos usa soluciones de *razonamiento automático* (las cuales son capaces de inferir hechos a partir de conocimiento representado de varias formas).

La Figura 4 muestra la distribución de los casos analizados en función del estado del proyecto:

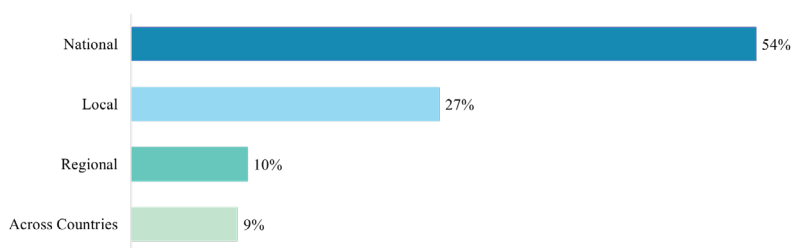
Figura 4. Distribución de casos de uso en base al estado del proyecto.



Fuente: JRC

La Figura muestra que un 38% de los casos se encuentran ya implementados y se usan en operaciones diarias del sector público. Este dato nos confirma que el uso de la Inteligencia Artificial en el sector público empieza a ser ya una realidad, aunque todavía hay un gran número de proyectos, bien en fase de pilotaje (30%), bien en desarrollo (25%). Cabe por último destacar que un 4% de los proyectos han dejado de utilizarse, bien porque el proyecto se descartó en su fase de pilotaje o bien porque la necesidad que el proyecto cubría desapareció.

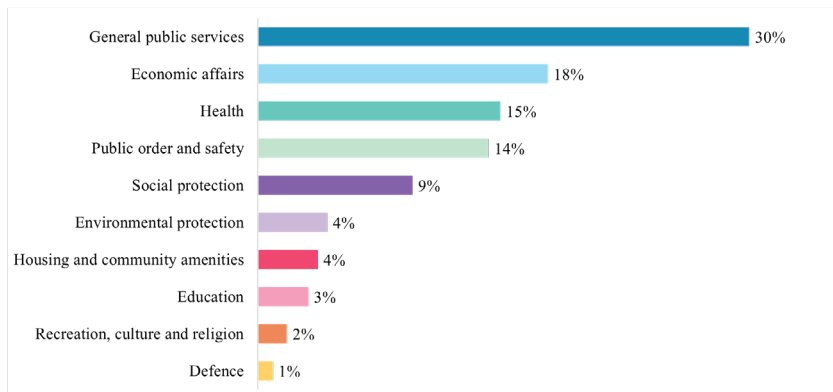
La Figura 5 muestra la distribución de casos según el ámbito de la administración que ha desplegado el proyecto:

Figura 5: Distribución de casos de uso en base al ámbito administrativo.

Fuente: JRC

En la citada Figura podemos observar cómo la mayoría de los casos analizados (un 54%) son proyectos desarrollados a nivel nacional. Una razón que puede explicar este resultado es el hecho de que actualmente los gobiernos nacionales disponen de mayores recursos para la realización de estos proyectos de innovación tecnológica y pueden actuar como “punta de lanza”. Sin embargo, aunque menor, el número de proyectos de las administraciones regionales y locales es también significativo, lo que muestra el interés y capacidad de todo tipo de administraciones a la hora de acometer proyectos de Inteligencia Artificial.

En la Figura 6 se muestra la distribución de los casos de uso en función de la función del gobierno a la que dan soporte:

Figura 6. Distribución de casos de uso en base a la función del gobierno a la que dan soporte.

Fuente: JRC

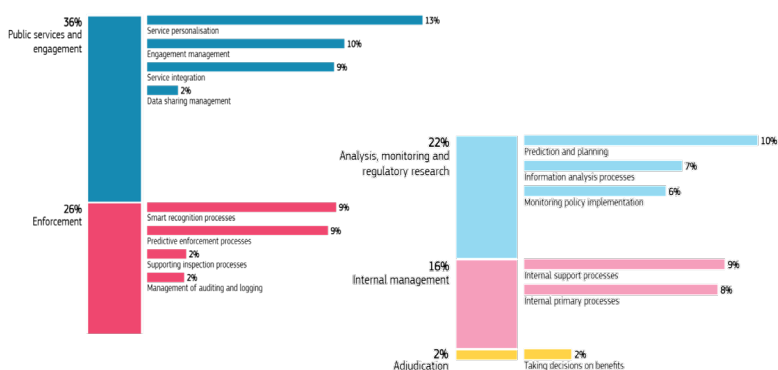
Tal y cómo se puede comprobar en la Figura 6, un tercio de los casos de uso están relacionados con los servicios públicos generales, seguido por un 18% dedicados a asuntos económicos, un 15% a salud y un 14% a orden público y seguridad.

Dentro de la categoría de servicios públicos generales podemos encontrar desde asistentes virtuales, tanto internos como externos, hasta sistemas de notificación y monitoreo, pasando por sistemas para la detección de desinformación, para la clasificación de documentos o para la detección del fraude. Es decir, todos aquellos servicios que permiten mejorar el funcionamiento de la administración general y su interacción con el ciudadano.

El significativo porcentaje de proyectos relacionados a asuntos económicos, salud y orden público, muestran el potencial de estas tecnologías de cara a la prestación de servicios públicos, así como la necesidad de contar con los marcos éticos adecuados que permitan controlar los riesgos mencionados anteriormente.

La Figura 7 muestra la distribución de los casos según el procedimiento administrativo al que dan soporte:

Figura 7. Distribución de casos de uso en base al procedimiento administrativo al que dan soporte.



Fuente: JRC

Como se puede observar en la Figura, el grupo de procesos más numeroso es el de servicios públicos relacionados con la vinculación con actores externos (36%), fundamentalmente ciudadanos y empre-

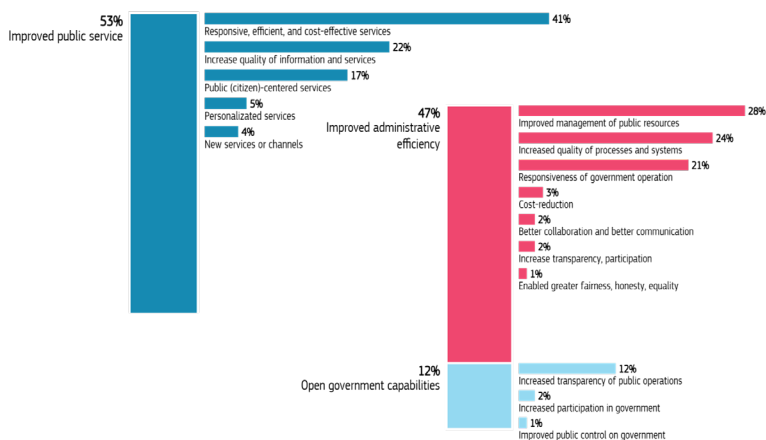
sas. Entre ellos, destacan soluciones relacionadas con la personalización de los servicios, lo que indica que uno de los usos de la Inteligencia Artificial tiene que ver con la adecuación de los servicios a las particularidades de los usuarios, buscando un enfoque más centrado en los mismos.

El segundo bloque es el de cumplimiento (26%), en el que vemos que la IA se utiliza para dar soporte al cumplimiento de las obligaciones públicas, incluyendo desde sistemas de auditoría hasta sistemas de reconocimiento biométrico. Un tercer grupo es el formado los sistemas de análisis y monitoreo (22%), que agrupa soluciones relativas a la realización de predicciones o a la monitorización de políticas públicas.

Cabe también destacar que el uso de Inteligencia Artificial para procedimientos de adjudicación (por ejemplo, para la asignación automática de beneficios sociales) es extremadamente bajo, lo que nos muestra que actualmente, el uso de la IA en los servicios públicos se centra muchos más en la automatización de tareas y la realización de predicciones que en la toma de decisiones automáticas.

Finalmente, la Figura 8 nos muestra la distribución de los casos en relación al valor aportado:

Figura 8. Distribución de casos de uso en base al procedimienti administrativo al que dan soporte.



Como podemos observar, la mayor parte de los casos (un 53%) se centran en la mejora de los servicios públicos, intentando proporcionar una mejor calidad, servicios más personalizados y centrados en el usuario o nuevos canales de interacción con la administración.

Una segunda categoría, también muy relevante es la de la mejora de la eficiencia administrativa (47%), en la que encontramos proyectos con el objetivo de mejorar la gestión pública, la eficiencia de los procesos o la reducción de costes. Por último, un 12% de los proyectos se centran en la mejora de las capacidades del gobierno, con el foco principal de incrementar la transparencia.

Además del análisis del panorama de uso de la inteligencia artificial en la administración pública europea es importante ayudar a las administraciones en la adopción de dichos casos de uso de una manera fiables. Para ello, el JRC ha publicado una propuesta de hoja de ruta para la adopción de proyectos de Inteligencia Artificial, elaborada en colaboración con representantes de los Estados Miembros de la Unión Europea (MANZONI, 2022) y centrada 16 recomendaciones en 4 áreas de intervención, tal y cómo se muestra en la Figura 9.

Figura 9. Áreas de intervención para la adopción de proyectos de Inteligencia Artificial por el sector público.

Area 1 Promote an EU value-oriented, inclusive, human-centric and trustworthy AI in the public sector 1.1 Harmonise and complement EU regulations to promote human-centric and trustworthy AI-enabled public services for all citizens. 1.2 Promote the adoption of ethical principles, the development of guidelines and the identification of mitigating measures to minimize the risks of deployment of AI by the public sector. 1.3 Develop and promote dedicated AI-enabled solutions based on co-creation approaches to increase relevance of services, and citizens' and businesses' trust and confidence in the use of AI by the public sector.	Area 2 Enhance coordinated governance, convergence of regulations, and capacity building 2.1 Create an EU-wide network of governance bodies for a streamlined management of AI in the public sector. 2.2 Design national and European capacity-building programs for public sector innovators aiming to develop and/or adopt AI in support of the digital transformation of public services. 2.3 Build upon and promote the use of regulatory sandboxes for public administrations, allowing experimentation of AI-enabled solutions in controlled environments. 2.4 Optimise funding in support of AI in the public sector to promote the spreading and scaling of reusable solutions. 2.5 Promote the development of multilingual guidelines, criteria, and tools for public procurement of AI solutions in the public sector throughout Europe.	Area 3 Build a shared and interactive AI digital ecosystem 3.1 Support multidisciplinary research and knowledge creation amongst European universities and Research and Development (R&D) institutions around AI for the public sector. 3.2 Build a common European Data Space for public sector bodies and their operators, drawing from the compilation of relevant AI datasets and related Registries throughout Europe. 3.3 Reinforce and advance existing initiatives on open data and interoperability. 3.4 Share reusable and interoperable AI components at all operational levels of European public administrations. 3.5 Create a European marketplace for GovTech solutions in support of public sector digital transformation.	Area 4 Apply and monitor sustainability through value-oriented AI impact assessment co-created frameworks. 4.1 Set up an EU observatory on AI, built on a pan-European network of AI national observatories to gather, share, and collectively manage best practices and experiences learned from different stakeholders in the public sector throughout Europe. 4.2 Develop and apply umbrella impact assessment frameworks based on key influencing factors to measure the use and impact of AI in the public sector. 4.3 Promote AI in the public sector in support of sustainability while developing sustainable AI, in compliance with environmental principles, and leveraging on civic engagement and participation.

Las 4 áreas de intervención son:

- **La promoción del uso en el sector público de una Inteligencia Artificial fiable, centrada en las personas y respetuosa con los valores Europeos**, con el objetivo de minimizar los riesgos específicos que el uso de estas tecnologías pueden traer a la administración.
- **Mejorar la gobernanza conjunta, la convergencia de los marcos regulatorios y de la necesaria capacidad para el desarrollo de soluciones** basadas en Inteligencia Artificial, con el objetivo de que Europa avance de manera coordinada en el despliegue de IA fiable.
- **Construir un ecosistema conjunto e interactivo de soluciones basadas en la Inteligencia Artificial**, que permita la reutilización, garantice la interoperabilidad y permita una libre circulación de datos respetando los derechos fundamentales de las personas.
- **Monitorizar la sostenibilidad del ecosistema a través de marcos de evaluación adecuados**, enfocados en la obtención de valor añadido, mediante observatorios y marcos de evaluación conjuntos.

4.- CONCLUSIONES Y RECOMENDACIONES

En las secciones anteriores hemos explicado qué es la Inteligencia Artificial, por qué es relevante para la transformación y mejora del sector público y hemos proporcionado una visión de alto nivel de cómo y dónde se está utilizando la Inteligencia Artificial en la administración europea en la actualidad. En esta sección, queremos presentar algunas conclusiones generales que se derivan de nuestro análisis y que pensamos que pueden ser de interés para los responsables públicos encargados de evaluar el uso de estas tecnologías dentro de la administración:

- **La Inteligencia Artificial es una realidad en el sector público.** Las organizaciones públicas deben empezar a considerar la Inteligencia Artificial no solo como algo relacionado con la

innovación y el desarrollo sino también como una alternativa real para la mejora de la maquinaria administrativa. Además, deberían comenzar a prepararse para un despliegue masivo de la misma en todas las áreas del sector público.

- **Es necesario disponer de conocimiento especializado interno.** Las organizaciones públicas deberían considerar el desarrollar conocimiento y capacidades internas relativas a la Inteligencia Artificial, que les ayuden al desarrollo de soluciones, a su adaptación y mantenimiento, así como a la adecuada compra pública de soluciones basadas en Inteligencia Artificial.
- **Es necesario incrementar el conocimiento general dentro de la administración.** El sector público debe considerar la Inteligencia Artificial como una tecnología que afectará al trabajo diario de la mayor parte de los empleados públicos, con lo que debe considerar la difusión masiva de un conocimiento mínimo que permita conocer el funcionamiento de los sistemas algorítmicos y la interpretación de sus resultados.
- **Buscar un equilibrio entre el conocimiento especializado interno y externo.** Muchas organizaciones públicas requerirán de ayuda para el desarrollo de sistemas basados en Inteligencia Artificial. Es fundamental que seleccionen de forma adecuada a socios y proveedores garantizando un equilibrio adecuado entre recursos internos y externos.
- **Necesidad de fiabilidad.** Cómo hemos comentado, la Inteligencia Artificial no está exenta de riesgos, los cuales deben ser evaluados en base a procedimientos estructurados y bien definidos, evitando usos discriminatorios e injustos que puedan afectar a la legitimidad de la administración. Adicionalmente, será necesario poner en marcha medidas de control que aseguren un uso de la Inteligencia Artificial centrado en las personas.
- **La tecnología es solo la mitad de la historia.** Las administraciones públicas deben saber que el esfuerzo técnico de desarrollar un sistema basado en Inteligencia Artificial es solo una pequeña parte del esfuerzo de implementación. Desplegar una solución basada en Inteligencia Artificial requiere gestión del cambio, conocimiento generalizado, así como nuevos roles y procedimientos en las organizaciones públicas.

Estas conclusiones iniciales que presentamos, fruto de nuestros primeros trabajos, nos dan una idea de la difícil tarea a la que se enfrenta el sector público, y que requerirá un esfuerzo coordinado por parte de los diferentes niveles administrativos, junto con el sector privado y la sociedad civil. Europa tiene que conseguir que esta sea su década digital, una década en la que las transiciones digitales y ecológicas nos lleven a una sociedad más próspera y justa. El uso adecuado de la Inteligencia Artificial será clave, y el papel del sector público fundamental.

BIBLIOGRAFÍA CITADA

- Martinez Plumed, F., Caballero Benítez, F., Castellano Falcón, D., Fernandez Llorca, D., Gomez Gutierrez, E., Hupont Torres, I., Merino, L., Monseerrat, C. and Hernández Orallo, J., AI Watch: Revisiting Technology Readiness Levels for relevant Artificial Intelligence technologies, EUR 31066 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-52328-4, doi:10.2760/495140, JRC129399.
- Samoili, S., Lopez Cobo, M., Delipetrev, B., Martinez-Plumed, F., Gomez Gutierrez, E. and De Prato, G., AI Watch. Defining Artificial Intelligence 2.0, EUR 30873 EN, Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-42648-6, doi:10.2760/019901, JRC126426.
- HLEG-AI, Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial (2019), Directrices Éticas para una Inteligencia Artificial fiable. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60423
- Comisión Europea. 2021. Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52021PC0206>
- Comisión Europea. 2018. Inteligencia Artificial para Europa. <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52018DC0237&from=EN>
- Comisión Europea. 2021. Fostering a European approach to Artificial Intelligence. <https://ec.europa.eu/newsroom/dae/redirection/document/75787>
- Comisión Europea, 2017. European Thematic Factsheet on Public Procurement. https://ec.europa.eu/info/sites/default/files/file_import/european-semester-thematic-factsheet-public-procurement_en_0.pdf
- Comisión Europea. 2021. The strategic use of public procurement for innovation in the digital economy. https://wbc-rti.info/object/document/21611/attach/KK0221251ENN_en_1_.pdf

- Tangi, L., Van Noordt, C., Combetto, M., Gattwinkel, D. and Pignatelli, F., AI Watch. European landscape on the use of Artificial Intelligence by the Public Sector, EUR 31088 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-53058-9, doi:10.2760/39336, JRC129301.
- Manzoni, M., Medaglia, R., Tangi, L., Van Noordt, C., Vaccari, L. and Gattwinkel, D., AI WatchRoad to the adoption of Artificial Intelligence by the Public Sector: A Handbook for Policymakers, Public Administrations and Relevant Stakeholders, EUR 31054 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-52131-0, doi:10.2760/693531, JRC129100.

SEGUNDA PARTE:
RÉGIMEN JURÍDICO

Sección primera:
Algoritmos y Derecho

Capítulo III

La naturaleza jurídica de los algoritmos

ANTONIO DAVID BERNING PRIETO¹

*Profesor Ayudante Doctor de Derecho Administrativo
Universidad Pablo de Olavide, de Sevilla*

Resumen: El presente trabajo pretende realizar un acercamiento a la posible naturaleza jurídica que ostentan los algoritmos en nuestro ordenamiento jurídico, partiendo de la base de diversas posiciones doctrinales existentes en la actualidad e intentando poner de manifiesto la problemática que subyace y las posibles soluciones que se proponen para dotar de un marco jurídico adecuado y suficientemente garantista a esta tecnología que se encuentra al servicio de la sociedad, pero cuya utilización por parte de los poderes públicos puede provocar inseguridad jurídica a los ciudadanos si no se dota de una regulación adecuada.

Abstract: The present work aims to make an approach to the possible legal nature that algorithms have in our legal system, starting from the basis of various doctrinal positions currently existing and trying to highlight the underlying problems and the possible solutions that are proposed for provide an adequate and sufficiently guaranteeing legal framework for this technology that is at the service of society, but whose use by public authorities can cause legal insecurity for citizens if it is not provided with adequate regulation.

Palabras-clave: algoritmo, administración pública, decisiones automatizadas, actividad administrativa automatizada, datos masivos.

Key-words: algorithm, public administration, automated decisions, automated administrative activity, big data.

SUMARIO.- 1.- INTRODUCCIÓN. 2. CONCEPTUALIZACIÓN DE LOS ALGORITMOS Y DISTINCIÓN DE FIGURAS AFINES. 3. POSICIONES DOCTRINALES EXISTENTES. 4. LA VERDADERA PROBLEMÁTICA QUE SUBYACE: LA NECESIDAD DE DOTAR DE GARANTÍAS JURÍDICAS SUFICIENTES AL USO DE ALGORITMOS POR PARTE DE LA ADMINISTRACIÓN. 4.1.- La única solución posible: dotar a los algoritmos de un régimen jurídico adecuado. 4.2.- ¿Deben los algoritmos formar parte de la motivación de los actos administrativos y, por tanto, constar expresamente en éstos? ¿y en el expediente administrativo? 4.3.- Del control de las decisiones adoptadas por los poderes públicos mediante el uso de algoritmos predictivos. 4.4.- De los diversos sesgos a que pueden dar lugar la actividad administrativa al hacer uso de los algoritmos. 4.4.1.- Posibles sesgos derivados del empleo de algoritmos en el procedimiento administrativo. 4.4.2.- Posibles sesgos derivados del empleo de algoritmos en la prestación de servicios públicos. 4.4.3.- Posibles sesgos en usos algorítmicos relacionados con el ejercicio de las

¹ ORCID: 0000-0002-7644-8210

potestades de investigación o comprobación de las administraciones públicas. 4.5.- De la necesidad de prever, en las licitaciones públicas, la cesión en favor de la Administración de los derechos de propiedad intelectual o industrial inherentes a los algoritmos que se desarrollen o se utilicen. 5.- CONCLUSIONES. BIBLIOGRAFÍA CITADA.

1.- INTRODUCCIÓN

En la era de la información, las nuevas tecnologías han traído consigo no sólo una modernización de la industria y la sociedad en general, sino que el modo de tomar decisiones ha cambiado drásticamente.

Entre otras tecnologías disruptivas, una de las que más relevancia tiene en la actualidad es la inteligencia artificial, que se nutre de datos masivos o *big data*, como es conocido, para la toma de decisiones en una gran variedad de materias y supuestos: desde la determinación de los ámbitos en los que se debe llevar a cabo una mayor labor investigadora de naturaleza fiscal, hasta la asignación de una plaza en un centro escolar público, pasando por el estudio de riesgos por parte de las entidades bancarias en el otorgamiento de préstamos hipotecarios.

Y es que las empresas, hasta tiempos no tan lejanos, tomaban sus decisiones sobre la base de criterios adoptados exclusivamente por personas que prestaban sus servicios en las mismas; no obstante, en la actualidad tales personas han sido relegadas a un segundo plano, de forma que la decisión se adopta por máquinas, y las personas únicamente intervienen para aplicar factores correctores que permitan, en determinados casos y bajo determinadas condiciones, reconducir posibles desviaciones algorítmicas. Y digo algorítmicas porque, en definitiva, todas esas máquinas se basan, en definitiva, en la aplicación a través de *software* (apoyado en un necesario *hardware* suficientemente potente) de unas prescripciones previamente dadas por un humano o, nuevamente, por una máquina, si se trata de datos obtenidos gracias a inteligencia artificial.

Como puede observarse, existe una serie de conceptos que se cruzan y pueden llegar a ser objeto de confusión, de forma que resulta esencial comenzar efectuando una breve pero necesaria distinción entre los mismos para conseguir adentrarnos en el fascinante mundo de los algoritmos sin errar en el intento.

2.- CONCEPTUALIZACIÓN DE LOS ALGORITMOS Y DISTINCIÓN DE FIGURAS AFINES

Un “algoritmo” puede ser definido, siguiendo el Diccionario de la Real Academia Española², como un “conjunto ordenado y finito de operaciones que permite hallar la solución de un problema”, o como “método y notación en las distintas formas del cálculo”.

Algunos organismos han proporcionado también sendas definiciones del algoritmo, como la de la Comisión de Garantía del derecho de acceso a la información pública de Cataluña, en la reclamación 123/2016, de 21 de septiembre, en la que lo definió como “procedimiento de cálculo que consiste en cumplir un conjunto ordenado y finito de instrucciones con unos datos especificados para llegar a la solución del problema planteado”, supuesto en que se solicitaba el acceso al algoritmo matemático aplicado para determinar los miembros de los tribunales de las pruebas de acceso a la universidad, caso en que se ordenó a la Administración proporcionar dicho acceso en lenguaje matemático y natural, de forma que fuese inteligible, en tanto en cuanto el algoritmo en este caso debía hacer uso de los datos o variables proporcionadas por el sistema en ejercicio de una potestad reglada, en ningún caso discrecional (en cuyo caso resulta por ejemplo prohibido adoptar un acto automáticamente en otros ordenamientos jurídicos, como el alemán³), y por tanto nada obsta, para la Comisión, a otorgar acceso al mismo a la persona interesada que lo requiere.

El código fuente, por su parte, fue definido por el Consejo de Transparencia y Buen Gobierno de la Administración General del Estado, en su resolución 701/2018, de 18 de febrero de 2009, como “el archivo o conjunto de archivos que tienen un conjunto de instrucciones muy precisas, basadas en un lenguaje de programación, que se utiliza para poder compilar los diferentes programas informáticos que lo utilizan y se puedan ejecutar sin mayores problemas. Los usuarios pueden usar el software sin mayores preocupaciones gracias a una interfaz gráfica sencilla que se basa en el desarrollo del código fuente”.

² Versión *online* accesible a través de la dirección web: <https://dle.rae.es/algoritmo?m=form>

³ Así lo dispone el §35 VwVfG.

Atendiendo a su definición, un algoritmo puede servir a muchas finalidades diversas (y expresarse de distintas formas: diagramas de flujo, el pseudocódigo y el propio lenguaje natural, como advierte LÓPEZ PULIDO⁴), no sólo en el ámbito de las matemáticas o la informática, donde tiene su origen⁵ como en el de la vida social, como sucede con las redes sociales, máximo exponente de su continuo empleo para ofrecer información acorde a nuestros intereses, entre otros, según nuestro perfil y experiencia de usuario. Nótese cómo plataformas de *streaming* actuales como Netflix® emplean algoritmos (en este caso predictivos, que citaremos posteriormente) para ofrecer a sus clientes recomendaciones adaptadas a sus preferencias de visionado, llegándose a calificar por CARRILLO⁶ como “entretenimiento de algoritmo”.

Estas funcionalidades aparecen también en la aplicación de algoritmos al Derecho Administrativo. Y es que, teniendo en cuenta que el algoritmo se emplea en la informática para establecer las reglas de funcionamiento de los programas o software, ninguna duda cabe de que, al utilizar aplicaciones informáticas en la tramitación de procedimientos administrativos, se están utilizando algoritmos bien como auxilio en las tomas de decisiones, bien para tomar las decisiones correspondientes.

No obstante, de acuerdo con la opinión de HUERGO⁷, aunque se puede decir que los procedimientos administrativos o judiciales obedecen a reglas algorítmicas (teniendo en cuenta la definición de

⁴ LÓPEZ PULIDO, J.P. (2021): “El derecho de acceso a los algoritmos y al código fuente, en la actuación administrativa automatizada y su afectación en las entidades locales”, *La Administración Práctica* n°. 11/2021, Ed. Thomson-Reuters Aranzadi (BIB\2021\5197), p. 6.

⁵ El origen de los algoritmos se atribuye generalizadamente al matemático y astrónomo *Muhammad ibn Mūsā al-Khwārizmī*, autor del primer Tratado de Álgebra cuya traducción *Algoritmi de numero Indorum* dio lugar al término actual de “algoritmo” (puede consultarse su biografía en la *Encyclopedia Britannica*, accesible online en la dirección web: <https://www.britannica.com/biography/al-Khwarizmi>).

⁶ CARRILLO BERNAL, J. (2019): *Paradigma Netflix: entretenimiento de algoritmo*, Ed. UOC, Barcelona, pp. 178 ss.

⁷ HUERGO LORA, A. (2020): “Una aproximación a los algoritmos desde el Derecho Administrativo”, en HUERGO LORA, A. (Dir.): *La regulación de los algoritmos*, Ed. Thomson-Reuters Aranzadi, Cizur Menor (Navarra), 1ª ed., p. 27.

algoritmo), al componerse de pasos con un contenido determinado que deben seguirse para llegar a la resolución del procedimiento, no lo hacen de forma idéntica a un programa informático, toda vez que el instructor tiene una vinculación jurídica distinta de la vinculación mecanicista que sigue el segundo, ya que mientras que el programa tendrá que seguir todos los pasos preestablecidos, el instructor podría omitir alguno, sin perjuicio de las consecuencias en la validez o eficacia del acto que ello pudiese conllevar.

Por tanto, para el autor, al no sustituir el algoritmo en ningún caso a la norma, sino automatizar su aplicación, en caso de error o defecto que origine un resultado incorrecto que se materializa en un acto administrativo por estar mal configurado el algoritmo, se originará la invalidez del mismo, en tanto en cuanto prevalece la norma sobre aquél.

No obstante, deben distinguirse en este punto los algoritmos predictivos de los que no lo son. Estos últimos obedecen a un patrón previamente establecido, de forma que sustituyen en la práctica la actividad mecánica que anteriormente realizaba un funcionario manualmente. Por tanto, realmente automatizan procedimientos cuyo esquema estaba previamente establecido normativamente, de forma que no existe ninguna innovación real en la práctica.

Por su parte, los algoritmos predictivos obedecen a una finalidad distinta, toda vez que no sólo pueden automatizar procedimientos, sino que pueden llegar a incorporar novedades y, por tanto, innovar, ya que, a partir de datos obtenidos por el sistema, el algoritmo es capaz de adaptarse a los mismos en función de diversas variables. Esta posibilidad, en determinados casos, podría hacer plantearnos si realmente están innovando el ordenamiento jurídico y, por ende, adquiriendo naturaleza normativa.

Esta aplicación de los algoritmos lleva tiempo utilizándose, especialmente en el ámbito tributario, por ejemplo, para la detección de irregularidades fiscales o en la presentación de declaraciones relativas a impuestos, si bien, como indican autores como BULL⁸, el riesgo radica en el empleo del *Big Data* o datos masivos puesto que son la base

⁸ BULL, H.P. (2017): *Der vollständig automatisiert erlassene Verwaltungsakt. Zur Begriffsbildung und rechtlichen Einhegung von E-Government*, DvBl 2017 (Ausgabe 7), p. 416.

de las predicciones basadas en datos que, posteriormente, se emplean en la generación automática de actos administrativos.

Este tipo de algoritmos, los predictivos, son un tipo específico de algoritmo que están concebidos y por tanto configurados para procesar datos en aras a la consecución de una finalidad concreta: obtener resultados ajustados a las necesidades del creador del mismo basados en datos masivos previamente puestos a su disposición.

Y es que, como apunta PONCE SOLÉ⁹, los algoritmos han pasado de ser estáticos (no predictivos) a ser dinámicos, pudiendo distinguirse en esta dinamicidad aquellos de aprendizaje automático o *machine learning*, que tienen capacidad de aprender con el tiempo y la experiencia, dando lugar a soluciones distintas a las inicialmente programadas (pero con cierta trazabilidad) y los de aprendizaje profundo o *deep learning*, que emulan redes neuronales complejas a partir de datos masivos cuya trazabilidad prácticamente es inexistente. Es por ello que devienen opacos, según el autor, incluso para sus creadores, a los que resulta imposible conocer cómo ha sido adoptada una decisión, sea correcta o errónea.

En definitiva, como apunta HUERGO¹⁰, las predicciones algorítmicas han llegado a sustituir los juicios de valor individuales emitidos por expertos, ya que tienen una capacidad de adaptarse a los cambios sociales mucho más alta que las personas, motivo por el cual su mayor ámbito de aplicación es en el sector privado, como los que utilizan las aseguradoras para calcular los índices de siniestralidad en aras a establecer la cuantía de las primas de los seguros, las entidades bancarias para analizar riesgos en la concesión de préstamos hipotecarios, o tiendas *online* para ofrecer automáticamente determinados productos a sus clientes.

La distinción fundamental entre algoritmos predictivos y no predictivos tiene especial relevancia en cuanto a su conceptualización y régimen jurídico, ya que las garantías que cada uno de ellos debe proporcionar serán diferentes en cada caso y deberán adaptarse, entre

⁹ PONCE SOLÉ, J. (2019): “Inteligencia artificial, derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico”, en *Revista General de Derecho Administrativo* nº 50, Ed. Iustel, p. 7.

¹⁰ HUERGO LORA, A. (2020): “Una aproximación a...”, *op. cit.*, p. 43.

otros, a la posible incidencia que cada uno de ellos pueda tener sobre terceros, en este caso ciudadanos, al ser los principales destinatarios de la actividad de los poderes y servicios públicos en que se utilizan.

3.- POSICIONES DOCTRINALES EXISTENTES

En la actualidad existen fundamentalmente dos fuertes posiciones doctrinales que recientemente han tratado de dar respuesta a la cuestión relativa a la naturaleza jurídica de los algoritmos.

Una de ellas es la mantenida por BOIX, quién ha defendido que los algoritmos deben ser considerados reglamentos, otorgándoles por tanto las máximas garantías posibles en cuanto que su aprobación, de reconocerse esta categoría, debería estar sujeta a todas las reglas establecidas para la aprobación de disposiciones de naturaleza reglamentaria, y su control estaría igualmente sujeto al mismo régimen jurídico que los reglamentos.

La otra es mantenida por HUERGO, quién aboga por no conferir naturaleza reglamentaria a los algoritmos ya que, para el autor, el concepto de algoritmo es demasiado amplio como para generalizar la aplicación de las reglas propias de las disposiciones reglamentarias a los mismos y, además, entiende que tampoco limitaría el uso de algoritmos, sino más bien todo lo contrario.

Para sustentar su posición, este autor efectúa una lúcida distinción entre ambas tipologías algorítmicas, y a tal efecto afirma que los algoritmos no predictivos carecen de contenido innovador, por lo que no son normas jurídicas sino instrumentos para facilitar su aplicación (incluso puede que la propia norma prevea su utilización, por lo que los algoritmos realmente forman parte de la norma, al aplicarla); se aplican al ejercicio de potestades regladas, para preparar la ejecución de un acto administrativo reglado o uno de sus elementos (como ejemplifica el autor con los certificados de la Seguridad Social); no son decisiones automatizadas, ya que no es lo mismo usar algoritmos para obtener datos que posteriormente se tendrán en cuenta a la hora de adoptar una decisión, que la adopción automática de una decisión; si hay un error en la formulación del algoritmo y el resultado fuese contrario a Derecho, no se pueden tener en cuenta elementos adicionales; y por

último examina el derecho a la exhibición del algoritmo, aspecto al que se volverá posteriormente.

En cuanto a los algoritmos predictivos, señala que se pueden utilizar, entre otras, para las siguientes finalidades:

- Para decidir el inicio de un procedimiento administrativo, caso en el que se crearía contenido nuevo a partir de otros datos de elaboración propia, sirviendo para establecer objetivos de políticas públicas o finalidades de policía predictiva.
- Para establecer criterios de actuación en ámbitos de muy escasa programación normativa, como sucede con los actos de trámite o decisiones que no son expuestas a los ciudadanos.
- Para determinar el contenido de decisiones administrativas. Esta utilidad es la más delicada, toda vez que parte del ejercicio de una potestad atribuida a la Administración Pública por una norma, en aras a aplicar la normativa a un caso concreto. En el ejercicio de potestades regladas existe menor margen de maniobra, ya que al supuesto de hecho previsto por la norma se aplican las consecuencias previstas por la misma, mientras que en las discrecionales existe cierto margen de maniobra para la Administración. No obstante, en este último supuesto, de intervenir el algoritmo, nos encontraríamos con la dificultad adicional de control de la motivación del acto dictado en ejercicio de tales potestades discrecionales.

Concluye el autor afirmando que “los algoritmos son algoritmos y los reglamentos son reglamentos”¹¹, puesto que la Administración, para ejercer su potestad normativa debe hacerlo a través de la aprobación de una disposición reglamentaria, mientras que un algoritmo bien tiene una función meramente auxiliar, bien debe tener autorización normativa para ir más allá, por lo que no se puede permitir, a juicio del autor, que la Administración regule mediante algoritmos.

Sin embargo, esta cuestión nos lleva a plantearnos la verdadera naturaleza jurídica de los algoritmos, toda vez que, como ha puesto

¹¹ HUERGO LORA, A. (2020): “Una aproximación a...”, *op. cit.*, p. 66

de manifiesto CERRILLO¹², la propia tecnología puede llegar a ser un instrumento de regulación (carácter regulatorio), pues el propio desarrollo tecnológico condiciona el alcance de la normativa reguladora y aquéllos, en determinados casos, cuando son empleados por las administraciones públicas, pueden llegar a producir efectos jurídicos.

Otros autores como ARROYO¹³ mantienen igualmente su oposición al tratamiento de los algoritmos como normas reglamentarias, sobre la base de que no responden a las características de estas. En este sentido, para el autor un reglamento programa el comportamiento de sus destinatarios mediante determinaciones abstractas, por lo que si un instrumento lo hace a través de determinaciones concretas o no están reconocidas por el sistema jurídico vigente, no podremos encontrarnos ante un reglamento.

El autor cita como ejemplo determinadas decisiones administrativas habilitantes de las administraciones públicas con función regulatoria que no son reglamentos en tanto en cuanto tienen un objeto concreto, así como los actos materiales concretos con igual función reguladora que pueden crear restricciones sobre las libertades económicas a pesar de no producir efectos jurídicos sobre sus destinatarios al regular un determinado sector.

Para este autor, los algoritmos se asemejan más a instrumentos de *soft law*, toda vez que no afectan directamente a la validez de los actos administrativos en que se materializan; esto es, su incumplimiento puede generar efectos anulatorios indirectos (si por ejemplo la norma impone el uso de un algoritmo determinado y la Administración se separa del mismo en un caso concreto para dar cumplimiento al principio de igualdad, sin la motivación que exigía dicha norma), efectos interpretativos o resarcitorios (si la Administración se separa de la reiterada aplicación de un algoritmo determinado en el pasado, lo

¹² CERRILLO I MARTÍNEZ, A. (2019): “El impacto de la inteligencia artificial en el Derecho Administrativo: ¿nuevos conceptos para nuevas realidades técnicas?”, *Revista General de Derecho Administrativo*, Ed. Iustel, nº 50, p. 10.

¹³ ARROYO JIMÉNEZ, L. (2020): “Algoritmos y reglamentos”, *Blog Almacén de Derecho*, 25 de febrero de 2020, accesible en la dirección web: <https://almacenderecho.org/algoritmos-y-reglamentos> (última consulta efectuada el 16/06/2022).

cual sugiere la aplicación de la doctrina del precedente administrativo en determinados supuestos de empleo de algoritmos), pero no va a determinar por sí solo la invalidez de los actos en que se concrete porque, a juicio del autor, al no ser reglamentos, no disfruta de su inderogabilidad singular.

En definitiva, lo que para el autor resulta claro es que, constatada la función reguladora que los algoritmos pueden adoptar, deben quedar sujetos a un régimen jurídico adecuado a esa tarea, sin necesidad para ello de aplicar en bloque el régimen jurídico de los reglamentos, que habría de ser modificado y adaptado a esta realidad tecnológica. Será por tanto imprescindible tener en cuenta la función institucional que cumple el algoritmo, es decir, aquello para lo que va a ser usado el mismo, para someter la adopción de decisiones administrativas a las garantías debidas (incluso las tradicionalmente asociadas a los reglamentos), propias de un Estado democrático de Derecho, como son la transparencia, publicidad, explicabilidad, participación o estabilización¹⁴ a las decisiones adoptadas mediante el uso de algoritmos.

4.- LA VERDADERA PROBLEMÁTICA QUE SUBYACE: LA NECESIDAD DE DOTAR DE GARANTÍAS JURÍDICAS SUFICIENTES AL USO DE ALGORITMOS POR PARTE DE LA ADMINISTRACIÓN

4.1.- La única solución posible: dotar a los algoritmos de un régimen jurídico adecuado

De lo expuesto anteriormente se colige que el esfuerzo llevado a cabo por parte de la doctrina se dirige a dotar de garantías suficientes el empleo de los algoritmos por parte de las Administraciones Públicas y, a tal fin, se han llegado incluso a proponer tratarlas como

¹⁴ Resulta interesante la referencia que hace el autor a la posibilidad de que un ciudadano invoque el precedente administrativo o el principio de confianza legítima si la Administración se separa de su reiterada actuación anterior de aplicación sistemática de un algoritmo determinado provocando un perjuicio al ciudadano que legítimamente ha confiado en esta previa actuación del poder público.

disposiciones de naturaleza reglamentaria, para así conseguir dotar de un régimen jurídico estable y ya existente a esta nueva realidad sobreenvenida que se va extendiendo a un ritmo ciertamente vertiginoso.

No obstante, a pesar de tratarse sin duda de disquisiciones dignas de elogio, creo que, ante realidades nuevas, merece la pena detenerse a crear un nuevo escenario sobre el que el algoritmo pueda actuar de forma jurídicamente adecuada, dentro de ciertos límites, para lo cual deberá ser regulado de forma conveniente en función de la concreta tipología algorítmica ante el que nos encontremos. En caso contrario, puede que cometamos nuevamente el error que se verificó con la aprobación de las Leyes 39 y 40/2015, consistente en la adaptación de figuras existentes a una nueva realidad electrónica, las cuales ya advirtieron antes de su entrada en vigor autores como ALAMILLO¹⁵, dando lugar a forzar el ajuste del régimen jurídico de los algoritmos a una regulación preexistente como es la de los reglamentos, la cual está pensada para este tipo de disposiciones, no para algoritmos.

Por tanto, mi propuesta en este sentido sería el desarrollo de una regulación jurídica específicamente destinada a la regulación del uso de los algoritmos en el sector público, de forma que sea suficientemente garantista como para salvaguardar los derechos e intereses de ciudadanos e interesados, y se cumplan las exigencias en materia de publicidad y transparencia con este tipo de tecnología cuando se aplique en este ámbito. Dotar a los algoritmos en el sector público de un régimen jurídico específico sería, sin duda, la solución a los problemas que más se están dando en la actualidad, sin necesidad de hacer uso de esquemas preestablecidos para instituciones diferentes, como sería por ejemplo la aplicación a los algoritmos del marco normativo previsto para los reglamentos, siendo realidades sustancialmente diferentes que obedecen a un origen y finalidades totalmente distintos y con una casuística tan variada que impediría una aplicación uniforme a todos los algoritmos de las reglas propias de la aprobación y aplicación de disposiciones reglamentarias.

¹⁵ ALAMILLO DOMINGO, I. (2017): “La regulación de la tecnología: la superación del modelo papel como elemento de transformación digital innovadora”, en MARTÍN DELGADO, I. (Dir.): *La reforma de la administración electrónica: una oportunidad para la innovación desde el Derecho*, Ed. INAP, Madrid, p. 126.

De regularse los algoritmos de la misma forma que los reglamentos, las diferencias existentes entre los reglamentos y los distintos algoritmos existentes abogarían en la práctica por una clara prevalencia de las excepciones frente a la regla general de regulación de esta realidad, lo cual no es nada deseable en aras a la verificación del principio de seguridad jurídica.

A tal fin, se torna necesario categorizar los algoritmos en nuestro Derecho como lo que son, algoritmos, sin necesidad de intentar encajarlos en una de las instituciones preexistentes, dado que son algo distinto merecedor de una regulación jurídico-pública que solvente toda la problemática que subyace en torno a su aplicación por parte de los poderes públicos, en tanto en cuanto puedan tener incidencia en derechos de terceros, teniendo en cuenta además para ello no sólo las cuestiones jurídicas que surgen, sino también las cuestiones técnicas, por lo que la interdisciplinariedad en esta regulación es imprescindible, como ha tenido ocasión de apuntar CERRILLO¹⁶ para el caso de la inteligencia artificial la cual, en definitiva, necesita algoritmos para nutrirse y que los sistemas gestionen la información de la forma pretendida.

Todo ello quiere decir que el empleo de algoritmos no predictivos o meramente de automatización, podrán seguir las reglas propias de cada categoría algorítmica, de forma que observen las normas actuales en materia de actividad administrativa automatizada previstas en el artículo 41 LRJSP si se dan los presupuestos necesarios para ello (actividad exclusivamente por medios electrónicos sin intervención humana), mientras que todo uso de algoritmos predictivos o que sean susceptibles de extralimitarse a la mera aplicación de fórmulas previstas en la norma que establezca su empleo deberán regirse no sólo por dicho precepto (ya que, al responder a la definición de actividad administrativa automatizada, deben dar cumplimiento al artículo 41 LRJSP), sino además debería complementarse con esta normativa específica que se propone desarrollar que cuente con las garantías suficientes para los ciudadanos y respete las prescripciones que desde instancias europeas se están desarrollando, como se desprende del

¹⁶ CERRILLO I MARTÍNEZ, A. (2019): “El impacto de...”, *op. cit.*, p. 9.

reciente informe del grupo JRC de la Comisión Europea¹⁷, en el que se hace especial énfasis en la problemática que acaece con la opacidad de los algoritmos y en la necesidad de efectuar un control y reglamentación adecuadas en el ámbito del sector público, entre otras medidas que se proponen.

Dicha regulación deberá prever, con carácter de mínimos:

- Ámbito objetivo de actuación, teniendo en cuenta para ello la distinción entre algoritmos predictivos y no predictivos.
- Ámbito subjetivo de actuación: centrándose en las entidades del sector público.
- Determinación de los órganos competentes para la aprobación del algoritmo a emplear.
- Establecimiento de reglas de publicidad del algoritmo y el código fuente en que se inserta.
- Derechos y garantías derivadas del empleo de algoritmos, especialmente en el seno del procedimiento administrativo.
- Deberes de transparencia relacionados con el algoritmo.
- Dotar de preferencia en el uso de algoritmos de caja blanca frente a los de caja negra.
- Establecimiento, con carácter básico, de la necesidad de aprobación de los programas y aplicaciones que vayan a operar con la actividad administrativa automatizada, de forma que se pueda verificar la existencia de un control efectivo sobre los mismos por parte del órgano competente, al igual que sucede en el ámbito tributario¹⁸. En la actualidad se prevé en el artículo

¹⁷ MANZONI, M., MEDAGLIA, R., TANGI, L., VAN NOORDT, C., VACCARI, L. y GATTWINKEL, D. (2022), *AI Watch. Road to the adoption of Artificial Intelligence by the public sector*. Publications Office of the European Union, Luxembourg.

¹⁸ Así lo dispone el artículo 96.4 de la Ley 58/2003, de 17 de diciembre, General Tributaria: “los programas y aplicaciones electrónicos, informáticos y telemáticos que vayan a ser utilizados por la Administración tributaria para el ejercicio de sus potestades habrán de ser previamente aprobados por ésta en la forma que se determine reglamentariamente”, sin que exista una norma similar en la normativa vigente reguladora del procedimiento administrativo común.

3 RAE sólo para la Administración General del Estado, por lo que sería conveniente extender al resto su vigencia.

4.2.- *¿Deben los algoritmos formar parte de la motivación de los actos administrativos y, por tanto, constar expresamente en éstos? ¿y en el expediente administrativo?*

El deber de motivación de los actos administrativos en los supuestos regulados en el artículo 35 LPAC resulta fundamental, como apunta NAVARRO¹⁹, toda vez que, como indica la autora, constituye no solo un deber de la Administración sino también una garantía de los administrados, que da a conocer a través de ella las razones por las que se dictan los actos que afectan a sus derechos e intereses legítimos, adquiriendo especial relevancia para la defensa de ambos debido a la información que contiene en la notificación, la propia motivación, aunque, como advierte acto seguido, una notificación defectuosa no es causa para anular un acto administrativo alegando falta de motivación siempre que la información relativa a la misma se deduzca del propio expediente.

Por tanto, en los supuestos en que cobra virtualidad el deber de motivación de los actos administrativos, la motivación debe constar en el propio acto o derivarse del expediente administrativo. Con el uso de algoritmos, resulta evidente que, si nos encontramos en uno de los supuestos del artículo 35 LPAC anteriormente citado, la motivación deberá contener los elementos básicos que permitan conocer las razones por la que se ha adoptado una determinada decisión administrativa. En caso contrario, dichas razones deberán desprenderse del expediente administrativo.

Ello obliga a examinar lo que el artículo 70.2 LPAC considera que forma parte del mismo, a saber: documentos, pruebas, dictámenes, informes, acuerdos, notificaciones y demás diligencias deban integrarlos, así como un índice numerado de todos los documentos que contenga cuando se remita y copia electrónica certificada de la resolución adoptada.

¹⁹ NAVARRO GONZÁLEZ, R.M. (2017): *La motivación de los actos administrativos*, Ed. Thomson-Reuters Aranzadi, 1ª ed., Cizur Menor -Navarra-, pp. 200 ss.

Por su parte, el artículo 70.3 LPAC contiene una delimitación negativa de todo aquello que no forma parte del expediente, concretamente toda información que tenga carácter auxiliar o de apoyo, como la contenida en aplicaciones, ficheros y bases de datos informáticas, notas, borradores, opiniones, resúmenes, comunicaciones e informes internos o entre órganos o entidades administrativas, así como los juicios de valor emitidos por las Administraciones Públicas, salvo que se trate de informes, preceptivos y facultativos, solicitados antes de la resolución administrativa que ponga fin al procedimiento.

Dicha regulación debe complementarse con el artículo 70.1 LPAC, el cual dispone que el expediente administrativo comprende el conjunto ordenado de documentos y actuaciones que sirven de antecedente y fundamento a la resolución administrativa, así como las diligencias encaminadas a ejecutarla.

Del tenor literal del precepto puede concluirse, en relación con la integración del algoritmo en el expediente administrativo, es que, en principio, la información contenida en aplicaciones se considera de carácter auxiliar o de apoyo por el legislador, de forma que, *a priori*, la Administración no está obligada a incorporarlo al expediente; no obstante, si el algoritmo ha servido de antecedente a la adopción de la resolución administrativa sí deberá incorporarse²⁰.

Sin embargo, dicha consideración sería correcta si nos encontrásemos ante algoritmos de naturaleza no predictiva, o ante la mera aplicación de fórmulas algorítmicas que se encuentran reguladas en la norma y la aplicación únicamente efectúa el cálculo correspondiente, facilitando la tarea (actividad administrativa automatizada *stricto sensu*), caso en que se deberán observar los requisitos recogidos en el artículo 41 LRJSP. Pero, si el algoritmo es de naturaleza predictiva, no será tan sencillo, y en tal caso la exclusión del expediente administrativo no resulta adecuada en ningún caso, en tanto en cuanto el destinatario de la resolución administrativa tiene derecho a la explicabilidad de las razones que han llevado a adoptar una decisión, haya

²⁰ Ello obligaría incluso a replantearse la procedencia de incorporar el algoritmo al archivo electrónico, toda vez que pueden darse supuestos en que el propio algoritmo tenga un valor histórico y deba ser objeto de conservación a largo plazo.

sido una persona humana o un algoritmo el que haya llegado a la conclusión correspondiente, es decir, la motivación.

Y esto trae como consecuencia que el propio interesado tenga derecho al acceso al expediente administrativo tanto durante la sustanciación del procedimiento (artículo 53.1.a LPAC) como en el seno del proceso judicial (artículo 48 LJCA), por lo que la falta de previsión por parte del legislador español de esta circunstancia al regular el expediente administrativo ha dado lugar a serios planteamientos en la práctica, como el conocido caso en que la Fundación CIVIO solicitó el acceso al código fuente del programa BOSCO, empleado por las empresas energéticas en aras a determinar los beneficiarios del bono social.

En el caso mencionado, se parte de la base de que, mediante Resolución de 15 de noviembre de 2017, de la Secretaría de Estado de Energía del Ministerio de Energía, Turismo y Agenda Digital se puso en funcionamiento la aplicación telemática que permite al comercializador de referencia comprobar que el solicitante del bono social cumple los requisitos para ser considerado consumidor vulnerable y, por tanto, beneficiario del citado bono social.

La Fundación CIVIO detectó que existían errores en la aplicación que daban lugar a que, en la práctica, a personas que tenían derecho al bono social, no les fuese reconocido, motivo por el que solicitó a la Administración poder acceder al código fuente de la aplicación y, por ende, al algoritmo que utiliza, para conocer los motivos que daban lugar a la denegación de la condición de determinadas personas como consumidores vulnerables y, por tanto, se les denegaba el derecho al bono social.

Dicha solicitud presentada en septiembre de 2018 requería, en relación con la aplicación, su especificación técnica, resultado de las pruebas efectuadas y su código fuente, sobre la base de la normativa vigente en materia de transparencia. Tuvo que ser recurrida ante el Consejo de Transparencia y Buen Gobierno en noviembre posterior²¹, siendo estimada parcialmente en febrero de 2019, si bien el Consejo consideró que el acceso al código fuente se encuentra limitado por el artículo 14.1.j) de la 19/2013, de 9 de diciembre, de Transparencia,

²¹ Procedimiento nº R/0701/2018.

acceso a la información pública y buen gobierno (LT, en adelante), relativo a la propiedad intelectual, por lo que permitió el acceso a la especificación técnica y al resultado de las pruebas, pero no al código fuente de la aplicación.

Ante la negativa de acceso al código fuente de la aplicación, se interpuso recurso contencioso-administrativo ante el Juzgado Central de lo Contencioso-Administrativo, al entender que en dicho código se encuentra la motivación real del acto administrativo dictado, que en este caso concede o deniega el bono social solicitado. El Juzgado dictó sentencia el 30 de diciembre de 2021²² desestimando las pretensiones de la Fundación sobre la base de que el acto administrativo no es dictado por una aplicación, sino por un órgano administrativo, pudiendo ser impugnado por el interesado de estimarlo conveniente. Lo que el juzgador en este supuesto concluye, es que la aplicación se utiliza en una fase del procedimiento que no es la terminación del mismo, por lo que lo considera un acto de trámite y, al efecto, incluso cita el artículo 75.1 y 2 LPAC, en tanto en cuanto regula los actos de instrucción del procedimiento dictados haciendo uso de aplicaciones informáticas.

No obstante, recientemente se observa un cambio de tendencia sobre el particular, ya que dicha sentencia ha sido recurrida en apelación por la citada Fundación ante la Audiencia Nacional, y el propio Consejo de Transparencia y Buen Gobierno ha virado en su posición, al presentar éste ante la sala de lo contencioso-administrativo un escrito de no oposición al recurso de apelación (que previamente la Abogacía del Estado había formulado en representación de aquél), al entender que la sentencia recurrida no resulta ajustada a Derecho. Habrá de seguirse el *íter* procesal de este asunto para conocer la evolución interactiva sobre el particular.

Volviendo a la sentencia ahora recurrida en apelación, en este caso la aplicación se entiende utilizada instrumentalmente para dictar el acto administrativo posterior por parte del órgano competente para ello. Por otra parte, se hace referencia en dicho pronunciamiento judicial al informe del Subdirector General de Tecnologías de la Información y de las Comunicaciones del Ministerio de Industria, Comercio y Turismo, en el que se justifica la utilización de aplicaciones específicas

²² Sentencia nº 143/2021, dictada en el procedimiento ordinario 18/2019.

que no gozan de los beneficios de licencia o código abierto (*open source*), al no ser objeto de reutilización por el sector público (ni privado), en tanto en cuanto las competencias en materia de energía no están transferidas a las comunidades autónomas.

Adicionalmente, en el mismo informe se pone de manifiesto que la “seguridad por diseño” debe equilibrarse con la “seguridad por oscuridad”, y en este caso prevalece esta última por tratarse de materias relativas a seguridad pública y nacional, aspecto con el que podemos estar de acuerdo en determinados supuestos, pero no es posible generalizar sin provocar una generalizada inseguridad jurídica ante cualquier decisión administrativa que adopten los poderes públicos.

Debe tenerse en cuenta que el propio Instituto Nacional de Ciberseguridad²³, organismo dependiente de la Administración General del Estado, define la seguridad por oscuridad como un “concepto que pretende emplear el secreto de implementación de un dispositivo o programa; es decir, encubrir cómo está construido interiormente para evitar sufrir ataques o vulnerabilidades y tratar de aumentar así el nivel de seguridad”. No obstante, advierte que “esta técnica ha sido ampliamente discutida, demostrando que no es efectiva y que incluso es contraproducente, ya que pueden existir vulnerabilidades solo conocidas por unos pocos que permitirían romper la seguridad de lo que se pretende encubrir”.

Lo que resulta evidente es que, en aquellos supuestos en que verdaderamente se utilicen las aplicaciones como herramienta instrumental o auxiliar, podrá tratárselas como tales en determinados supuestos, siempre y cuando sea posible conocer la motivación de la decisión y recurrirla a pesar del uso del algoritmo, pero en aquellos otros en que dichos requisitos no se verifiquen, debería ser permitido el acceso al código fuente y, más concretamente, al algoritmo empleado.

²³ Glosario de términos de ciberseguridad. Instituto Nacional de Ciberseguridad. Accesible a través de la dirección web: https://www.incibe.es/sites/default/files/contenidos/guias/doc/guia_glosario_ciberseguridad_2021.pdf (última consulta efectuada el 24/06/2022).

4.3.- *Del control de las decisiones adoptadas por los poderes públicos mediante el uso de algoritmos predictivos*

Sin duda la posibilidad, por parte de los poderes públicos, de adoptar decisiones basados en modelos predictivos en lugar de hechos, a pesar de ser ciertamente discutible y criticado por la doctrina²⁴, es una realidad y, por tanto, debería ser objeto de regulación expresa para garantizar que los modelos predictivos están siendo correctamente aplicados y no se vulneran derechos de terceros. En este sentido, para HUERGO no debería ser posible, toda vez que una decisión administrativa debe basarse en conductas de las personas, y los algoritmos predictivos se basan en juicios probabilísticos.

Sin embargo, ciertamente descendiendo al plano práctico es una realidad que existen decisiones administrativas relativas a la prestación de servicios públicos que se basan en modelos probabilísticos. Piénsese que una decisión administrativa no siempre se plasmará en actos administrativos, aunque sea lo habitual como forma normal de poner fin al procedimiento administrativo, ya que existen decisiones que inciden en derechos de los ciudadanos que no se materializan en este tipo de actos, pero no por ello tiene menor incidencia en su esfera de derechos, como por ejemplo la utilización de algoritmos predictivos para determinar los pacientes que tienen más probabilidades de tener éxito en una operación, usando para el cálculo de dicha probabilidad datos masivos o *big data*, y, por tanto, se les confiere preferencia a la hora de ser citados e intervenidos.

En todos estos supuestos, tanto cuando se trate de una resolución administrativa dictada en el seno de un procedimiento administrativo, como cuando se trate de decisiones relativas a la prestación de servicios públicos con incidencia en derechos de los ciudadanos, uno de los mayores problemas en la utilización de los algoritmos en la adopción de decisiones por parte de los poderes públicos es, efectivamente, su control, tanto *ex ante* como *ex post*, toda vez que en su aprobación no se somete a publicidad y posteriormente no es posible acceder al código fuente del algoritmo ni a su funcionamiento interno real, motivo

²⁴ HUERGO LORA, A. (2020): “Una aproximación a...”, *op. cit.*, p. 86.

por el que a este tipo de algoritmos (normalmente de tipo predictivo) se le ha llegado a denominar “caja negra” o *black box*²⁵.

En este punto podrían distinguirse dos clases de usos o decisiones: algunas que no incidirían directamente en derechos o intereses de terceros, como la decisión de llevar a cabo determinadas actuaciones en ámbitos o sectores concretos (inspecciones, controles, etc.), más propias de la aplicación general de políticas públicas o planes de actuación, caso en el que la aplicación del algoritmo se mantiene en el ámbito interno de la Administración y no debe ser objeto de control externo (evidentemente, no sería lógico informar a determinados sujetos de las líneas de actuación inspectoras o de seguridad, por ejemplo), y otras que incidirían directamente en terceros, al tratarse de verdaderas decisiones de los poderes públicos con directa afectación de los interesados.

En estas últimas, actualmente la única vía de control sería recurrir la decisión administrativa, es decir, el acto administrativo resultante del procedimiento de que se trate, recurso en el que se podría discutir el algoritmo empleado, al igual que pueden cuestionarse informes técnicos que son tenidos en cuenta a la hora de dictarlo, para lo cual se podría hacer uso de una prueba pericial técnica. Ahora bien, incluso en este supuesto es imprescindible conocer el funcionamiento del algoritmo, al menos los criterios que utiliza a la hora de adoptar la decisión y las variables que tiene en cuenta para ello.

La propia Red de Derecho Administrativo e Inteligencia Artificial (Red DAIA, en adelante) ya concluyó en 2019 que son los derechos fundamentales y constitucionales los que exigen transparencia, audibilidad y justificación en el uso de algoritmos por los poderes públicos, toda vez que el derecho de acceso a la información, el debido proceso y garantías judiciales y administrativas o el derecho a la protección de datos incluyen el acceso a los algoritmos y a los sistemas en que se integran, a los datos que manejan y su calidad, especialmente los motivos y razones por las que se adopta una decisión que afecta a

²⁵ PASQUALE, F. (2015): *The black box society. The secret algorithms that control money and information*. Ed. Harvard University Press, Cambridge, Massachusetts, London, England, p. 4.

una o varias personas, evitando así la vulneración del principio constitucional de interdicción de la arbitrariedad²⁶.

4.4.- De los diversos sesgos a que pueden dar lugar la actividad administrativa al hacer uso de los algoritmos

La casuística puede ser muy variada en cuanto a los posibles sesgos a que pueden dar lugar las decisiones basadas en algoritmos. Como se ha examinado, la principal problemática subyace en la dificultad de acceder a la formación de la decisión, especialmente cuando se trata de algoritmos predictivos, y conocer por tanto cómo se ha adoptado una determinada decisión.

El propio gobierno de Reino Unido ha tenido ocasión de manifestar su posición al respecto²⁷, admitiendo que los algoritmos, al buscar y tratar patrones de datos, pueden producir decisiones defectuosas o sesgadas, discriminando desproporcionadamente a determinados grupos, al igual que puede ocurrir con decisiones humanas, motivo por el cual deben ser susceptibles de control, como se tuvo ocasión de afirmar anteriormente. De hecho, para controlar entre otras cosas este tipo de situaciones, se ha previsto en Reino Unido la creación del *Centre for Data Ethics & Innovation* (Centro de Ética de los Datos e Innovación) de corte estatutario al que se podrán efectuar consultas y que deberá vigilar la actividad pública en este sentido, lo que sin duda demuestra el interés que está suscitando en dicho país la adopción de decisiones basadas en algoritmos por parte de los poderes públicos.

²⁶ Conclusión 9 del I Seminario Internacional Derecho Administrativo e Inteligencia Artificial celebrado en la Universidad de Castilla-La Mancha el 1 de abril de 2019 (p. 4).

²⁷ Informe del Comité de Ciencia y Tecnología de la Cámara de los Comunes: *Algorithms in decisionmaking, Fourth Report of Session 2017-19*. Disponible en la dirección web: https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/351/35108.htm#_idTextAnchor087 (última consulta efectuada el 9/6/2022).

4.4.1.- Posibles sesgos derivados del empleo de algoritmos en el procedimiento administrativo

Cuando la norma reguladora del concreto procedimiento de que se trate regule las variables y cálculos necesarios para la toma de decisión, la verificación de la adecuación del resultado tras aplicar el algoritmo resultará más sencilla, al emplearse éste como medio auxiliar. Sin embargo, en relación a resoluciones administrativas que ponen fin a procedimientos administrativos, el problema puede ir más allá al tratarse de sistemas avanzados que utilizan algoritmos predictivos (y que, por tanto, los algoritmos son entrenados para inferir determinados modelos a partir de un conjunto de datos, a fin de determinar las acciones que se requieren para alcanzar un objetivo determinado y siguen aprendiendo mientras se utilizan²⁸), a veces el sistema de aprendizaje podría avanzar dando por válidas decisiones anteriores, de modo que podrían darse situaciones en la práctica equiparables a dictar un acto administrativo sobre la base de un acto previo inválido.

En estos casos, al igual que he tenido ocasión de afirmar en relación con los actos administrativos dictados en soporte electrónico²⁹, habrá de distinguirse si se trata de un supuesto de incomunicabilidad electrónica de la invalidez *ex* artículo 49 LPAC, limitando la extensión de la invalidez de un acto a los sucesivos que no dependan del mismo, o si realmente la relevancia de la decisión algorítmica debe dar lugar a la invalidez del acto, para lo cual deberá determinarse en primer lugar si el algoritmo incurre en algún tipo de ilegalidad y se ha contaminado el acto que lo aplica por haber sido esencial en el mecanismo decisonal previo.

Otro de los posibles problemas que pueden darse, y que han sido puestos de manifiesto por parte de la doctrina es la relativa a la competencia cuando las decisiones se adoptan automáticamente basándose en algoritmos, al tratarse de una máquina, y no una persona, la que adopta materialmente la decisión. El cuestionamiento de la

²⁸ Apartado C del Libro Blanco sobre Inteligencia Artificial de la Comisión Europea, COM(2020) 65 final, p. 20.

²⁹ BERNING PRIETO, ANTONIO D. (2019): *Validez e invalidez de los actos administrativos en soporte electrónico*, Ed. Thomson-Reuters Aranzadi, Cizur Menor -Navarra-, pp. 311 y ss.

teoría del órgano y del acto administrativo, en tanto en cuanto la tradicional conceptualización estaba siendo superada por la intervención tecnológica, fue puesta de manifiesto por autores como VALERO TORRIJOS³⁰, ALAMILLO o URIOS³¹, y posteriormente abordada en profundidad con la vigencia de la Ley 11/2007 por MARTÍN DELGADO³², para quién la actividad administrativa automatizada se encuentra perfectamente alineada tanto con la teoría del órgano como con la del acto administrativo.

En este sentido, el autor defiende, de un lado, que el hecho de que la decisión la adopte una máquina no quiere decir que no pueda imputarse a un órgano, ya que la autoría del acto recaerá sobre el órgano administrativo que ejerce la potestad y tiene encomendada la competencia *ex lege*; y, de otro, que un acto administrativo dictado por una máquina sigue respondiendo a su definición, esto es, si se cumple el supuesto de hecho previsto en la norma, se aplica la consecuencia jurídica en ella contenida y, mediante este simple proceso, se satisface el interés general, radicando simplemente la diferencia en que este proceso intelectual lo realiza una máquina en lugar de una persona, ya que la imputación de la voluntad al órgano realmente es una ficción, sea dictado automática o manualmente.

Dicha afirmación resulta claramente aplicable al ámbito de la actuación administrativa automatizada basada en algoritmos no predictivos, pero con el uso de los algoritmos predictivos es necesario detenerse a examinar si se mantiene la validez de la teoría del órgano y del acto administrativo. Y es que, como acertadamente apunta MARTÍN DELGADO³³, de la incorporación de nuevas tecnologías de este tipo a las administraciones públicas se derivan consecuencias,

³⁰ VALERO TORRIJOS, J. (2007): *El régimen jurídico de la e-Administración. El uso de los medios informáticos y telemáticos en el procedimiento administrativo común*, Ed. Comares, Granada, 2ª ed., pp. 72 y ss.

³¹ ALAMILLO DOMINGO, I. y URIOS APARISI, X. (2007): “Los límites a la utilización del sello del órgano por parte de las Administraciones Públicas en la Ley 11/2007”, Comunicación presentada a las *Jornadas Tecnimap*, Gijón, 27-30 de noviembre de 2007, p. 4.

³² MARTÍN DELGADO, I. (2009): “Naturaleza, concepto y régimen jurídico de la actividad administrativa automatizada”, *Revista de Administración Pública* n° 180, pp. 8 y ss.

³³ *Ibíd.*, p. 14.

pero no para los conceptos de órgano o acto administrativo (ni, añadiendo, las disposiciones reglamentarias, que no deben ver acrecentadas sus variantes, ya de por sí suficientemente numerosas), sino sobre su régimen jurídico; no obstante lo cual, quizás haya llegado el momento de repensar la teoría del órgano en el sentido de imputar la actividad basada en algoritmos a la propia entidad en lugar de a un órgano concreto de la misma.

En este sentido, se ha defendido en el presente trabajo la necesidad de dotar de un régimen jurídico adecuado al empleo de algoritmos predictivos por parte de la administración pública, ya que si bien existen unos requisitos establecidos en cuanto a la actividad administrativa automatizada, éstos están orientados a aplicaciones preprogramadas que realizan las operaciones indicadas (programadas) previamente, pero no cuando tales aplicaciones aprenden con el tiempo y su actividad y llegan a tomar decisiones que pueden ser distintas a las inicialmente configuradas, como resultado de su propio aprendizaje.

En la actualidad, la utilización de algoritmos predictivos debe regularse de forma específica y dotarse de un elevado nivel de garantías, evitando así que se cercenen derechos y garantías procedimentales propias de nuestro ordenamiento jurídico-administrativo pues, de otro modo, se vulnerarían derechos de terceros. En este sentido, si bien resulta claro que los algoritmos se pueden aplicar tanto a actos reglados como de trámite, en el ejercicio de potestades regladas es más fácil entender que podría hacerse uso de los mismos, mientras que en las discrecionales no³⁴; sin embargo, cabría plantearse si es precisamente en el ejercicio de estas últimas en las que podría desplegar toda su virtualidad práctica, al poder dotar de la máxima objetividad decisiones que en la actualidad podrían estar subjetivizadas, siempre y cuando pudiesen motivarse adecuadamente.

³⁴ Así lo establecen otros ordenamientos jurídicos, como el alemán, como se desprende del §35 VwVfG.

4.4.2.- Posibles sesgos derivados del empleo de algoritmos en la prestación de servicios públicos

Otra de las cuestiones fundamentales relativas al uso de algoritmos es su empleo en la prestación de servicios públicos por parte de las administraciones públicas. La actividad de servicio público puede definirse con BERMEJO VERA³⁵ como aquellas actividades o servicios indispensables y esenciales para la comunidad de titularidad de la administración pública (sin perjuicio de su posible prestación a través de terceros como forma de gestión indirecta de servicios públicos, como por ejemplo la concesión) y que se presta de forma regular y continua.

De los múltiples servicios públicos existentes, uno de los más importantes es el sanitario, que se prestan bien de forma directa (como regla general, a través de centros de titularidad pública), bien de forma indirecta³⁶, a través de la colaboración de centros privados en la atención sanitaria por medio de un concierto para la prestación de determinados servicios o la vinculación de hospitales privados a la red pública mediante convenio³⁷, estando el centro (a pesar de mantener la titularidad privada) en este último caso sometido a las mismas condiciones, inspecciones y controles que un hospital de titularidad pública.

Los algoritmos en el ámbito sanitario no son tan ajenos a la realidad como cabría pensar, de hecho, se llevan años empleando. Sin embargo, los algoritmos utilizados hasta la fecha han sido simples, esto es, no predictivos, y ordenados a establecer pautas de conducta ante determinadas situaciones, esto es, los pasos que deben darse ante supuestos de dolor torácico, ante personas inconscientes o en un accidente de tráfico. Sin embargo, como se ha ido comentado a lo largo del presente trabajo, la inteligencia artificial está adentrándose en este

³⁵ BERMEJO VERA, J. (2009): Derecho Administrativo. Parte Especial. Ed. Thomson-Civitas, 7ª ed., Cizur Menor -Navarra-, p. 63

³⁶ Se han creado nuevas formas de colaboración indirecta en aras a los principios de eficacia, celeridad, economía y flexibilidad, si bien no es objeto del presente trabajo y por tanto nos circunscribiremos a las principales y más amplia y tradicionalmente admitidas.

³⁷ *Ibíd.*, p. 197

sector de forma rápida, y se están empezando a emplear algoritmos más avanzados en este ámbito, ya que permite agilizar la adopción de decisiones en cuanto a la prestación de servicios públicos con los ciudadanos.

Además, dada la robotización y automatización que se está produciendo en este sector³⁸, ninguna duda cabe de que las máquinas acabarán empleando algoritmos predictivos incluso en las intervenciones quirúrgicas, toda vez que, si son capaces de llevar a cabo una operación, también lo serán para aplicar toda una base de datos de experiencias previas a la hora de tomar una decisión acerca del mejor ángulo para diseccionar un tumor o cómo se han tenido mejores resultados al cauterizar una arteria, por ejemplo.

En este tipo de supuestos será imprescindible conocer el funcionamiento de los algoritmos no sólo en aras a la transparencia y a la evolución de la propia ciencia, ya que interesará conocer a los propios facultativos médicos las razones que han llevado al robot a adoptar una u otra decisión en cada supuesto concreto, sino porque si hay un resultado lesivo, podría dar lugar a supuestos de responsabilidad patrimonial de la Administración, y habrán de probarse no sólo la existencia de una acción u omisión derivada de la prestación de un servicio público sanitario que provoca un resultado lesivo efectivo, evaluable, individualizable y antijurídico que deriva de un funcionamiento normal o anormal del mismo sino, muy especialmente, la relación de causalidad entre la acción u omisión y el daño causado, para lo cual será imprescindible acceder a la forma en que el robot ha adoptado las sucesivas decisiones durante la ejecución de la intervención.

³⁸ La primera laparoscopia ejecutada por robots sin intervención humana es ya una realidad cuya extensión será cuestión de tiempo únicamente: <https://hub.jhu.edu/2022/01/26/star-robot-performs-intestinal-surgery/> (última consulta efectuada el 13/06/2022). Se explica detalladamente por DE MOMI, E. Y SEGATO, A. (2022): “Autonomous robotic surgery makes light work of anastomosis”, *Science Robotics*, nº 62, p. 7

4.4.3.- Posibles sesgos en usos algorítmicos relacionados con el ejercicio de las potestades de investigación o comprobación de las administraciones públicas

La potestad administrativa puede definirse, con GAMERO³⁹, como “un poder jurídico unilateral, atribuido conforme al principio de legalidad para la satisfacción de un interés general, sometido a la ley y a control jurisdiccional (judicial y constitucional)”. Sin necesidad de profundizar en un tema tan profuso y extenso como el de las potestades administrativas, por exceder el objeto del presente trabajo, debe remarcarse que una de las principales garantías de su ejercicio es su sometimiento a control jurisdiccional, atribuyéndose por la Ley 29/1998, de 13 de julio, reguladora de la Jurisdicción Contencioso-administrativa (también LJCA) al orden contencioso-administrativo, tanto si la potestad es ejercida por Administraciones Públicas *stricto sensu*, como si es llevada a cabo por órganos constitucionales o estatutarios en su actividad materialmente administrativa (artículo 1), así como en caso de corporaciones de Derecho Público en ejercicio de funciones públicas y actos de concesionarios dictados en ejercicio de potestades administrativas (artículo 2), debiendo extenderse, como acertadamente indica GAMERO⁴⁰, a las entidades de naturaleza jurídico-privada en caso de ejercicio de potestades administrativas, en consonancia con lo previsto en la LPAC y LRJSP, dada la dicción literal de los artículos 2.1.b) *in fine* en ambos textos legales.

En el seno de dicho control judicial, si el ejercicio de potestades administrativas se ha llevado a cabo mediante fórmulas algorítmicas, ninguna duda cabe de que será imprescindible facilitarlas al juzgador, a efectos de su control, para poder efectuar la correspondiente pericial relativa a su adecuación con el ordenamiento jurídico-administrativo vigente. Piénsese que se puede hacer uso de esta tecnología en supuestos tan variados como la potestad de investigación en materia de bienes públicos (artículo 45 de la Ley 33/2003, de 3 de noviembre,

³⁹ GAMERO CASADO, E. (2021): “Delimitación conceptual de la potestad administrativa”, en GAMERO CASADO, E. (Dir.): *La potestad administrativa. Concepto y alcance práctico de un criterio clave para la aplicación del Derecho Administrativo*, Ed. Tirant Lo Blanch, Valencia, p. 73.

⁴⁰ *Ibíd.*, p. 140.

del Patrimonio de las Administraciones Públicas⁴¹) incluso por las entidades locales, como ha tenido ocasión de examinar GOSÁLBEZ⁴², prevista en el artículo 4.1.b) de la Ley 7/1985, de 2 de abril, reguladora de las Bases del Régimen Local (LRBRL), e incluso en materia de subvenciones, como ha tenido ocasión de examinar profusamente TOSCANO⁴³.

Este autor diferencia dos tesis ante las que nos podríamos encontrar en el ámbito de las subvenciones, que resulta claramente inspirador al tratar los algoritmos: considerar que la comprobación de la subvención supone su consideración como una potestad administrativa, por reunir todas sus notas características, o bien considerar que la comprobación es una facultad auxiliar o de soporte de otras potestades, o incluso una mera actuación material o de ejecución. La clave para el autor reside en que el acto resultado de la comprobación incide en todo caso en la esfera jurídica del interesado, produciendo efectos jurídicos sobre esta por sí solo, unilateralmente, sin ser necesarios requisitos adicionales, sin perjuicio de que en ocasiones precise del apoyo de otras potestades administrativas en determinados supuestos.

Quizás esta doble tesis pueda resultar útil en aras a determinar si el ejercicio de la potestad administrativa de investigación o comprobación emplea algoritmos, especialmente predictivos (por ejemplo, para investigar los bienes de las administraciones públicas que se encuentran sin un destino claro y podrían estar siendo objeto de prescripción adquisitiva por parte de terceros mediante la detección de los bienes que tienen mayor peligro de ser ocupados por terceros, como recién

⁴¹ Con el siguiente tenor literal: “Las Administraciones públicas tienen la facultad de investigar la situación de los bienes y derechos que presumiblemente formen parte de su patrimonio, a fin de determinar la titularidad de los mismos cuando ésta no les conste de modo cierto”.

⁴² GOSÁLBEZ PEQUEÑO, H.: “La recuperación posesoria de los bienes locales: ¿desahucio o interdicto administrativo?”, *La Administración al Día*, Instituto Nacional de Administración Pública, accesible en la dirección web: <https://laadministracionaldia.inap.es/noticia.asp?id=1104536> (última consulta efectuada el 15/06/2022).

⁴³ TOSCANO GIL, F. (2021): “Las potestades administrativas en el ámbito de las subvenciones”, en GAMERO CASADO, E. (Dir.): *La potestad administrativa...*, *op. cit.*, pp. 848 y ss.

temente ha ocurrido⁴⁴), en aras a facilitar la tarea, lo cual implicaría su necesario control judicial.

Pero las potestades de investigación no se limitan a estos ámbitos, sino que se extienden a materias como la tributaria, seguridad social y otras que con seguridad son desconocidas por la ciudadanía en general, dado que se aplican en fases previas o de asistencia a personas humanas, lo cual ha sido puesto de manifiesto por la Red DAIA, en cuya Declaración Final 3^a concluye que es necesario conocer estas aplicaciones algorítmicas también, para que no escapen de la aplicación del Derecho y permita determinar el grado concreto de intervención humana de la que gozan⁴⁵, por lo que se afirma que debe haber una clara trazabilidad y transparencia del uso de sistemas algorítmicos en cualquier fase y tipo de actuación administrativa, ya sea formal o informal.

Este aspecto conecta con lo examinado en el caso SyRI, ampliamente examinado por COTINO⁴⁶, aplicación que la administración holandesa empleaba con fines de investigación, para determinar si alguno de los sujetos que la misma consideraba de riesgo debían ser objeto de una investigación, haciendo uso para ello de algoritmos predictivos y datos ilimitados de los ciudadanos, incluyendo datos relativos a trabajo, sanciones administrativas, datos fiscales, bienes muebles e inmuebles, beneficios sociales, datos comerciales, de vivienda, de identificación, de integración cívica, cumplimiento de la legislación, educación, pensiones, deudas, permisos y exenciones, entre otras.

A pesar de que el sistema contaba con múltiples garantías, entre las que se encontraban las de anonimización, división funcional,

⁴⁴ Algunos diarios se hicieron eco de la noticia: https://www.elconfidencial.com/espana/2022-03-25/ocupas-castillo-corbanera-santander-usucapion_3397793/ (última consulta efectuada el 15/06/2022).

⁴⁵ Conclusión n° 3 de la Declaración Final del “II Seminario Internacional Derecho Administrativo e Inteligencia Artificial en el Sector Público: la importancia de las garantías jurídicas”, celebrado en la Universidad de Valencia los días 10 y 11 de octubre de 2019.

⁴⁶ COTINO HUESO, L. (2020): “SyRI, ¿a quién sanciono? Garantías frente al uso de inteligencia artificial y decisiones automatizadas en el sector público y la sentencia holandesa de febrero de 2020”, en *La Ley Privacidad*, n° 4 (mayo 2020), Ed. Wolters Kluwer.

borrado o de confidencialidad, un Tribunal de La Haya examinó el funcionamiento del sistema en la sentencia de 5 de febrero de 2020⁴⁷, en la que reconoció los esfuerzos por ofrecer ciertas garantías en el uso de la aplicación, pero concluyó que la misma era contraria al artículo 22 RGPD y al artículo 8 del Convenio Europeo de Derechos Humanos (CEDH, en adelante) , en tanto en cuanto encontró carencias como la ausencia de auditorías externas independientes sobre el funcionamiento del algoritmo y la posible incidencia en derechos fundamentales de los ciudadanos, así como los problemas derivados de la opacidad o “caja negra” (*black box*) del sistema, reconociendo el Tribunal que el principio de transparencia es garante de la protección de datos y los derechos fundamentales de los ciudadanos. En ocasiones, el propio sistema discriminaba a determinadas zonas o vecindarios por su nivel bajo de vida, motivo por el cual mucha actividad investigadora se centraba en tales zonas, lo cual resulta a todas luces contrario al CEDH.

Conscientes de la relevancia que se debe otorgar a la transparencia en la aplicación de los algoritmos, en la regulación jurídica que se desarrolle en un futuro deberán tenerse en cuenta tales advertencias, y trabajar, de acuerdo con la opinión de COTINO⁴⁸, en el desarrollo de algoritmos que respondan a “cajas blancas”, esto es, más transparentes, aunque sus resultados puedan ser más reducidos, en aras a dar cumplimiento a los principios y derechos constitucionalmente reconocidos a los ciudadanos y, en definitiva, para proporcionar una buena administración.

4.5.- De la necesidad de prever, en las licitaciones públicas, la cesión en favor de la Administración de los derechos de propiedad intelectual o industrial inherentes a los algoritmos que se desarrollen o se utilicen

El empleo por parte de los poderes públicos de aplicaciones desarrolladas por terceros trae consigo una problemática adicional: la protección de derechos de propiedad intelectual o industrial relacio-

⁴⁷ Rb. Den Haag-C/09/550982/HA ZA 18/388

⁴⁸ COTINO HUESO, L.: “SyRI, ¿a quién...”, *op. cit.*, p. 3.

nadas con el algoritmo o el código fuente empleado para ello por el desarrollador de la aplicación, lo cual trae consigo la imposibilidad de acceder a los mismos ni tan siquiera por la propia entidad pública que hace uso de dicho software, por lo que tampoco podrá ponerla a disposición de los ciudadanos o, en su caso, del juzgador, de ser solicitado en virtud de la normativa vigente en materia de transparencia o requerido para ello durante la sustanciación de un proceso judicial, respectivamente.

Esta situación ocasiona problemas de oscurantismo, sin duda, que ha sido puesto de manifiesto, entre otros, por DE LA CUEVA⁴⁹, para quién esta problemática se suscita en dos ámbitos funcionales: en los de gestión y en los de control, de forma que la modificación de una determinada aplicación podría modificar de forma directa la Ley o el Reglamento a que su funcionamiento obedece, transgrediendo de forma directa el principio de legalidad y alterando el sistema de fuentes, llegando incluso a afirmar la posibilidad de encontrarnos ante casos que supongan una verdadera *dictadura de la máquina*. Si bien me parece excesiva dicha apreciación, no puede obviarse que el control del funcionamiento de las aplicaciones es un verdadero reto para los poderes públicos que hacen uso de ellas, ya que pasaría por formar grupos de expertos interdisciplinares que aunasen a informáticos y juristas (algo cada vez más necesario en la administración electrónica) para verificar periódicamente que el funcionamiento se ajusta a la norma.

Pero la complejidad aparece si no resulta posible, ni tan siquiera formando estos grupos de expertos, acceder al código fuente o al algoritmo que aplica, por lo que en este punto es imprescindible que los poderes públicos, toda vez que deberán licitar la adquisición de programas de software que den respuesta a sus necesidades, prevean en los Pliegos de Cláusulas Administrativas Particulares la cesión por parte del adjudicatario de los derechos de propiedad intelectual o industrial referidos a las aplicaciones que se desarrollen durante la ejecución del contrato, de forma que se pueda tener acceso a los mismos y controlar, de esta forma, su adecuación a la normativa

⁴⁹ DE LA CUEVA, J. (2018): “Código fuente, algoritmos y fuentes del Derecho”, *Revista El Notario del Siglo XXI*, Colegio Notarial de Madrid, n° 77.

que están llamados a aplicar, e incluso de este modo, llegado el caso, pueda ser susceptible de ser adaptado a nuevas necesidades o cambios normativos.

Esta necesidad se verá en la práctica limitada por situaciones en que lo que se adquiere no es el desarrollo de una aplicación, sino la adquisición de licencias para su uso por la entidad, caso en que no será posible obligar a la cesión de derechos sobre el código fuente o el algoritmo, por lo que, salvo supuestos muy excepcionales, siempre y cuando el destino final de la aplicación no sea producir efectos directos en la esfera de derechos de los ciudadanos, se deberá apostar en el futuro por el desarrollo de aplicaciones para uso exclusivo por las administraciones públicas y posterior compartición con el resto a través de las herramientas en materia de transferencia de tecnología reguladas en el artículo 158 LRJSP y que tiene su máximo exponente en el Centro de Transferencia de Tecnología de la Administración General del Estado⁵⁰, o bien haciendo uso de aplicaciones *open source* o de código abierto, que permiten un acceso libre y gratuito con posibilidad de modificación ulterior por parte de las entidades que hagan uso de ellas.

Y es que el Comité de Ministros del Consejo de Europa es sensible a esta problemática y, en este sentido, ya afirmó en su Declaración de 8 de abril de 2020⁵¹ que “los Estados deberían establecer niveles adecuados de transparencia con respecto a los criterios y métodos básicos de contratación pública, utilización, diseño y procesamiento de los sistemas algorítmicos aplicados por y para ellos, o por agentes del sector privado. Los marcos legislativos para la propiedad intelectual o los secretos comerciales no deben impedir esa transparencia, ni los Estados o los particulares deben tratar de explotarlos con este fin. Los niveles de transparencia deben ser lo más altos posible y proporcionales a la gravedad de los impactos adversos en los derechos humanos, incluidas las etiquetas o sellos éticos de los sistemas algorítmicos para permitir a los usuarios navegar entre sistemas. El uso de sistemas algorítmicos en los procesos de toma de decisiones que conllevan altos

⁵⁰ Accesible en la dirección web: <https://administracionelectronica.gob.es/ctt/solucionesCTT.htm#.YqsTkKLP3IU> (última consulta efectuada el 16/06/2022)

⁵¹ Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems.

riesgos para los derechos humanos debe estar sujeto a normas particularmente estrictas en lo que respecta a la explicabilidad de los procesos y los resultados”.[†]

Por tanto, resulta inexcusable proporcionar a los ciudadanos una transparencia efectiva y la trazabilidad debida de los sistemas algorítmicos empleados por los poderes públicos en su toma de decisiones, especialmente las que afectan a los ciudadanos, sin que deban objetarse derechos de propiedad intelectual, secreto comercial o criterios de seguridad por oscuridad para evitarlo.

5.- CONCLUSIONES

En mi opinión, las posiciones doctrinales existentes en cuanto a la naturaleza jurídica de los algoritmos no son estériles, más al contrario, ponen en el objetivo la necesidad de examinar con detenimiento este nuevo fenómeno que se está implementando poco a poco en el sector público para determinar con claridad no sólo el régimen jurídico aplicable, sino cómo debe procederse en aras a su aprobación una vez desarrollado técnicamente, y con evidentes efectos en materias como la transparencia o la información pública a que los ciudadanos tienen derecho a acceder.

Las escasas resoluciones judiciales que han tenido ocasión de pronunciarse al respecto no han sabido hasta la fecha dar respuesta a esta necesidad, toda vez que no han dado luz verde a la posibilidad de acceder al código fuente de los algoritmos cuando ha sido necesario, como se ha tenido ocasión de examinar, tendencia que paulatinamente deberá cambiar para adaptarse a esta nueva realidad, que exige las debidas garantías en la medida de lo posible, ni han ofrecido alternativa alguna a dicho acceso, limitándose a la negativa por defecto.

Y es debe reconocerse que, si bien resulta relativamente sencillo conocer cómo se ha adoptado una determinada decisión cuando se han empleado para ello algoritmos no predictivos, no lo es tanto cuando nos encontramos ante algoritmos predictivos, en los que existe, de un lado, un componente preestablecido y unos datos precargados y, de otro, una serie de variables que van siendo elaboradas de forma inteligente por el sistema gracias a su aprendizaje a lo largo del tiempo. Es por ello que, aunque en el sector privado esto no suponga graves

consecuencias, más allá de la decisión de la entidad, por ejemplo, de otorgar un préstamo o no, o de establecer una prima de seguro superior o inferior, es necesario tener en cuenta que en el sector público la implementación de tales sistemas basados en algoritmos predictivos debe dotarse de las suficientes garantías para que los ciudadanos no vean cercenados sus derechos, ya que este caso, se puede tratar de la asignación de una plaza escolar o el establecimiento del orden de prevalencia en la citación para una intervención quirúrgica en función de su edad, antecedentes o estimación de su esperanza de vida en función de múltiples variables.

La transparencia se erige como vía principal para poder acceder al algoritmo empleado en la toma de decisiones con carácter general, tal y como se ha puesto de manifiesto por el Consejo de Europa⁵², por lo que sería, tal y como establece la Comisión Europea en su Libro Blanco sobre IA, una forma de abordar la opacidad de los algoritmos pero, en el ámbito español, a la vista de las escasas resoluciones administrativas y judiciales existentes sobre la apertura pública del código fuente y los algoritmos empleados en aplicaciones, ha resultado insuficiente.

Por tanto, habría que fomentar un clima de transparencia más exigente en relación con la adopción de decisiones públicas con base en algoritmos, especialmente predictivos, dado que la normativa jurídico-pública actualmente vigente daría una respuesta parcial, al menos en cuanto a los algoritmos no predictivos se refiere, en relación con la actividad administrativa automatizada.

Una posible solución a los problemas de acceso al código fuente sería permitir el acceso parcial al mismo, de forma que no se comprometería la seguridad del sistema, al no permitir el acceso a la información contenida en el mismo relativa a la seguridad, interoperabilidad, interconexión, etc. En definitiva, la solución podría encontrarse en la propia Ley de Transparencia que prevé dicho acceso parcial, permitiendo visualizar y examinar la parte del código fuente en que consta

⁵² Committee of experts on internet intermediaries (MSI-NET), *Algorithms and human rights. Study on the human rights dimensions of automated data processing techniques and possible regulatory implications*, Council of Europe study DGI(2017)12, p. 38.

el algoritmo y su aplicación correspondiente, y sin poner en riesgo la seguridad.

Complementariamente, se podría ofrecer la traducción del algoritmo matemático al lenguaje natural, teniendo en cuenta que los algoritmos pueden expresarse de diversas formas (lenguaje matemático, diagramas de flujo, pseudocódigo o lenguaje natural), por lo que sería posible en la práctica.

BIBLIOGRAFÍA CITADA

- ALAMILLO DOMINGO, I. y URIOS APARISI, X. (2007): “Los límites a la utilización del sello del órgano por parte de las Administraciones Públicas en la Ley 11/2007”, Comunicación presentada a las *Jornadas Tecnimap*, Gijón, 27-30 de noviembre de 2007
- ALAMILLO DOMINGO, I. (2017): “La regulación de la tecnología: la superación del modelo papel como elemento de transformación digital innovadora”, en MARTÍN DELGADO, I. (Dir.): *La reforma de la administración electrónica: una oportunidad para la innovación desde el Derecho*, Ed. INAP, Madrid.
- ARROYO JIMÉNEZ, L. (2020): “Algoritmos y reglamentos”, *Blog Almacén de Derecho*, 25 de febrero de 2020, accesible en la dirección web: <https://almacendederecho.org/algoritmos-y-reglamentos> (última consulta efectuada el 16/06/2022).
- BERMEJO VERA, J. (2009): *Derecho Administrativo. Parte Especial*. Ed. Thomson-Civitas, 7ª ed., Cizur Menor -Navarra-
- BERNING PRIETO, ANTONIO D. (2019): *Validez e invalidez de los actos administrativos en soporte electrónico*, Ed. Thomson-Reuters Aranzadi, Cizur Menor -Navarra-
- BULL, H.P. (2017): *Der vollständig automatisiert erlassene Verwaltungsakt. Zur Begriffsbildung und rechtlichen Einhegung von E-Government*, DvBl 2017 (Ausgabe 7)
- CARRILLO BERNAL, J. (2019): *Paradigma Netflix: entretenimiento de algoritmo*, Ed. UOC, Barcelona
- CERRILLO I MARTÍNEZ, A. (2019): “El impacto de la inteligencia artificial en el Derecho Administrativo: ¿nuevos conceptos para nuevas realidades técnicas?”, *Revista General de Derecho Administrativo*, Ed. Iustel, nº 50
- COTINO HUESO, L. (2020): “SyRI, ¿a quién sanciono? Garantías frente al uso de inteligencia artificial y decisiones automatizadas en el sector público y la sentencia holandesa de febrero de 2020”, en *La Ley Privacidad*, nº 4 (mayo), Ed. Wolters Kluwer

- DE LA CUEVA, J. (2018): “Código fuente, algoritmos y fuentes del Derecho”, *Revista El Notario del Siglo XXI*, Colegio Notarial de Madrid, nº 77
- DE MOMI, E. Y SEGATO, A. (2022): “Autonomous robotic surgery makes light work of anastomosis”, *Science Robotics*, 62
- GAMERO CASADO, E. (2021): “Delimitación conceptual de la potestad administrativa”, en GAMERO CASADO, E. (Dir.): *La potestad administrativa. Concepto y alcance práctico de un criterio clave para la aplicación del Derecho Administrativo*, Ed. Tirant Lo Blanch, Valencia
- GOSÁLBEZ PEQUEÑO, H.: “La recuperación posesoria de los bienes locales: ¿desahucio o interdicto administrativo?”, *La Administración al Día*, Instituto Nacional de Administración Pública
- HUERGO LORA, A. (2020): “Una aproximación a los algoritmos desde el Derecho Administrativo”, en HUERGO LORA, A. (Dir.): *La regulación de los algoritmos*, Ed. Thomson-Reuters Aranzadi, Cizur Menor (Navarra), 1ª ed.
- LÓPEZ PULIDO, J.P. (2021): “El derecho de acceso a los algoritmos y al código fuente, en la actuación administrativa automatizada y su afectación en las entidades locales”, *La Administración Práctica* nº. 11/2021, Ed. Thomson-Reuters Aranzadi (BIB\2021\5197)
- MANZONI, M., MEDAGLIA, R., TANGI, L., VAN NOORDT, C., VACCARI, L. y GATTWINKEL, D. (2022): *AI Watch. Road to the adoption of Artificial Intelligence by the public sector*. Publications Office of the European Union, Luxembourg
- MARTÍN DELGADO, I. (2009): “Naturaleza, concepto y régimen jurídico de la actividad administrativa automatizada”, *Revista de Administración Pública* nº 180
- NAVARRO GONZÁLEZ, R.M. (2017): *La motivación de los actos administrativos*, Ed. Thomson-Reuters Aranzadi, 1ª ed., Cizur Menor -Navarra-
- PASQUALE, F. (2015): *The black box society. The secret algorithms that control money and information*. Ed. Harvard University Press, Cambridge, Massachusetts, London, England
- PONCE SOLÉ, J. (2019): “Inteligencia artificial, derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico”, en *Revista General de Derecho Administrativo* nº 50, Ed. Iustel
- TOSCANO GIL, F. (2021): “Las potestades administrativas en el ámbito de las subvenciones”, en GAMERO CASADO, E. (Dir.): *La potestad administrativa. Concepto y alcance práctico de un criterio clave para la aplicación del Derecho Administrativo*, Ed. Tirant Lo Blanch, Valencia
- VALERO TORRIJOS, J. (2007): *El régimen jurídico de la e-Administración. El uso de los medios informáticos y telemáticos en el procedimiento administrativo común*, Ed. Comares, Granada, 2ª ed.

Capítulo IV

La aplicación del principio de transparencia a la actividad administrativa algorítmica

ISAAC MARTÍN DELGADO¹

Catedrático de Derecho Administrativo

Director del Centro de Estudios Europeos “Luis Ortega Álvarez”

Universidad de Castilla-La Mancha

Resumen: En un contexto de creciente tecnologización de la organización y de los procedimientos de actuación de nuestras Administraciones públicas con clara incidencia en los derechos de los ciudadanos, particularmente a través del empleo de técnicas y herramientas de inteligencia artificial, resulta necesario plantearse si los estándares de control de legalidad que nos hemos dado siguen siendo válidos y reflexionar acerca de cómo regular la implantación de esta tecnología en el sector público. Con este enfoque como premisa y sobre la base del principio de buena administración, en el presente trabajo se analiza la aplicación de las exigencias legales en materia de transparencia a la actividad administrativa algorítmica con el fin de identificar aspectos susceptibles de mejora en nuestro ordenamiento jurídico y se plantean algunas consideraciones que podrían ayudar a reforzar su vigencia.

Abstract: In a context of increasing technologicalisation of the organisation and procedures of our public administrations with a clear impact on citizens' rights, particularly through the use of artificial intelligence, it is convenient to consider whether our legal standards are still valid and to reflect on how to regulate the implementation of this technology in the public sector. With this approach as a premise and taking into account the principle of sound administration, this paper analyses the application of the transparency legal requirements to the algorithmic administrative activity in order to identify aspects that could be improved and adds some considerations that could help to strengthen the transparency of the algorithmic systems.

Palabras-clave: Transparencia, algoritmos, acceso al código fuente, buena administración

Key-words: Transparency, algorithms, access to the source code, sound administration

SUMARIO.- 1. A MODO DE INTRODUCCIÓN.- 1.1. El gran reto para el jurista.- 1.2. Algunas preguntas previas.- 2. CONCEPTO Y CONTEXTO.- 3. PRETEXTO.- 4. LA

¹ ORCID: 0000-0003-3515-1725. Este trabajo se enmarca también en el Proyecto “Administración Pública e Inteligencia Artificial: regulación y uso de la AI en el ámbito de la contratación pública” (TED2021-130682B-I00), financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea NextGenerationEU/PRTR

TRANSPARENCIA DE LA ACTIVIDAD ADMINISTRATIVA ALGORÍTMICA.- 4.1. Premisas.- 4.2. Un ejercicio práctico: regular la transparencia de la actividad administrativa algorítmica.- 4.3. Publicidad activa y derecho de acceso en relación con el uso de IA en el sector público.- 5. A MODO DE CONCLUSIÓN: ALGUNAS GARANTÍAS NECESARIAS.-

1.- A MODO DE INTRODUCCIÓN

1.1.- *El gran reto para el jurista*

La inteligencia artificial (en adelante, IA) es una tecnología transformadora de carácter disruptivo que está impactando –y continuará haciéndolo en el futuro inmediato– a todos los sectores de la sociedad. También a la Administración Pública.

Desde su nacimiento como disciplina científica en los años 50 ha ido evolucionando, rápidamente en unos casos, más lentamente en otros, hasta incrementarse exponencialmente su desarrollo en tiempos recientes como consecuencia de tres factores, principalmente: el crecimiento de la cantidad de datos disponibles (la IA se nutre de ellos); el incremento de la capacidad de computación y de almacenamiento; y el desarrollo de nuevas técnicas. Sin embargo, puede detectarse que su implantación y uso se están volcando más en el ámbito privado –comercio, servicios financieros, turismo, industria, medios de comunicación– que en el ámbito público. Aunque, en puridad, ello no debe sorprender, por la lentitud que suelen demostrar los poderes públicos para incorporar novedades y por la mayor rigidez del Derecho Público frente al Derecho Privado, llama la atención, ciertamente, el hecho de que, incluso en los documentos oficiales sobre IA, la atención se centre más en el mundo de la empresa y en la sociedad que en las propias organizaciones públicas. Éstas también necesitan de transformación digital y, sobre todo, han de ponderar igualmente cómo aplicar en su organización y procesos de actuación técnicas y herramientas de IA que les hagan más eficaces, flexibilicen y simplifiquen sus relaciones con los ciudadanos, mejoren la prestación de los servicios públicos digitales y minimicen los riesgos que su uso conlleva.

La implantación de las TIC –y, más en particular, de la IA– en el ámbito de la Administración Pública debe ir precedida de una reflexión doctrinal a la altura del reto; y, desde luego, apegada a la realidad existente. Sin embargo, ha de reconocerse abiertamente que este campo de estudio, hasta hace muy poco, había quedado relegado a un segundo plano por considerarse algo muy técnico y con escaso impacto en lo que verdaderamente importaba a la disciplina². Curiosamente –lo cual no deja de ser paradójico–, pueden encontrarse reflexiones pioneras realizadas más o menos a la par del surgimiento y desarrollo de la tecnología, tanto en nuestro país como en países de nuestro entorno. V. FROSINI estudiaba en 1984 la automatización administrativa como “la toma en posesión, por parte de la Administración pública misma, de los métodos y de los instrumentos de la tecnología informática actual en vista de una aplicación en la Administración pública”. En ese trabajo pueden encontrarse reflexiones de interés acerca de todos y cada uno de los retos principales que preocupan ahora, 40 años después, a la doctrina: el poder transformador de la información gestionada por programas informáticos, la necesaria reconfiguración del empleo público y la reorganización de tareas y funciones, la traducción del código jurídico a código informático, la transparencia de la actuación administrativa, la responsabilidad por mal funcionamiento de las máquinas, el control de la actividad administrativa informatizada y los riesgos que de todo ello pueden derivarse para el principio de legalidad, los derechos e intereses de los ciudadanos y la independencia de la Administración respecto de la empresa, sobre la base de la premisa de que quien es poderoso se hace cada vez poderoso, aun en el error³. Lo mismo puede señalarse

² Muy significativa, en este sentido, es la afirmación de S. MUÑOZ MACHADO en el número especial de *El Cronista del Estado Social y Democrático de Derecho* dedicado a Inteligencia Artificial, cuando señala que “en el número 0 fijé un catálogo provisional de los principales problemas que afectaban a nuestras instituciones (...) quizás la más notable de las ausencias de aquella lista primera de preocupaciones fuera la revolución digital y sus consecuencias y retos para el Derecho”. MUÑOZ MACHADO, S. (2022): “Prólogo”, en *El Cronista del Estado Social y Democrático de Derecho*, n° 100, pp. 5.

³ FROSINI, V. (1984): “Informática y Administración Pública”, *Revista de Administración Pública*, n° 105, p. 447 y ss. También en esa época pueden encontrarse enfoques preventivos ante el uso de computadores. Vid. FERNÁNDEZ, T.-R. (1971): “Jurisprudencia y computadores”, *Revista de Administración Pública*, n°

respecto de los estudios de A. MASSUCCI y G. DUNI en Italia, por señalar otro ejemplo⁴.

El gran reto en el momento presente radica en encontrar los puntos de conexión entre la inteligencia artificial y la Administración Pública y en formular la doble pregunta acerca de, por un lado, qué novedad puede aportar el uso de este conjunto de nuevas tecnologías y, por otro, si las normas y garantías que hemos desarrollado en el ámbito del Derecho Administrativo entendido como ciencia jurídica son todavía útiles o, por el contrario, han de buscarse otras. Ello exige partir de la realidad práctica actualmente existente y del estado actual de la técnica, planteando preguntas no futuristas⁵, sino realistas, con el fin de encontrar respuestas útiles en el contexto actual.

64, pp. 327-336. Ilustra el planteamiento la siguiente afirmación “hay que decir que la lectura de estas frases (referidas a un proyecto para el tratamiento de la jurisprudencia por medio de ordenadores) produce verdaderos escalofríos. Ni más ni menos que el problema de las vigencias, y con él, el de la interpretación de las normas, va a venir resuelto por el ordenador, que es quien va a decir cuál es la norma aplicable, cuál es la legislación actual y cuáles los preceptos derogados” (p. 331); termina el autor con una llamada a la cautela: “sin negar radicalmente la validez futura de los ordenadores en este campo –el tiempo y las experiencias ajenas podrán decirlo en su día y darán cuenta, al mismo tiempo, de sus límites concretos–, parece obligado apuntar desde ahora sus riesgos y sus peligros, frente a tanto entusiasmo apriorístico y tanto y tan ingenuo optimismo-maquinista, especialmente en un país como el nuestro, tan formalista en su talante jurídico” (p. 335).

⁴ DUNI, G. (1978): “L'utilizzabilità delle tecniche elettroniche nell'emanazione degli atti e dei procedimenti amministrativi. Spunto per una teoria dell'atto emanato in forma elettronica, *Rivista amministrativa della Repubblica Italiana*, n° 129, pp. 407 y ss.; MASSUCCI, A. (1993): *L'atto amministrativo informatico. Primi lineamenti di una ricostruzione*, Jovene editore, Napoli.

⁵ Comienzan a publicarse artículos en esta clave –que hemos denominado futurista– en los que sus autores se plantean cuestiones tales como si existirán derechos fundamentales en un mundo de IA cuando la persona pase a un segundo plano o si los algoritmos han de tener reconocidos derechos. En este sentido pueden verse, respectivamente, URBINA MENDOZA, E.J. (2020): “El Derecho público del algoritmo. Reflexiones sobre la transición de la modernidad jurídica crítico/lineal a la cuántica/fractal”, en *Revista de Derecho Público*, n° 161/162, p.11 y ss.; y OSÉS, G. (2020): “Algoritmos con derechos”, *Diario 16*, 8 de diciembre de 2020, disponible en <https://diario16.com/algoritmos-con-derechos/> (Última visita: 30/12/2022), quien considera que los algoritmos son los esclavos del siglo XXI y les deben ser reconocidos derechos. Presentan el valor de ayudar a reflexionar sobre el futuro que puede derivarse del uso de la IA en el ámbito

No en vano, “frente a estas realidades, no ancladas en un futurismo imaginativo, el jurista no puede permanecer ausente o limitarse a ignorarlas”⁶. Efectivamente, hemos de ser capaces de ofrecer una visión realista y desmitificada de la IA partiendo de sus capacidades actuales, puesto que las especulaciones sobre el futuro –que, aparentemente resultan más interesantes– no pueden hacerse a costa de distraerse de lo verdaderamente importante⁷. También equilibrada, porque los algoritmos no son ni buenos ni malos por sí mismos, sino que sus efectos –potenciales y concretos– dependen de su uso⁸. Y es que no podemos obviar el hecho de que no pocos proyectos de implantación de IA derivan en prototipos y simulaciones que no llegan a ser aplicables directamente por motivos diferentes –alto coste, implicaciones éticas o legales⁹...– ni tampoco la máxima de que, como regla de principio, la Administración digital puede funcionar igual que la Administración tradicional¹⁰. Junto con ello, el desarrollo actual de esta tecnología dista mucho del estado necesario para poder hablar de inteligencia en sentido estricto; los sistemas que se están creando producen resultados útiles, pero no inteligentes. Ciertamente, esa utilidad es muy relevante, sobre todo porque, en no pocos casos, permiten alcanzar resultados que escapan a la capacidad de la mente humana.

público, pero resultan poco útiles desde la perspectiva regulatoria. No será el enfoque que se adopte en estas páginas.

⁶ La expresión es de VILLAR PALASÍ y fue escrita en 1978, seguida de las siguientes palabras: “en este momento no se trata de problema de futurismo, sino de problemática ya actualmente en utilización y con una perspectiva de ampliación enorme”, VILLAR PALASÍ, J.L. (1978): “Aspectos jurídicos y políticos de la Telemática”, *Revista Española de Derecho Administrativo*, nº 19, p. 501. En este artículo, además, señala ya alguno de los riesgos que pueden derivarse de las nuevas tecnologías, centrándolos en la pérdida de libertades esenciales (en p. 502).

⁷ SURDEN, H. (2019) “Artificial Intelligence and Law: An Overview”, *Georgia State University Law Review*, nº 35, pp. 1306-1307.

⁸ FRY, H. (2019): *Hola mundo. Cómo seguir siendo humanos en la era de los algoritmos*, Blackiebooks, Barcelona, p. 4.

⁹ FERNÁNDEZ GIL, A. (2021): Introducción a la obra *Inteligencia Artificial (Umbrales éticos, Derecho y Administraciones Públicas)*, de Mar MORENO REBATO, Aranzadi, Cizur Menor, p. 13.

¹⁰ HUERGO LORA, A. (2020): “Una aproximación a los algoritmos desde el Derecho Administrativo”, en HUERGO LORA, A. (Dir.): *La regulación de los algoritmos*, Thomson-Aranzadi, Madrid, p. 26.

1.2.- Algunas preguntas previas

Innovar la Administración en clave tecnológica no es cuestión de tecnología ni de leyes únicamente; exige un planteamiento global y un enfoque integral de carácter multidisciplinar. Y ello ha de hacerse, en mi opinión, sin ideas preconcebidas ni prejuicios injustificados, sobre la base de una pregunta inicial: si se permite en el sector privado, *¿por qué* no lo incorporamos en el sector público? A la hora de plantear esta pregunta no se olvida que existen diferencias importantes entre uno y otro ámbito de implantación de la IA, sino que se busca más bien evidenciar el hecho de que desde el ámbito administrativo no se está llevando a cabo un eficaz aprovechamiento de la ingente cantidad de datos que generan, adquieren y conservan nuestras Administraciones Públicas para el ejercicio de las funciones a éstas encomendadas y en aras de la mejor satisfacción de los derechos e intereses de los ciudadanos y de la persecución del interés general. La diferencia que marca la vinculación positiva de la Administración Pública al principio de legalidad, particularmente en el ejercicio de las potestades ablatorias, no puede entenderse como una excusa para no impulsar el uso de la IA en aquellos ámbitos de acción en los que puede aportar valor añadido y, con ello, prestar un mejor servicio a los ciudadanos.

Junto con ello, resulta preciso formular una segunda pregunta de partida: *¿para qué* incorporar las tecnologías disruptivas y su capacidad transformadora en los diferentes niveles de Administración Pública? No me refiero tanto a la finalidad general como al objetivo en lo concreto. Una mirada rápida a la realidad permite demostrar que se está comenzando a hacer uso de estas herramientas en el ámbito de potestades ablatorias, y no tanto en beneficio de los ciudadanos. Baste con pensar en la posibilidad de automatizar sanciones, introducida por el Real Decreto Ley 2/2021, de 26 de enero, de refuerzo y consolidación de medidas sociales en defensa del empleo, que incorpora en la Ley de Infracciones y Sanciones en el Orden Social la posibilidad de extender actas de inspección de forma automatizada¹¹. Esta previsión

¹¹ Para un completo análisis de esta cuestión, vid. GOERLICH PESET, J. M. (2021): “Decisiones administrativas automatizadas en materia social: algoritmos en la gestión de la Seguridad Social y en el procedimiento sancionador”, *Labor*, vol. 2, nº 2, pp. 22-42. También TODOLÍ SIGNES, A. (2020): “Retos legales del uso

ha sido desarrollada por el Real Decreto 688/2021, de 3 de agosto, por el que se modifica el Reglamento general sobre procedimientos para la imposición de sanciones por infracciones de orden social y para los expedientes liquidatorios de cuotas de la Seguridad Social, de modo tal que es posible en nuestro sistema iniciar un procedimiento sancionador y tramitarlo íntegramente de forma automatizada, imponiendo una sanción ante una eventual infracción, siempre que no haya oposición por parte del interesado.

Ciertamente, la Administración tiene la función de inspeccionar, perseguir los incumplimientos normativos y sancionarlos; pero es mucho más que eso. Si compartimos la idea de que la Administración es una organización social al servicio del interés general, habremos de exigir igualmente que la innovación se encauce para la mejor satisfacción de los derechos e intereses de los ciudadanos, a los que sirve, y para la más eficaz prestación de servicios públicos, que justifica su existencia. Los principios de servicio efectivo a los ciudadanos y de simplicidad, claridad y proximidad a estos, reconocidos en el art. 3.1 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público (en adelante, LRJSP), no pueden referirse única y exclusivamente al pago de multas, la exacción de impuestos o la inspección en el ámbito de la Seguridad Social. La Administración se legitima ante los ciudadanos, principalmente, haciéndose útil a estos. Y, si es inútil, ¿cómo será vista en un contexto de superación de intermediarios? De la misma manera en que el *homo digitalis* implica un cambio sustancial en la forma en la que el ser humano se relaciona con el mundo, existe el riesgo real de que el *civis digitalis* acabe viendo a la Administración como una rémora del pasado, innecesaria en el mundo de la revolución digital. En definitiva, está en juego el papel central de la Administración como institución social que defiende el interés general y el bien común¹².

Finalmente, al por qué y al para qué hemos de añadir una última cuestión previa conectada con las dos anteriores: ¿cómo hacer uso

del *big data* en la selección de sujetos a investigar por la Inspección de Trabajo y de la Seguridad Social”, *Revista Galega de Administración Pública*, n° 59, pp. 313-337.

¹² RAMÍO, C. (2022): *Burocracia inteligente. Guía para transformar la Administración pública*, Catarata, Madrid, pp. 13 y ss.

de la IA en el marco de las relaciones entre Administración y ciudadanos? Esta tecnología trae consigo la promesa de la objetividad absoluta y la eficacia perfecta. Ello tiene un efecto positivo que permite mejorar, sin duda alguna, el ejercicio de la función administrativa: optimiza el modelo burocrático sin el elemento subjetivo que puede condicionarlo¹³. Pero, al mismo tiempo, no debe desconocerse que su implantación puede llegar a hacerse a costa de la libertad de los ciudadanos. El determinismo algorítmico¹⁴, el cumplimiento absoluto de la legalidad vigente, la administración del interés general de una forma rápida y eficaz, sin fallos ni fisuras, sin indefiniciones, la plena seguridad jurídica, ¿en qué medida es compatible con la libertad, entendida como libre determinación de las personas, consagrada en el art 10 de la Constitución?¹⁵.

Desde una perspectiva diferente –que podría denominarse interna–, puede apreciarse igualmente una clara pérdida de relevancia del papel del Derecho como cauce y guía de la implantación y uso de la IA por nuestras Administraciones. No sólo porque en los más importantes documentos de referencia, tanto a nivel europeo como a nivel nacional, se apueste más por las estrategias y por la ética que por la regulación, sino porque el Derecho es visto como un obstáculo al potencial transformador de la tecnología. El algoritmo puede convertirse en la práctica en el sustituto de la Ley¹⁶, por acción o por omisión. Por acción, como consecuencia de una voluntad consciente de prescindir del rol del Derecho a la hora de enfocar la aplicación de IA en el sector público, establecer los requisitos relativos al diseño de los sistemas algorítmicos y fijar las garantías de los derechos de los ciudadanos, en beneficio de la ética y la autorregulación; por omisión,

¹³ *Íbidem*, p. 61

¹⁴ LASALLE, J.M. (2019): *Ciberleviatán. El colapso de la democracia liberal frente a la revolución digital*, Arpa, Barcelona, p. 78.

¹⁵ En relación con esta cuestión, PRESNO considera que el recurso a sistemas de IA por parte de los poderes públicos debe hacerse sin que quede afectado ese mínimo invulnerable que garantiza el principio de la dignidad humana y, en positivo, promoviendo el libre desarrollo de la personalidad. PRESNO LINERA, M. (2022): *Derechos fundamentales e inteligencia artificial*, Marcial Pons-Fundación Jiménez Abad, Madrid, p. 104.

¹⁶ LASALLE afirma, con acierto, que “el algoritmo no puede ser una ley por la ausencia de la ley del Estado”, *Íbidem*, p. 158.

por la inactividad del Legislador y del Ejecutivo a la hora de ejercer su poder normativo y directivo. Junto con ello, en los casos en los que se está comenzando a regular el uso de la IA, los controles establecidos son más nítidos cuando se trata de entidades privadas que cuando se refiere a entidades públicas¹⁷.

En definitiva, si es cada vez mayor el uso de diferentes tecnologías para la ejecución de la Ley por parte de las Administraciones Públicas, hemos de entender que la intermediación de la programación informática constituye un elemento que ha de ser objeto específico de estudio –e, incluso, de regulación–, por lo que resulta imperativo afrontar el análisis en profundidad de los retos que de ello se derivan. Ciertamente, el código informático puede interferir en el código jurídico¹⁸, no podemos no tenerlo en cuenta.

Estamos en periodo de transición digital, con escasas normas específicamente pensadas para dar respuesta a los desafíos que plantean las tecnologías disruptivas, y ello provoca confusión y crea inseguridad jurídica acerca de cómo resolver los problemas que se plantean. Precisamente por esta razón resulta preciso abordar el reto desde la Academia.

Realizadas estas premisas, el análisis que se ofrecerá en las siguientes páginas se centra en lo relativo a la actividad formalizada de la Administración en el contexto del procedimiento administrativo. Aunque no se desconoce que la IA puede ser igualmente útil –y plantea también riesgos– en otros ámbitos de actuación administrativa, sin duda alguna el relativo al procedimiento es el que mejor permite testar las hipótesis que se plantean en ellas. No en vano, gran parte del Derecho Administrativo consiste en procedimiento y muchos de sus principios nucleares –buena fe, imparcialidad, equidad, racionalidad,

¹⁷ BOIX PALOP, A. (2022): “Transparencia en la utilización de inteligencia artificial por parte de la Administración”, en *El Cronista del Estado Social y Democrático de Derecho*, n° 100, p. 93.

¹⁸ Con este enfoque, véase WU, T. (2019): “Will artificial intelligence eat the law? The rise of hybrid social-ordering systems”, *Columbia Law Review*, Vol. 119, n° 7, pp. 2001-2028. Para un análisis concreto en un ámbito específico de actuación desde esta perspectiva, MICHELER, E. y WHALEY, A. (2018): “Regulatory technology: replacing Law with computer code”, *Law, Society, Economy Working Papers*, n° 14, pp. 1-28.

accesibilidad, transparencia, rendición de cuentas, participación...– poseen carácter procedimental¹⁹; junto con ello, el principal efecto transformador de las TIC afecta a los procedimientos de decisión. Todo ello conduce a la necesidad de repensar las categorías jurídicas y los parámetros normativos que nos hemos dado; a valorar adecuadamente qué aspectos de los mismos han de ser reforzados para mantener los logros conseguidos hasta el momento desde la perspectiva de la legalidad, de la legitimidad de la acción administrativa frente a los ciudadanos y de la satisfacción del interés general; y a determinar, en su caso, la conveniencia de crear nuevos principios y garantías.

Más en concreto, y a modo de aplicación práctica de esta metodología de aproximación a las relaciones entre Derecho Administrativo e Inteligencia Artificial, la reflexión se focalizará en una de las exigencias –ciertamente, fundamental– relativa a cómo implantar la IA en la actividad administrativa: la referente a la transparencia. Antes de ello, sin embargo, se llevará a cabo una doble reflexión acerca del concepto de IA aplicada a la Administración y del contexto de referencia, pues de uno y otro se derivan exigencias relevantes para la cuestión principal que será objeto de análisis.

2.- CONCEPTO Y CONTEXTO

Aprehender jurídicamente una realidad exige partir de una definición que haga posible identificar sus manifestaciones para poder establecer un marco jurídico adecuado. Tal definición, sin embargo, ha de ser lo suficientemente flexible –más aún cuando tiene que ver con realidades, como la IA, que evolucionan muy rápidamente– para perdurar en el tiempo²⁰.

De una manera sencilla, aunque existen definiciones doctrinales muy diferentes acerca de qué ha de entenderse por IA, puede partirse de la siguiente idea: IA es todo agente racional creado por humanos

¹⁹ HARLOW, C. y RAWLINGS, R. (2019): “Proceduralism and automation: challenges to the values of Administrative Law”, *Law, Society, Economy Working Papers*, nº 3, p. 2.

²⁰ En este sentido, MORENO REBATO, M. (2021): *Inteligencia Artificial (Umbrales éticos, Derecho y Administraciones Públicas)*, Aranzadi, Cizur Menor, p. 13.

que decide y actúa sobre la base de la percepción²¹, procesando información para producir un resultado a través de un razonamiento que emula el realizado por humanos. Es esa capacidad de percibir lo que le rodea y actuar en consecuencia e, incluso, de transformarlo, lo que le hace merecer el adjetivo de “inteligente”.

Con carácter general, la IA es una disciplina científica que incorpora varias técnicas complejas –aprendizaje automático, razonamiento automático e, incluso, la robótica– que permiten el diseño y puesta en práctica de sistemas de software y hardware que deciden o ayudan a decidir sobre la base del manejo e interpretación de datos. También que, en el momento actual, difícilmente puede hablarse de inteligencia tal y como la predicamos respecto de los seres humanos, esto es, con capacidades cognitivas similares a las de una persona.

Efectivamente, para hablar con propiedad de algoritmos y acción administrativa es necesario partir de los conceptos y de las consecuencias derivadas del uso de aquéllos en el ámbito de la Administración.

Como es sabido, el término algoritmo se define por la RAE como un conjunto ordenado y limitado de operaciones que permite encontrar la solución de un problema. Aplicado al ámbito de la IA, puede señalarse que, si bien cumple la misma función, lo hace sobre la base de instrucciones lógicas traducidas a código informático²² –creadas íntegramente por humanos y que operan de forma directa e inequívoca o parcialmente generadas por el propio sistema–, si bien sin comprender la información manejada como lo haríamos los humanos. En realidad, en función del tipo de tarea que se encomiende, la comprensión por la máquina del conocimiento que está generando es irrelevante; lo importante es que su aproximación estadística derivada del manejo computacional de una ingente cantidad de datos sea acertada para el objetivo buscado. Desde esta perspectiva, cuanto más susceptible de codificación sea el ámbito en el que se usan los sistemas algorítmicos, más eficientes y útiles serán estos. Y, precisamente por esa razón, para comprender verdaderamente la lógica de funcionamiento

²¹ RUSSELL, S. y NORVING, P. (2010): *Artificial Intelligence, a modern approach*, 3ª edición, Prentice Hall, New Jersey pp. 1-2.

²² LESSIG, L. (1999): *Code and other laws of cyberspace*, Basic Books, New York.

del sistema, será necesario dar a conocer cuáles son los objetivos reales a los que sirve²³.

Podemos definir genéricamente la acción administrativa algorítmica como la actuación de la Administración por medio de sistemas que integran procesos algorítmicos con el fin de automatizar el proceso decisorio humano²⁴, total o parcialmente²⁵. Esta definición incluiría tanto la automatización completamente programada, como el uso de la IA en sentido estricto. No obstante, se debe distinguir la una de la otra. En el primer caso, el producto de la máquina representa la voluntad humana, anticipada a través de la programación –constituyendo así una forma diferente de expresión de la voluntad, ya que la programación responde plenamente a las instrucciones del sujeto que los programa– que sintetiza los criterios normativos, transformados en algoritmos, y conduce a la decisión final a través de una secuencia lógica y condicionada predeterminada. En el segundo caso, no se trata simplemente de programación, sino de “pensamiento”, entendido como capacidad de formar un juicio u opinión sobre algo, de llevar a cabo procesos intelectuales propios²⁶. De hecho, como es sabido, el *machine learning* permite generar predicciones a través de su propio sistema de aprendizaje. Y el aprendi-

²³ CARDON, D. (2018): *Con qué sueñan los algoritmos*, Ediciones Dado, Madrid, p. 81.

²⁴ COBBE, J. (2019): “Administrative Law and the Machines of Government: Judicial Review of Automated Public-Sector Decision-Making”, *Legal Studies*, vol. 39, y disponible en el link <http://dx.doi.org/10.2139/ssrn.3226913>, p. 3 (Última visita: 30/12/2022).

²⁵ HUERGO LORA distingue entre actuación administrativa automatizada y predicciones algorítmicas. Para él, hay actuaciones automatizadas sin inteligencia artificial e inteligencia artificial sin actuación automatizada y, efectivamente, es así. Vid. HUERGO LORA, A. (2021): “Administraciones Públicas e inteligencia artificial: ¿más o menos discrecionalidad?”, *El Cronista del Estado Social y Democrático de Derecho*, n° 96-97, pp. 78-95. Sin embargo, en este trabajo el término automatización se usa en el sentido de reproducción de procesos intelectivos u operativos por parte de máquinas a través de sistemas de información, con independencia de que se empleen para adoptar actos administrativos –definitivos o de trámite– o para extraer información de interés basándose en datos.

²⁶ A lo largo de estas páginas se usarán indistintamente las expresiones sistema algorítmico, sistema de IA o, incluso, algoritmo, para hacer referencia al software diseñado por seres humanos con el fin de resolver problemas sobre la base de la interpretación de datos.

zaje es señal y requisito de la inteligencia²⁷. Estos algoritmos no funcionan sobre la base de la deducción, sino que hacen predicciones en términos de probabilidad. De este modo, tienen mayor capacidad para representar el mundo real.

No obstante, es preciso señalar que hablamos de programación y, por tanto, de optimización de sistemas para la consecución de objetivos concretos sobre la base de datos y procesos de clasificación. Dicho sencillamente, su enorme potencial en cuanto a la generación de conocimiento se compensa con una patente “estrechez de percepción”²⁸. Eso sí, con un impacto tan grande que pueden terminar afectando a la colectividad mucho más de lo que lo hacen decisiones adoptadas por humanos, puesto que estos sistemas que fabricamos, a su vez nos construyen, al poseer la capacidad de organizar y orientar nuestra realidad; dicho sencillamente: lo real es cada vez más manipulable²⁹. Del mismo modo, existe un segundo efecto pernicioso que puede derivarse de esa gran capacidad computacional: el incremento exponencial de los datos desborda la capacidad de asimilación del ser humano, lo que hace imprescindible la máquina, generando una importante dependencia de ella³⁰.

En definitiva, estamos viviendo un proceso de transformación que está derivando en que las decisiones adoptadas por el ser humano sobre la base de la información obtenida por él mismo sean cada vez más reducidas en beneficio de la adopción de decisiones que parten de información ofrecida por las máquinas. Ello afecta igualmente a la Administración.

En este contexto, ha de insistirse en que el concepto de IA se refiere a la tecnología que hace a las máquinas “inteligentes”, imitando algunas de las facultades intelectuales humanas. Esto presenta múl-

²⁷ TURING, A.M. (1950): “Computing Machinery and Intelligence”, *Mind* N.S., 59, p. 433 ss. Como indica BARRIO, lo artificial se refiere al proceso de llegar a la inteligencia, no al resultado final. BARRIO ANDRÉS, M. (2022), “Inteligencia artificial: origen, concepto, mito y realidad”, en *El Cronista del Estado Social y Democrático de Derecho*, n° 100, p. 18.

²⁸ CARDON, D. (2018): *Con qué sueñan los algoritmos...*, op. cit., p. 13; 21 y 58.

²⁹ TURING, A. M. (1950): “Computing Machinery and Intelligence”..., op. cit., p. 433 ss.

³⁰ LASALLE, J.M. (2019); *Ciberleviatán...*, op. cit., pp. 40-43.

tiples aplicaciones: robótica, automatización de procesos, capacidad decisoria ... Un conjunto de algoritmos es un código, una secuencia de instrucciones para resolver problemas que transforman datos en conocimiento con el fin de adoptar decisiones. La principal funcionalidad del algoritmo es que “permite hablar a los datos”³¹, porque busca datos e identifica patrones de actuación y correlaciones entre esos datos y la finalidad que se quiere conseguir. En resumen, las máquinas aprenden a generar descripciones, predicciones y prescripciones –y, por tanto, conocimiento– de acuerdo con los datos. No lo decide un programa basado en determinados algoritmos, sino que la decisión se explica sobre la base de los datos o, dicho de otra forma, hace que los datos tengan sentido³². Está claro que no todos los problemas jurídicos pueden ser resueltos a través de un algoritmo, porque muchos necesitan de la intuición y no únicamente de la capacidad analítica³³. Pero el Derecho Administrativo ofrece múltiples espacios tanto para las actividades materiales (guiar el tráfico mediante semáforos operados por máquinas, vigilar los espacios públicos con drones, demarcar la línea de una autopista y encontrar los terrenos más idóneos para la expropiación), como para las actividades formales (determinar qué empresas deben ser inspeccionadas por la Agencia Estatal de la Administración Tributaria, seleccionar empleados públicos para una Comunidad Autónoma, conceder subvenciones por entidades locales, verificar riesgos de incumplimiento normativo) que pueden ser optimizadas mediante el uso de los mismos.

En este sentido, hay que tener en cuenta que existen diferentes tipos de actividad administrativa y ámbitos de aplicación muy diversos. En lo que ahora interesa, debe centrarse la atención en tres de ellas:

³¹ YEUNG, K. K. (2018): “Algorithmic Regulation: a Critical Interrogation”, en *Regulation & Governance*, 12, p. 506

³² Como precisa HUERGO LORA, ante la aplicación de algoritmos predictivos desaparecen las decisiones subjetivas, que pasan a ser sustituidas por predicciones basadas en correlaciones descubiertas con el análisis de una gran cantidad de datos referidos a operaciones anteriores; pero, al mismo tiempo, también implica el desplazamiento de criterios racionales, porque esas predicciones sustituyen la causalidad por la correlación, HUERGO LORA, A. (2020): “Una aproximación a los algoritmos desde el Derecho Administrativo...”, *op. cit.*, p. 35.

³³ FROSINI, V. (2019): *Cibernética, Derecho, Internet y Sociedad*, Ediciones Olejnik, Santiago de Chile, p. 88.

la regulación; la adopción de actos administrativos (discrecionales y no discrecionales); y la actividad material en el ejercicio de potestades administrativas (como pueden ser la inspección, la organización del trabajo de los empleados públicos o la publicación de información para el cumplimiento de las obligaciones en materia de transparencia). Pues bien, el espacio y el papel de los algoritmos serán diferentes según el tipo de actividad. Es decir, las reflexiones sobre las relaciones entre IA y Administración Pública no se pueden hacer de forma general, sino que es imprescindible partir de la diferenciación entre áreas de acción, puesto que son muchos los ámbitos susceptibles de automatización: no es lo mismo hacerlo para el desarrollo de actividad formalizada (procedimientos de licitación) que para la prestación de servicios públicos (diagnóstico de enfermedades). Si entendemos que el uso de la IA en la Administración Pública es una opción, y no una obligación, le corresponde a la doctrina, al Legislador, al Gobierno, a los operadores jurídicos y técnicos –y, en última instancia, a la Justicia– identificar su espacio. Los algoritmos pueden servir para decidir o, simplemente, para apoyar las decisiones humanas. En el marco de las decisiones, pueden ser utilizados para ejercer potestades regladas o de discrecionalidad reducida³⁴. Pero no debemos olvidar que también será posible utilizar los algoritmos para crear, aplicar y reformar normas³⁵. El Derecho es parcialmente computable, y la aplicación de IA se sintetiza en la descripción de procesos de aprendizaje y otras ca-

³⁴ MARTÍN DELGADO, I. (2009): “Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada”, en *Revista de Administración Pública*, nº 189, p. 353 y ss.

³⁵ SOLUM, L.B. (2019): “Artificially Intelligent Law”, en *Rivista de BioDiritto*, 1, p. 53 ss.; CANALS, D. (2019): “El proceso normativo ante el avance tecnológico y de la transformación digital (inteligencia artificial, redes sociales y datos masivos)”, en *Revista General de Derecho Administrativo*, 50. En este sentido, MORENO REBATO precisa que la capacidad de los lenguajes informáticos ya existentes de traducir normas jurídicas en código informático para adoptar decisiones no debe condicionar el ejercicio de la potestad legislativa y la potestad reglamentaria; sin embargo, al mismo tiempo considera que en un futuro pudiera ocurrir que las normas jurídicas se elaborasen en dos formatos, lenguaje natural y lenguaje informático, haciendo posible su ejecución completa; MORENO REBATO, M. (2021): *Inteligencia Artificial (Umbralés éticos, Derecho y Administraciones Públicas)*..., op. cit., p. 129.

racterísticas de la inteligencia humana con exactitud para que puedan ser emulados por una máquina³⁶.

Esto nos sitúa ante la necesidad de llevar a cabo una reflexión profunda sobre los tipos de poderes y decisiones administrativas que pueden ser ejercidos y adoptadas mediante algoritmos. Y también acerca de cómo garantizar que el paso del lenguaje jurídico al lenguaje informático se mantenga fiel a la voluntad normativa del Legislador. Este último extremo es fundamental en tanto que aquellos a quienes corresponde escribir el código informático para trasladar al mismo el lenguaje natural –en este caso, el lenguaje de la norma– llevan a cabo una labor de intermediación que requiere interpretación y, por tanto, es susceptible de error³⁷. No en vano, su labor no sólo implica el diseño de un programa informático, sino que, desde una perspectiva jurídica, están construyendo efectivamente sistemas de adopción de decisiones³⁸.

La plasmación práctica de todos estos retos tiene un doble nexo común: de un lado, han de cumplir con el principio de buena administración; de otro, quedan condicionados por el principio de transparencia.

3.- PRETEXTO

Expuestos sintéticamente el concepto de IA y sus implicaciones para el Derecho Administrativo, así como su contexto aplicativo, conviene detenerse, siquiera brevemente, en el pretexto, es decir, en la fundamentación de su uso.

³⁶ DE LA SIERRA, S. (2022): “La matematización de la realidad y del Derecho Público”, en *Ibericonect*, 14 de marzo de 2022, disponible en <https://www.ibericonnect.blog/2022/03/la-matematizacion-de-la-realidad-y-del-derecho-publico/> (Última visita: 30/12/2022).

³⁷ CITRON, D.K. (2008): “Open Code Governance”, *University of Chicago Legal Forum*, nº 1, p. 366-367.

³⁸ HOGAN-DORAN, D. (2017): “Computer says “no”: automation, algorithms and artificial intelligence in Government decision-making”, *The Judicial Review*, nº 13, p. 8.

El término pretexto se emplea aquí en sentido positivo, no peyorativo. Resulta evidente que hacer uso de IA en el sector público no es una imposición, sino una opción. Pero una opción que se torna inexcusable si tomamos en serio un principio general de actuación de la Administración que ha sido reconocido, además, como un conjunto de derechos de los ciudadanos: el principio de buena administración.

Efectivamente, la buena administración debe ser el principio que guíe el empleo de IA en la organización y el procedimiento administrativo algo que, quizás, no ha sido interiorizado aún por nuestras Administraciones Públicas³⁹; al menos, no lo suficiente.

En líneas precedentes se ha insistido con carácter general en la necesidad de relacionar la implantación de IA con el mejor cumplimiento de la misión de satisfacer el interés general y contribuir al bien común que la Constitución encomienda a la Administración Pública. La forma de concretar esta idea desde la perspectiva jurídico-administrativa es conectar el uso de algoritmos en el sector público con la función de servicio a los ciudadanos que justifica la existencia de la Administración Pública como organización social.

Una buena Administración Pública es una Administración que cumple con las funciones legalmente encomendadas, pero que lo hace con transparencia, al servicio de los ciudadanos y de forma imparcial, racional y motivada; y las TICs en general –y la IA en particular– pueden ayudar eficazmente a la realización del principio/derecho a una buena administración⁴⁰.

Como es sabido, desde una perspectiva dogmática, el procedimiento administrativo es considerado en el ordenamiento español como una institución jurídica autónoma de relevancia constitucional llamada a cumplir una triple función: instrumento para la satisfacción del interés general, cauce de legalidad de la actuación administrativa

³⁹ PONCE SOLÉ, J. (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad; algoritmos y procedimiento administrativo debido tecnológico”, *Revista General de Derecho Administrativo*, n° 50, p. 6.

⁴⁰ GALETTA, D.U. (2020): “Digitalizzazione e diritto ad una buona amministrazione (il procedimento amministrativo, fra Diritto UE e tecnologie ICT)”, en CAVALLO PERIN, R.; GALETTA, D.U. (ed), *Il Diritto dell'Amministrazione Pubblica digitale*, Giappichelli, Turín, pp. 85-86.

y medio de participación de los ciudadanos en la toma de decisiones administrativas.

Efectivamente, esta triple función tiene su proyección en la propia Constitución Española, de la que se derivan principios y derechos aplicables al procedimiento administrativo; de este modo, el procedimiento administrativo se configura como instrumento de cumplimiento de los principios constitucionales relacionados con la actuación administrativa.

La Constitución proclama en su art. 9.3 el principio de interdicción de la arbitrariedad y en el art. 31 los principios de eficiencia y economía. Más en concreto, establece en el 103 que “*La Administración Pública sirve con objetividad los intereses generales y actúa de acuerdo con los principios de eficacia, jerarquía, descentralización, desconcentración y coordinación, con sometimiento pleno a la ley y al Derecho*”. Junto con ello, remite a la Ley la regulación de las garantías de imparcialidad de los funcionarios públicos en el ejercicio de sus funciones⁴¹. Aunque no referido específicamente a los poderes públicos, por su alcance general ha de mencionarse por último la encomienda que el art. 18.4 dirige al Legislador para limitar el uso de la informática para garantizar el pleno ejercicio de los derechos de los ciudadanos.

El procedimiento administrativo es el cauce formal que permite a la Administración Pública dar satisfacción a las necesidades públicas contempladas por la Constitución y las leyes y respetar el principio de legalidad. Desde esta óptica, la Exposición de Motivos de la LPAC señala que “*la esfera jurídica de derechos de los ciudadanos frente a la actuación de las Administraciones Públicas se encuentra protegida a través de una serie de instrumentos (...) de carácter preventivo, a través del procedimiento administrativo, que es expresión clara de que la Administración Pública actúa con sometimiento pleno a la Ley y al Derecho*”. Junto con ello, el principio de objetividad se proyecta

⁴¹ Para una visión general de las relaciones entre Constitución y tecnología, con interesantes propuestas sobre la necesidad de reconfigurar el sistema de fuentes, reordenar el sistema de derechos fundamentales a fin establecer una conexión directa entre éstos y las técnicas digitales y, más ampliamente, la constitucionalización de la tecnología, puede verse BALAGUER CALLEJÓN, F. (2022): *La Constitución del algoritmo*, Fundación Manuel Jiménez Abad, Zaragoza.

particularmente sobre el procedimiento administrativo, como se verá inmediatamente.

Aunque no existe en el texto constitucional un reconocimiento explícito del principio general de buena administración, tanto la doctrina como la jurisprudencia han derivado del mismo el deber de buena administración, la obligación legal de seguimiento de un procedimiento administrativo debido y el derecho a un procedimiento administrativo debido, todo ello con la finalidad de lograr decisiones administrativas de calidad. Se supera con ello la concepción clásica del procedimiento como proceso para la adopción de actos administrativos, vigente en el siglo XIX y a principios del siglo XX, para abrir paso a la idea del procedimiento como garantía de buena administración⁴².

En relación con este aspecto, la constante jurisprudencia del Tribunal Supremo ha derivado de diversos artículos de la Constitución la existencia de un principio implícito de buena administración, en línea con el art. 41 de la Carta de Derechos Fundamentales de la Unión Europea, que impone una obligación jurídica de debido cuidado o debida diligencia en el desarrollo de las funciones administrativas⁴³.

⁴² Por todos, PONCE SOLÉ, J. (2001): *Deber de buena administración y derecho al procedimiento administrativo debido. Las bases constitucionales del procedimiento administrativo y del ejercicio de la discrecionalidad*, Lex Nova, Valladolid; y, más recientemente, (2019): *La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida*, Cuadernos de la Cátedra de Democracia y Derechos Humanos, n° 15, Universidad de Alcalá de Henares-Defensor del Pueblo, Madrid. Del mismo autor, con un enfoque específico en lo relativo al uso de IA, PONCE SOLÉ (2018): “La prevención de riesgos de mala administración y corrupción, la inteligencia artificial y el derecho a una buena administración”, *Revista Internacional de Transparencia e Integridad*, n° 6, pp. 1-19.

⁴³ Un ejemplo lo encontramos en la Sentencia del 15 de octubre de 2020: “Es sabido que el principio de buena administración está implícito en nuestra Constitución (artículos 9.3, 103 y 106), ha sido positivizado en la Carta de Derechos Fundamentales de la Unión Europea (artículos 41 y 42) (...) constituye, según la mejor doctrina, un nuevo paradigma del Derecho del siglo XXI referido a un modo de actuación pública que excluye la gestión negligente (...) y –como esta misma Sala ha señalado en anteriores ocasiones– no consiste en una pura fórmula vacía de contenido, sino que se impone a las Administraciones Públicas, de suerte que el conjunto de derechos que de aquel principio derivan (audiencia, resolución en plazo, motivación, tratamiento eficaz y equitativo de los asuntos, buena fe) tiene –debe tener– plasmación efectiva y lleva aparejado, por ello, un

A nivel normativo, como es sabido, la LRJSP ha procedido a codificar en su art. 3 varios principios generales referidos a la actuación de las Administraciones Públicas. En particular, merecen mención especial los principios de participación, objetividad y transparencia de la actuación administrativa, así como los principios de buena fe, confianza legítima y lealtad institucional. Simplemente se mencionan, no se desarrollan sus contenidos. Asimismo, el Tribunal Supremo ha derivado también de este artículo el principio de buena administración⁴⁴.

Finalmente, con carácter general, ha de destacarse particularmente una idea ya mencionada. El art. 75 LPAC establece que *“los actos de instrucción necesarios para la determinación, conocimiento y comprobación de los hechos en virtud de los cuales deba pronunciarse la resolución, se realizarán de oficio y a través de medios electrónicos, por el órgano que tramite el procedimiento, sin perjuicio del derecho de los interesados a proponer aquellas actuaciones que requieran su intervención o constituyan trámites legal o reglamentariamente establecidos”*. Con ello se está configurando el derecho a un procedimiento debido e incorporando el deber de la Administración de actuar racionalmente, como ha sido afirmado con contundencia en la STS de 4 de abril de 2021⁴⁵.

correlativo elenco de deberes plenamente exigible por el ciudadano a los órganos públicos”.

⁴⁴ En la STS de 19 de febrero de 2019 se dice que: “Ya en otras ocasiones hemos hecho referencia al principio de buena administración, principio implícito en la Constitución, arts. 9.3 y 103, proyectado en numerosos pronunciamientos jurisprudenciales y positivizado, actualmente, en nuestro Derecho común, art. 3.1.e) de la Ley 40/2015; principio que impone a la Administración una conducta lo suficientemente diligente como para evitar definitivamente las posibles disfunciones derivadas de su actuación, sin que baste la mera observancia estricta de procedimientos y trámites, sino que más allá reclama la plena efectividad de garantías y derechos reconocidos legal y constitucionalmente al contribuyente y mandata a los responsables de gestionar el sistema impositivo, a la propia Administración Tributaria, observar el deber de cuidado y la debida diligencia para su efectividad y de garantizar la protección jurídica que haga inviable el enriquecimiento injusto”.

⁴⁵ STS de 14 de abril de 2021 (Rec. 28/2020): “el derecho al procedimiento administrativo debido, que es corolario del deber de buena administración, garantiza que las decisiones administrativas (...) se adopten de forma motivada y congruente con el iter procedimental, sin incurrir en desviación del procedimiento, en la medida que se requiere que no haya discordancias de carácter sustancial

En definitiva, el procedimiento –y los trámites y garantías contemplados legalmente en torno al mismo– es el cauce de realización de los principios de buena administración, transparencia y participación contemplados constitucionalmente.

Esta sucinta expresión de las derivadas del principio de buena administración referidas al procedimiento administrativo y de las normas de Derecho positivo a través de las cuales se ha concretizado en nuestro sistema, así como del alcance de uno y otras en la jurisprudencia, permite entender mejor una idea que ya ha sido anticipada: el uso de IA por la Administración contribuye a este ideal; pero no puede hacerse de cualquier modo.

Efectivamente, la pregunta que hemos de hacernos es por qué las eventuales ilegalidades derivadas de un uso deficiente de IA por parte de la Administración afectan de manera distinta a los derechos de los ciudadanos respecto a como lo hacen las actuaciones llevadas a cabo por el personal a su servicio. La respuesta es, cuanto menos, doble: de un lado, su impacto es mayor desde una perspectiva colectiva⁴⁶ y, junto con ello, se emplean en el ejercicio de potestades exorbitantes, lo que exige mayores cautelas⁴⁷; de otro, por la falta de empatía de las máquinas, lo que les dificulta la adopción de decisiones acordes con

entre los datos fácticos relevantes, la fundamentación jurídica obrante en el expediente y el contenido de la decisión administrativa. Este deber constitucional de buena administración, (...), exige que las Administraciones Públicas respeten el deber de motivación y los principios de objetividad, transparencia y racionalidad, conforme a lo dispuesto en los artículos 35 y 129 de la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas. En este contexto, referido al deber de la Administración de cumplir las garantías procedimentales, consideramos que se ha infringido el derecho al procedimiento debido, cuyo reconocimiento, como derecho procedimental de naturaleza constitucional, se infiere del artículo 24 de la Constitución, que garantiza el derecho a la tutela judicial efectiva, que se incardina en el derecho general de protección jurídica, y del artículo 103 del Texto Fundamental, que enuncia como principios rectores del actuar administrativo el principio de objetividad....”.

⁴⁶ PRESNO LINERA, M. (2022): *Derechos fundamentales e inteligencia artificial*, op. cit., p. 41.

⁴⁷ BOIX PALOP, A. (2022): “Transparencia en la utilización de inteligencia artificial...” op. cit., p. 91.

el principio de buena administración cuando existe cierto margen de valoración⁴⁸.

Como concreción específica de estas premisas, las Administraciones Públicas tienen la obligación de hacerse comprender ante actuaciones basadas en IA⁴⁹, lo que conduce a revisar y, en su caso, reforzar las exigencias de transparencia y motivación. Pero, junto con ello, no podemos no tener presente una clara conexión con el principio de buena administración: la necesidad de tramitar un procedimiento adecuado para la utilización de un sistema de IA en el ámbito administrativo⁵⁰ como garantía de transparencia (publicidad del sistema), participación (legitimidad del sistema) y legalidad (evaluación de impacto).

Estas garantías jurídicas han de ser reivindicadas y, cuando proceda, incorporadas en una visión de conjunto que integre también la perspectiva de la ciencia política que, en lo que ahora interesa, se puede sintetizar en los siguientes aspectos –en esencia, coincidentes–: el diseño de los sistemas ha de realizarse con cautela y mediante equipos multidisciplinares; deben ser objeto de evaluación previa por parte de una agencia especializada que los acredite; han de ser transparentes y permitir la motivación de las decisiones públicas adoptadas por medio de los mismos; y, en todo caso, en su configuración y funcionamiento debe intervenir personal propio de las Administraciones Públicas expertos en IA⁵¹.

En definitiva, el principio de buena administración se concreta en el mejor servicio al interés general combinado con el respeto de los intereses particulares legalmente reconocidos para la adopción de la

48 En este sentido, PONCE, J. (2022): “Reserva de humanidad y supervisión humana de la Inteligencia artificial”, en *El Cronista del Estado Social y Democrático de Derecho*, n° 100, p. 61. Sobre la empatía en relación con el uso de las TIC por parte de los poderes públicos, RANCHORDÁS, S. (2022): “Empathy in the Digital Administrative State”, en *Duke Law Journal*, n° 71, pp. 1341-1389.

49 PONCE, J. (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad...”, *op. cit.*, p. 40.

50 VALERO, J (2019): “Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración”, *Revista Catalana de Dret Public*, n° 58, p. 88 y PONCE, J. (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad...”, *op. cit.*, p. 35.

51 RAMIÓ, C. (2019): *Inteligencia Artificial y Administración Pública: Robots y humanos compartiendo el servicio público*, Catarata, Madrid, pp. 116-117.

mejor decisión en Derecho a través del procedimiento oportuno⁵². Ello debe proyectarse sobre la implantación de técnicas y herramientas de IA en el sector público y, en particular, sobre la aplicación del principio de transparencia.

4.- LA TRANSPARENCIA DE LA ACTIVIDAD ADMINISTRATIVA ALGORÍTMICA

4.1.- Premisa

Ciertamente, no son pocos los desafíos jurídicos que se derivan de la implantación de técnicas y herramientas de IA en el sector público. Sin embargo, sobre la base de las consideraciones anteriormente formuladas, puede concluirse que, desde la perspectiva formal, el principal de ellos es el relativo a la transparencia⁵³.

Estrechamente relacionado con el principio de buena administración se encuentra el principio de buena fe en las relaciones entre Administración y ciudadanos, igualmente reconocido en el art. 3 LRJSP. En ese contexto, puede ubicarse el principio de transparencia, entendido, en lo que ahora interesa, como posibilidad de conocer la existencia de procesos decisionales automatizados⁵⁴ y de su lógica de funcionamiento⁵⁵.

⁵² MENÉNDEZ SEBASTIÁN, E. (2021): *De la función consultiva clásica a la buena administración. Evolución en el Estado Social y Democrático de Derecho*, Marcial Pons, Madrid, p. 28.

⁵³ COBBE, J. (2019): “Administrative Law and the Machines of Government...”, op. cit., p. 5.

⁵⁴ GALETTA, D.U.: (2020): “Digitalizzazione e diritto ad una buona amministrazione...”, op. cit., p. 99.

⁵⁵ Para una completa visión sobre el principio de transparencia ante el empleo de IA y el uso de datos masivos, con análisis de algunos casos prácticos, puede verse COTINO, L. (2021): “Hacia la Transparencia 4.0: el uso de la Inteligencia Artificial y *big data* para la lucha contra el fraude y la corrupción y las (muchas) exigencias constitucionales”, en RAMIÓ, C. (Coord.), *Repensando la Administración digital y la innovación pública*, Instituto Nacional de Administración Pública, Madrid, 2021, pp. 169-196. Acerca del alcance de las obligaciones de transparencia y sus conexiones en relación con el uso de IA, COTINO, L. (2022): “Transparencia y explicabilidad de la inteligencia artificial y “compañía” (comu-

Hablar de transparencia de la actividad administrativa en el contexto del uso de IA por las Administraciones requiere poner de manifiesto una cuestión inicial: los humanos estamos lejos de ser transparentes en nuestras actuaciones⁵⁶. Hemos diseñado en nuestro ordenamiento jurídico un sistema formal que entiende la transparencia como la publicidad de un listado de informaciones variadas que se consideran relevantes a los efectos de justificar la acción administrativa y rendir cuentas ante los ciudadanos y, junto con ello, la posibilidad de pedir acceso –siempre mediado por la Administración– a información en poder de ésta. Ciertamente la transparencia es un principio que resulta igualmente aplicable a la actuación administrativa algorítmica y que, como en el caso anterior, tiene sus propios retos y desafíos. Pero es posible construir un conjunto de derechos y obligaciones que permitan aceptar su puesta en funcionamiento en el ámbito de las Administraciones Públicas con garantías específicamente pensadas para tal fin.

4.2.- Un ejercicio práctico: regular la transparencia de la actividad administrativa algorítmica

El reto de la transparencia es susceptible de ser abordado desde el Derecho de múltiples formas. Considero que puede resultar interesante enfocarlo desde la perspectiva regulatoria y, en particular, analizar cómo ha sido tratada la cuestión en la Carta de Derechos Digitales aprobada por el Gobierno de España a través de la Secretaría de Estado de Digitalización e Inteligencia Artificial con la colaboración de un

nicación, interpretabilidad, auditabilidad, testabilidad, comprobabilidad, simulabilidad...). Para qué, para quién y cuánta”, en COTINO, L. y CASTELLANOS, J. (Eds.): *Transparencia y explicabilidad de la inteligencia artificial*, Tirant Lo Blanch, Valencia, 2022, pp. 25-70.

En relación con el sistema italiano, OROFINO, A. G. (2020): “The Implementation of the Transparency Principle in the Development of Electronic Administration”, *European Review of Digital Administration & Law*, vol. 1, pp. 123-142, disponible en <https://www.erdalreview.eu/free-download/978882553896012.pdf> (Última visita: 30/12/2022).

⁵⁶ COGLIANESE, C. y LAI, A. (2022): “Algorithm vs. algorithm”, *Duke Law Journal*, nº 71, p. 1313.

grupo de expertos⁵⁷. Resulta de interés hacerlo así porque se trata de la primera tentativa de regulación general de esta cuestión y, al mismo tiempo, constituye una oportuna aplicación práctica para comprender mejor el papel del Derecho en la regulación de la tecnología. En definitiva, permite testar la metodología presentada al inicio de estas páginas⁵⁸.

Ciertamente, uno de los grandes riesgos derivados del uso de IA es la opacidad. Los documentos europeos sobre la materia⁵⁹ parten de

⁵⁷ Disponible en https://www.lamonicloa.gob.es/presidente/actividades/Documents/2021/140721-Carta_Derechos_Digitales_RedEs.pdf (Última visita: 30/12/2022).

⁵⁸ Existen otros posibles enfoques, también desde la perspectiva regulatoria. Quizás el más relevante de ellos es el que se deriva de la aplicación de la normativa –europea y nacional– reguladora de la protección de datos de carácter personal, que contiene importantes exigencias de transparencia acerca del uso de sistemas automatizados para la toma de decisiones y de acceso a los datos empleados. Sobre esta cuestión, puede verse MEDINA GUERRERO, M. (2022): “El derecho a conocer los algoritmos utilizados en la toma de decisiones. Aproximación desde la perspectiva del derecho fundamental a la protección de datos personales”, en *Teoría y Realidad Constitucional*, nº 49, pp. 141-171.

⁵⁹ En particular, Directrices Éticas para una IA fiable (<https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>, Última visita: 30/12/2022), publicado en abril de 2019 y elaborado por el Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea en junio de 2018. Tal y como se indica en el mismo, “la fiabilidad de la IA se apoya en tres componentes que deben satisfacerse a lo largo de todo el ciclo de vida del sistema: a) la IA debe ser lícita, es decir, cumplir todas las leyes y reglamentos aplicables; b) ha de ser ética, de modo que se garantice el respeto de los principios y valores éticos; y c) debe ser robusta, tanto desde el punto de vista técnico como social, puesto que los sistemas de IA, incluso si las intenciones son buenas, pueden provocar daños accidentales. Cada uno de estos componentes es en sí mismo necesario pero no suficiente para el logro de una IA fiable. Lo ideal es que todos ellos actúen en armonía y de manera simultánea. En el caso de que surjan tensiones entre ellos en la práctica, la sociedad deberá esforzarse por resolverlas”. En el propio documento se especifican los contenidos principales de cada uno de estos tres pilares. De este modo, por IA lícita ha de entenderse aquella que observa las obligaciones jurídicas negativas (qué no puede hacer) y positivas (qué debe hacerse), que deberán estar fundadas en los derechos fundamentales (respeto de la dignidad humana, la libertad individual, la democracia, la justicia y el Estado de Derecho, la igualdad, no discriminación y solidaridad y, más ampliamente, los derechos de los ciudadanos ante los poderes públicos); la IA ética es la que resulta coherente con las normas éticas

derivadas de los principios de respeto de la autonomía humana, la prevención de daños, la equidad y la explicabilidad; la IA robusta, finalmente, es la que logra que los sistemas funcionen de manera segura a nivel técnico. Junto con ello, se especifican en el documento una serie de principios que articulan la realización de la IA fiable: acción y supervisión humanas (evaluación del impacto sobre los derechos fundamentales y mantenimiento de la autonomía de los destinatarios en relación con los sistemas de IA, así como garantía de control sobre su funcionamiento y resultados); solidez técnica y seguridad (enfoque preventivo de los riesgos y protección frente a vulnerabilidades); gestión de la privacidad y de los datos (garantía de calidad de los datos y protección de la intimidad); transparencia (trazabilidad, explicabilidad y comunicación); diversidad, no discriminación y equidad (evitación de sesgos y accesibilidad y diseño universal, con participación de los interesados en el diseño y durante el despliegue); bienestar social y ambiental (sostenibilidad y enfoque dirigido a la mejora de la sociedad); y rendición de cuentas (auditabilidad de los sistemas, conocimiento de sus riesgos y compensación ante eventuales daños).

Más recientemente ha sido aprobada la Declaración Europea sobre los Derechos y Principios Digitales para la Década Digital (COM (2022) 28 final, de 26 de enero), acompañada de una Comunicación de la Comisión en la que se explica la iniciativa, más en la línea del marketing jurídico que de la innovación en Derecho (Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social y al Comité de las Regiones estableciendo una Declaración Europea de Derechos Digitales y principios para la Década Digital, COM (2022) 27 final, de 26 de enero). En ella se insiste en la necesidad de situar al ciudadano en el centro de la transición digital y en su aplicabilidad a las Administraciones Públicas. Sin embargo, se opta por una declaración de derechos de naturaleza política y se confiesa que no todos los principios que en ella se recogen se corresponden con derechos directamente aplicables. Llamen la atención dos extremos relativos al planteamiento. De un lado, el enfoque transmite la idea de que el mundo digital es diferente al mundo físico; sin embargo, uno y otro forman parte de nuestro mundo real, en el que nos movemos y actuamos las personas, en el que nos relacionamos entre nosotros y con las Administraciones Públicas. De otro, que se opte por explicar a través de ella intenciones políticas y no por regular nuevos derechos para proteger a los ciudadanos frente a los potenciales riesgos de la tecnología y para incorporar obligaciones que garanticen el cumplimiento de los mismos. A ello ha de sumarse el carácter declarativo. En cualquier caso, todo texto jurídico, aun no siendo vinculante, está llamado a producir efectos jurídicos. Por ello debe valorarse positivamente que se integre en la Declaración el valor de las personas como núcleo de la transformación digital; el carácter vicarial de la tecnología, que ha de estar siempre al servicio y usarse para el beneficio de los ciudadanos; y, sobre todo, el acceso a la totalidad de los servicios públicos. Adicionalmente, resulta de interés el Capítulo III que, bajo el título “Libertad de elección”, contiene un apartado relativo a las interacciones con algoritmos y sistemas de inteligencia artificial. En él se integra un principio y varios compromisos –políticos–: el principio de empoderamiento de toda perso-

tres exigencias a la hora de aplicar la transparencia a las técnicas de IA: trazabilidad (capacidad para llevar a cabo un seguimiento documentado de los datos, proceso y desarrollo de despliegue de un sistema de IA), explicabilidad (capacidad de explicar los procesos técnicos de un sistema de IA y las decisiones humanas asociadas al mismo) y auditabilidad (capacidad de un sistema de IA de someterse a la evaluación de sus algoritmos, datos y procesos de diseño). Efectivamente, resulta fundamental documentar los datos de los que se nutre el sistema y el proceso que sigue para su funcionamiento, dar a conocer a las personas afectadas cómo opera, así como sus capacidades y limitaciones y, más en general, participar en su diseño y aplicación.

En este sentido, y por contraposición, la opacidad de los sistemas algorítmicos puede ser de tres tipos⁶⁰: jurídica, esto es, derivada de las exigencias de protección de otros derechos e intereses (por ejemplo, la propiedad intelectual); sociológica, es decir, derivada de la escasa capacidad para conocer y comprender cómo funcionan desde la perspectiva tecnológica; e intrínseca, derivada del funcionamiento del propio sistema. En mi opinión, el primero de ellos es necesario; el segundo ha de irse corrigiendo paulatinamente, fomentando las habilidades tecnológicas de los empleados públicos y de los ciudadanos; el tercero puede mitigarse a través de la investigación y el desarrollo de la propia tecnología. Sin embargo, junto con todos ellos, el tipo de opacidad que ha de combatirse desde el Derecho, porque carece de justificación, es la que podríamos calificar de opacidad intencionada, esto es, la derivada de la ocultación por la Administración del uso

na para beneficiarse de las ventajas de la IA a fin de tomar sus propias decisiones con conocimiento de causa en el entorno digital y las obligaciones de velar por la transparencia en el uso de algoritmos; de asegurar que los sistemas algorítmicos se basen en conjuntos de datos adecuados para evitar la discriminación y permitan la supervisión humana de los resultados que afectan a personas; de garantizar que no se emplean para predeterminar las decisiones de las personas; y de salvaguardar que estos sistemas son seguros y se utilizan con pleno respeto a los derechos fundamentales de las personas.

⁶⁰ COBBE, J. (2019): “Administrative Law and the Machines of Government...”, *op. cit.*, p. 5. En el mismo sentido, CERRILLO, A. (2020): “El impacto de la inteligencia artificial en las Administraciones Públicas: estado de la cuestión y una agenda”, en A. CERRILLO I MARTÍNEZ; M. PEGUERA POCH: *Retos jurídicos de la inteligencia artificial*, Thomson Reuters Aranzadi, Cizur Menor, pp. 83-84.

de estos sistemas algorítmicos y de los detalles acerca de la forma de funcionamiento de los mismos.

Como punto de partida, han de tomarse en consideración dos premisas:

- A nivel práctico, las aplicaciones y soluciones de IA utilizadas en estos momentos por la Administración, en términos generales, no han sido aprobadas previamente con carácter formal y ni siquiera son conocidas, en la mayor parte de los casos, más allá de notas de prensa o documentos técnicos que hacen mención a las mismas. Ello genera opacidad, inseguridad jurídica y desconfianza, influyendo decisivamente sobre la legitimidad de la actuación de los poderes públicos.
- Desde una perspectiva teórica, como ha sido anticipado, el empleo de IA en el ámbito público es una manifestación del deber de la Administración de perseguir con eficacia y objetividad la satisfacción de los intereses generales previsto en el art. 103 CE y del principio de buena administración, que ha de conciliarse con los derechos y garantías que la propia Constitución y el resto del ordenamiento jurídico otorgan a los ciudadanos en sus relaciones con los poderes públicos y, en particular, de los que se derivan del derecho a una buena administración. En definitiva, es un instrumento, no un sustituto del concepto de Administración Pública tal y como la configura nuestro texto constitucional.

A ello se ha de unir una tercera idea general: el uso de técnicas y herramientas de IA no nos viene dado ni impuesto. Somos nosotros, desde el Derecho, la Política y la Técnica, quienes hemos de determinar los instrumentos que mejor se adapten a nuestras necesidades, el espacio de actuación, el método de aplicación y las garantías exigibles. Una de ellas, fundamental, es la relativa a la transparencia conectada con la motivación y el control.

La transparencia, como se ha recordado con anterioridad, es un principio general de la actuación administrativa (consagrado en el art. 3 LRJSP) y un derecho constitucional (reconocido en el art. 105 CE y desarrollado por la LTBG). Como principio y como derecho, posee múltiples manifestaciones y aplicaciones prácticas en diferentes

sectores del ordenamiento jurídico-administrativo. Pero son manifestaciones de un mismo principio y de un mismo derecho. La trazabilidad y la explicabilidad –entendidas como capacidad para explicar los procesos técnicos seguidos y las decisiones de ellos derivadas– resultan imprescindibles no sólo para generar confianza en el sistema algorítmico⁶¹, sino también y sobre todo para su reconocimiento legal cuando a través del mismo se adoptan decisiones. No en vano, la capacidad de interpretar por parte de los humanos el proceso tecnológico seguido por el sistema guarda conexión directa con el deber de motivación en tanto que permite explicar el contexto fáctico⁶².

En consecuencia, hablar de transparencia en el uso de la IA en el sector público exige tomar en consideración una doble manifestación: de un lado, la transparencia externa, referida a la difusión de la información relativa a la existencia de sistemas que hacen uso de técnicas de IA en el sector público, los organismos responsables de su configuración y gestión, las empresas que los han diseñado, las condiciones de la contratación, etc.; de otro, la transparencia interna, referida al funcionamiento práctico de la solución de la que se hace uso: espacio de aplicación, carácter decisorio o no decisorio, tipo de tecnología empleada, forma de razonamiento lógico del sistema, etc. La primera guarda conexión con la transparencia en sentido estricto, esto es, con la obligación de dar a conocer determinada información por ser relevante para los ciudadanos desde la perspectiva de la comprensión de qué hace y cómo actúa el poder público⁶³; la segunda, por el contrario, con la

⁶¹ MORENO REBATO, M. (2021): *Inteligencia Artificial (Umbrales éticos, Derecho y Administraciones Públicas)*..., op. cit., p. 77.

⁶² PALMIOTTO ETTORE, F. (2022): “The Right to Contest Automated Decisions”, *The Digital Constitutionalist*, disponible en <https://digi-con.org/the-right-to-contest-automated-decisions/> (Última visita: 30/12/2022)

⁶³ El Real Decreto 203/2021, de 30 de marzo, por el que se aprueba el Reglamento de actuación y funcionamiento del sector público por medios electrónicos ha incorporado en su art. 13 una doble novedad relevante para el tema que nos ocupa (si bien solo aplicable a la Administración General del Estado): de un lado, el deber de autorizar la determinación de una actuación administrativa como automatizada por resolución del titular del órgano administrativo competente; de otro, la obligación de publicarla en su sede electrónica. En mi opinión, partiendo del concepto amplio de actividad administrativa algorítmica que se defiende en estas páginas, esta doble exigencia resulta de aplicación respecto de los sistemas algorítmicos.

motivación de las decisiones adoptadas por o sobre la base de sistemas algorítmicos para evidenciar por qué lo hace⁶⁴.

Sobre la base de estas premisas, ha de tenerse presente que la transparencia en sentido estricto tiene una doble vertiente: proactiva y reactiva. Dos son los preceptos de la Carta de Derechos Digitales en los que vale la pena detenerse a la hora de abordar esta cuestión: el XVIII y el XXV. El primero de ellos, referido a los derechos digitales de la ciudadanía en sus relaciones con las Administraciones Públicas, incorpora dos exigencias de carácter general que resultan poco novedosas, por estar reconocidas en nuestro ordenamiento jurídico por normas con rango legal: en primer lugar, la aplicabilidad del principio de transparencia al entorno digital y, en particular, la garantía del derecho de acceso a la información pública y del cumplimiento de las obligaciones de publicidad activa; y, en segundo lugar, la necesidad de identificar a los órganos responsables de toda actividad pública realizada en el entorno digital. Constituye una novedad muy relevante, por el contrario, lo dispuesto en los apartados 6 y 7, que contienen una consideración general y tres garantías específicas directamente relacionadas con la transparencia y el correcto funcionamiento de los algoritmos aplicados en el ámbito administrativo.

La consideración general es que los derechos de la ciudadanía en relación con la inteligencia artificial reconocidos en la Carta resultarán de aplicación igualmente en el marco de la actuación administrativa. Efectivamente, el art. XXV requiere que se asegure un enfoque centrado en la persona y precisa que en el desarrollo y ciclo de vida de los sistemas de inteligencia artificial se deberá garantizar el derecho a la no discriminación algorítmica; la transparencia, audibilidad, aplicabilidad y trazabilidad; y la accesibilidad, usabilidad y fiabilidad. Junto con ello, reconoce el derecho a solicitar una supervisión e intervención humana y a impugnar las decisiones automatizadas basadas

⁶⁴ En este sentido, G. COGLIANESE y D. LEHR diferencian entre transparencia “de escaparate” y transparencia razonada, considerando que la primera permite entender qué hace la Administración y la segunda por qué lo hace; ciertamente, una y otra están relacionadas entre sí: ésta depende de aquélla, en tanto que para explicar por qué hace algo, la Administración debe previamente publicitar qué está haciendo. COGLIANESE, G. y LEHR, D. (2019): “Transparency and algorithmic governance”, *Administrative Law Review*, n° 71, pp. 9-15.

en sistemas de IA que produzcan efectos en su esfera personal y patrimonial.

Como garantías específicas –relacionadas con la transparencia interna–, de un lado se contempla el derecho a obtener una motivación comprensible en lenguaje natural de las decisiones que se adopten en el entorno digital, con justificación de las normas jurídicas relevantes, de la tecnología empleada y de los criterios de aplicación de las mismas. De otro, se incorpora la necesidad de evaluar el impacto sobre los derechos digitales del diseño de los algoritmos para la toma de decisiones automatizadas o semiautomatizadas y se reserva con carácter general la adopción de decisiones discrecionales a personas. Finalmente, se ha incorporado el derecho a la transparencia sobre el uso de instrumentos de IA, su funcionamiento y alcance en cada procedimiento, con especial atención a los datos utilizados, margen de error, ámbito de aplicación y carácter decisorio o no decisorio; además, se remite a la ley la determinación de las condiciones de acceso al código fuente con el objeto de verificar que no produce resultados discriminatorios⁶⁵.

Más allá de valorar muy positivamente estas previsiones, por lo que implican, pueden formularse dos consideraciones. En primer lugar, se trata de garantías de orden formal que, sin embargo, deberían ir acompañadas de una garantía de carácter institucional: la atribución a un órgano específico, en cada nivel de Administración –o, previo convenio, con competencia sobre todas ellas–, de la función de verificación y aprobación previa de los sistemas algorítmicos. En segundo lugar, en este tipo de previsiones se aprecia muy bien la debilidad de las proclamaciones de la Carta al carecer de carácter jurídico vinculante. Las novedades que trae consigo, no contempladas en otras normas, podrán ser desconocidas por las diferentes Administraciones sin que de ello se derive consecuencia alguna. Dicho sencillamente, se trata de meras indicaciones. Aun siendo conscientes de que la Carta

⁶⁵ Estas novedades recogen parcialmente la propuesta planteada por la Red DAIA en el proceso de consulta pública. Vid. <http://reddaia.org/documentos/> y el listado de propuestas recibidas https://portal.mineco.gob.es/es-es/ministerio/participacionpublica/audienciapublica/Paginas/SEDIA_Carta_Derechos_Digitales.aspx (Última visita: 30/12/2022).

ha sido redactada como si fuera a tener carácter jurídico vinculante, con la finalidad de inspirar y orientar las distintas normas que regulan estas cuestiones –y, precisamente por ello, la aprobación y entrada en vigor de la misma generará en el futuro efectos jurídicos positivos a través de su aplicación por los distintos operadores jurídicos y, en particular, por el recurso a la misma como fuente de inspiración y recurso de interpretación por parte de jueces y tribunales o, incluso, a través de su incorporación a normas jurídicas vinculantes–, no menos cierto es que ello dependerá de operadores jurídicos, Jueces y Tribunales, Gobiernos y Legisladores, y no directamente de la voluntad de la propia Carta. El Derecho no lo puede todo y necesita de políticas públicas para lograr los fines que se buscan con las normas; pero el carácter jurídico vinculante es la principal herramienta de la que se vale el Derecho para perseguir esos fines. Y el contexto en el que estamos exige la entrada en juego de previsiones normativas vinculantes sobre estos extremos.

En cualquier caso, resulta fundamental establecer una exigencia preliminar para la viabilidad jurídica de los sistemas algorítmicos: documentar el método de programación inicial, los datos de entrada recopilados y seleccionados por los mismos, la forma en que se hizo, los ensayos efectuados y su validación⁶⁶. En concreto, los deberes de publicidad activa deben abarcar igualmente los fallos que se detecten en el funcionamiento del sistema. Todo ello resulta imprescindible no sólo desde la perspectiva de la transparencia en sentido estricto, sino –y sobre todo– desde la óptica de la motivación y el control. De hecho, la posibilidad de conocer estos extremos es premisa para el control de legalidad de los mismos.

4.3.- Publicidad activa y derecho de acceso en relación con el uso de IA en el sector público

Como ha sido anticipado, las previsiones de la Carta de Derechos Digitales no poseen carácter jurídico vinculante. Sin embargo, ello no significa que no resulte exigible a día de hoy ningún deber de trans-

⁶⁶ MORENO REBATO, M. (2021): *Inteligencia Artificial (Umbrales éticos, Derecho y Administraciones Públicas)*..., op. cit., p. 78.

parencia –en la doble manifestación indicada– a las Administraciones Públicas que ejercen sus competencias basándose en sistemas algorítmicos. Antes al contrario, tanto la LTBG como la LPAC contienen previsiones claras que resultan de aplicación en este ámbito.

En primer lugar, puede sostenerse la existencia de una obligación de publicidad activa –respecto de la operatividad de los propios sistemas algorítmicos, de su finalidad, de qué datos se nutren y de cómo operan–, derivada de lo dispuesto en el art. 5.1 LTBG. En virtud del mismo, las Administraciones Públicas tienen la obligación de publicar de forma periódica y actualizada *“la información cuyo conocimiento sea relevante para garantizar la transparencia de su actividad relacionada con el funcionamiento y control de la actuación pública”*. El uso de sistemas algorítmicos para la toma de decisiones administrativas, ciertamente, es una categoría relevante para estos fines y, por tanto, la información relativa a su existencia, ámbito de actuación y alcance debería estar disponible en las sedes electrónicas de los organismos públicos que hacen uso de los mismos⁶⁷.

Desde esta perspectiva, ha de destacarse la novedad introducida por la Ley 1/2022, de 13 de abril, de la Generalitat, de Transparencia y Buen Gobierno de la Comunitat Valenciana, cuyo artículo 16 contempla entre las obligaciones de publicidad activa relativas a información de relevancia jurídica *“la relación de sistemas algorítmicos o de inteligencia artificial que tengan impacto en los procedimientos administrativos o la prestación de los servicios públicos con la descripción de manera comprensible de su diseño y funcionamiento, el nivel de riesgo que implican y el punto de contacto al que poder dirigirse en cada caso, de acuerdo con los principios de transparencia y explicabilidad”*.

Junto con ello, no resulta descabellado entender que resulta de aplicación igualmente lo previsto en el art. 7 a) LTBG. Efectivamente, el empleo de sistemas algorítmicos en la toma de decisiones supone la concreción de la aplicación de la Ley en un sector específico de actuación de una concreta Administración Pública o, dicho de otra manera, una específica interpretación del Derecho con efectos jurídicos. Con

⁶⁷ Así lo ha defendido PONCE SOLÉ, J. (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad...”, p. 45.

base en esta idea, deberían hacerse públicos todos aquellos casos de uso de IA por parte del sector público que se empleen para la toma de decisiones administrativas⁶⁸.

En relación con esta cuestión, si bien de forma muy genérica, el art. 23 de la reciente Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación⁶⁹ ha apostado por la transparencia de la actividad administrativa algorítmica a través de una triple línea de acción: la incorporación de mecanismos que tengan en cuenta el principio de transparencia y rendición de cuentas respecto de los algoritmos involucrados en la toma de decisiones por parte de las Administraciones Públicas; la interpretabilidad de las decisiones adoptadas sobre la base de algoritmos; y la introducción de sellos de calidad y de evaluaciones de impacto (que, por su propia naturaleza de información administrativa, quedan incorporados en el ámbito objetivo de aplicación de la LTBG). A pesar de su vocación de conver-

⁶⁸ BOIX invoca el principio de seguridad jurídica para defender la publicidad total de los sistemas algorítmicos, incluyendo su código fuente; BOIX A. (2020): “Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la Administración para la adopción de decisiones”, *Revista de Derecho Público: Teoría y Método*, Vol. I, pp. 254-255.

⁶⁹ El precepto, bajo el título “Inteligencia Artificial y mecanismos de toma de decisión automatizados”, establece lo siguiente:

1. *En el marco de la Estrategia Nacional de Inteligencia Artificial, de la Carta de Derechos Digitales y de las iniciativas europeas en torno a la Inteligencia Artificial, las administraciones públicas favorecerán la puesta en marcha de mecanismos para que los algoritmos involucrados en la toma de decisiones que se utilicen en las administraciones públicas tengan en cuenta criterios de minimización de sesgos, transparencia y rendición de cuentas, siempre que sea factible técnicamente. En estos mecanismos se incluirán su diseño y datos de entrenamiento, y abordarán su potencial impacto discriminatorio. Para lograr este fin, se promoverá la realización de evaluaciones de impacto que determinen el posible sesgo discriminatorio.*

2. *Las administraciones públicas, en el marco de sus competencias en el ámbito de los algoritmos involucrados en procesos de toma de decisiones, priorizarán la transparencia en el diseño y la implementación y la capacidad de interpretación de las decisiones adoptadas por los mismos.*

3. *Las administraciones públicas y las empresas promoverán el uso de una Inteligencia Artificial ética, confiable y respetuosa con los derechos fundamentales, siguiendo especialmente las recomendaciones de la Unión Europea en este sentido.*

4. *Se promoverá un sello de calidad de los algoritmos.”.*

tirse en el mínimo común normativo que contenga las definiciones y las garantías fundamentales del Derecho antidiscriminatorio español –expresada en su Exposición de Motivos–, el alcance de tal apuesta por cuanto se refiere al uso de IA es muy limitado, por dos razones: De un lado, el cumplimiento de estas exigencias queda sujeto a la posibilidad técnica de la incorporación de tales mecanismos; de otro, la propia generalidad de la literalidad del texto anima a una interpretación flexible de las obligaciones que incorpora. No obstante, debe valorarse muy positivamente el hecho de que paulatinamente se vayan incorporando exigencias legales específicas en relación con el uso de IA por la Administración.

En segundo lugar, el art. 35 de la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas es claro al establecer el deber de motivación de los actos administrativos que afecten a los derechos o intereses de los ciudadanos⁷⁰. Efectivamente, desde una perspectiva formal, el uso de sistemas algorítmicos ni libera de la exigencia de motivar ni supone un cambio sustancial en cuanto al contenido de la motivación de aquellos actos administrativos adoptados por o sobre la base de los mismos⁷¹. Sin embargo, aunque las exigencias existen y pueden entenderse aplicables al uso de IA por la Administración, resulta conveniente reforzar las garantías que se derivan de ellas. Así, se deben fortalecer las obligaciones de transparencia relativas a la elaboración del modelo y al diseño del sistema⁷², incluyendo expresamente entre el catálogo de obligaciones de publicidad activa de la Ley básica estatal la información relativa las especificaciones técnicas del sistema algorítmico,

⁷⁰ Sobre esta cuestión, GAMERO CASADO, E. (2021): “Compliance (o cumplimiento normativo) de desarrollos de Inteligencia Artificial para la toma de decisiones administrativas”, en *Diario La Ley*, n° 50, 19 de abril de 2021, in totum.

⁷¹ CARULLO, G. (2021): “Decisione amministrativa e intelligenza artificiale”, *Diritto dell'informazione e dell'informatica*, n° 3, p. 440.

⁷² VALERO, J. (2019): “Las garantías jurídicas de la inteligencia artificial ...”, *op. cit.*, p. 88.; en el mismo sentido, en relación con el sistema italiano, PINOTTI, G. (2021): “Amministrazione digitale algoritmica e garanzie procedimentali”, *Labour & Law Issues*, vol. 7, n° 1, p. 92. Para HUERGO, en la medida en que el algoritmo aporte algún elemento nuevo en la aplicación de la norma, que no sea replicable sin el algoritmo, es exigible la total transparencia respecto a su funcionamiento, HUERGO LORA, A. (2020), “Una aproximación a los algoritmos desde el Derecho Administrativo...”, *op. cit.*, p. 75.

los datos de los que se nutre, los resultados de su entrenamiento y las eventuales auditorías que se hayan podido realizar; todo ello, para evitar interpretaciones contrarias al principio de transparencia en la práctica; junto con ello, la motivación, aun breve y sucinta, debe ser razonable y adaptarse a la singularidad de la tecnología⁷³; adicionalmente, desde el momento en el que la predicción algorítmica ha servido, directa o indirectamente, para adoptar una decisión administrativa, ha de integrarse en la motivación y, además, el algoritmo deberá formar parte del expediente⁷⁴.

Finalmente, debe valorarse convenientemente si resulta imprescindible desde la perspectiva de la motivación dar acceso al código fuente⁷⁵. Resulta patente que sólo a través del acceso y valoración del có-

⁷³ VALERO, J. (2021): “Las garantías jurídicas de la inteligencia artificial ...”, *op. cit.*, p. 88.

⁷⁴ HUERGO, A. (2021): “Administraciones Públicas e inteligencia artificial...”, *op. cit.*, p. 89.

⁷⁵ Partidarios de ofrecer acceso al código fuente se muestran DE LA CUEVA, J. (2019): “La importancia del código fuente”, en CAPILLA RONCERO, F.S. (Dir.), *Derecho digital: Retos y cuestiones actuales*, Thomson Reuters Aranzadi, Cizur Menor, pp. 109-127; y GÓMEZ JIMÉNEZ, M.L. (2021): *Automatización procedimental y sesgo electrónico: el procedimiento administrativo electrónico desde la Inteligencia Artificial*, Thomson-Reuters Aranzadi, Cizur Menor, p. 150. Por el contrario, en relación con los algoritmos no predictivos, HUERGO LORA entiende que, con la regulación vigente, no es necesario dar a conocer al ciudadano que la decisión ha sido elaborada con la ayuda de una aplicación informática ni, con carácter general, poner a su disposición el código fuente (cuestión distinta es que una norma expresamente lo ordenara); por el contrario, en el caso de los algoritmos predictivos –por su carácter potencialmente innovador derivado del hecho de que añaden contenido de elaboración propia– sí considera que la Administración debe hacer público su uso cuando sirven para la adopción de decisiones, en tanto que constituyen antecedente y fundamento de la decisión administrativa y, por ello, deben incorporarse al expediente, HUERGO, A. (2020): “Una aproximación a los algoritmos desde el Derecho Administrativo...”, *op. cit.*, pp. 72 y 85. En la misma línea, G.M. DÍAZ GONZÁLEZ entiende que la publicación del código fuente no es una solución satisfactoria en todos los casos por la falta de conocimientos por parte del ciudadano para entender el funcionamiento del sistema y por el hecho de que el elevado grado de complejidad de algunos de ellos puede implicar imposibilidad de comprender su funcionamiento, a parte de que puede afectar negativamente a la propia función administrativa que se busca con el algoritmo, DÍAZ GONZÁLEZ, G. M. (2020): “Algoritmos y actuación policial: la policía predictiva”, en HUERGO LORA, A. (Dir.): *La regulación de los algoritmos*, Thomson-Aranzadi, Cizur Menor, p. 189.

digo fuente puede conocerse cómo opera el sistema desde el punto de vista tecnológico y, en su caso, detectar eventuales errores⁷⁶; al mismo tiempo, hace factible la posibilidad de comprobar si la programación del sistema algorítmico es o no conforme con lo previsto por la norma a cuya aplicación sirve⁷⁷; es más, se ha sostenido la existencia de una conexión directa entre la apertura del código desde la perspectiva de su diseño –esto es, optar por sistemas algorítmicos de código abierto– y el principio de Estado democrático e, incluso, el principio de jerarquía en el seno de la organización administrativa, en tanto que, de lo contrario, los niveles superiores de empleados públicos, a quienes se responsabiliza de las decisiones, dependerían de quienes tengan encomendada la programación o supervisión del sistema⁷⁸. Sin embargo, una decisión de esta naturaleza tiene inconvenientes, derivados esencialmente de la posibilidad de “engañar” al sistema⁷⁹.

Esta cuestión se ha planteado ya en nuestro país desde la perspectiva del ejercicio del derecho de acceso a la información pública en aplicación de la normativa sobre transparencia. No procede dar aquí

Un trabajo muy exhaustivo sobre esta cuestión, con un relevante análisis de Derecho Comparado, puede verse en GUTIÉRREZ DAVID, M.E. (2021): “Administraciones inteligentes y acceso al código fuente de los algoritmos públicos. Conjurando riesgos de cajas negras decisionales”, *Derecom*, n° 30, pp. 143 y ss.

⁷⁶ CITRON, D.K. (2008): “Open Code Governance”..., *op. cit.*, p. 357.

⁷⁷ CARULLO, G. (2021): “Decisione amministrativa e intelligenza artificiale”..., *op. cit.*, p. 441. Para este autor, el código fuente debería incorporarse en la motivación como texto de la misma, al menos la parte que permita verificar la correspondencia entre las previsiones normativas y su traducción a lenguaje informático e, incluso, publicarse en la sede electrónica correspondiente.

⁷⁸ CITRON, D.K. (2008): “Open Code Governance”..., *op. cit.*, pp. 358 y ss.

⁷⁹ GUTIÉRREZ DAVID sintetiza las críticas a la apertura total en cinco extremos: el impacto en la privacidad, por el riesgo de filtración de datos personales confidenciales; la posibilidad de engaño de los modelos orientados al cumplimiento normativo; el riesgo de prácticas contrarias a la competencia derivadas de la accesibilidad de los algoritmos por terceros con afectación al derecho de propiedad intelectual y al secreto comercial; la naturaleza dinámica y adaptativa de los modelos, que se reajustan sobre la base de los datos de entrada; y la opacidad inherente a los modelos de decisión algorítmica; GUTIÉRREZ DAVID, M.E. (2021): “Administraciones inteligentes y acceso al código fuente...”, *op. cit.*, p. 177.

puntual cuenta de esos casos, suficientemente analizados por la doctrina⁸⁰, sino más bien presentar una síntesis del enfoque.

Efectivamente, el Consejo de Transparencia y Buen Gobierno ha tenido ocasión de pronunciarse acerca de este concreto tema con motivo de varias reclamaciones frente a desestimaciones de solicitudes de acceso al código fuente de determinadas aplicaciones informáticas. En primer lugar, ha de mencionarse la Resolución 701/2018, de 18 de febrero de 2019, en la que se afirma que el código fuente es información pública y, por tanto, puede ser objeto del derecho de acceso, aunque en el caso concreto se deniega por entender que resulta aplicable el límite de la propiedad intelectual, a pesar de que la aplicación –sistema BOSCO, utilizado por parte de los comercializadores de referencia en el ámbito energético para determinar si un consumidor tiene la condición de vulnerable y, por tanto, ha de recibir bonificación en la correspondiente factura– había sido diseñada por la propia Administración⁸¹. Pero, sobre todo, de particular interés es la Resolución 58/2021, de 20 de mayo de 2021, relativa a una solicitud de acceso al algoritmo de cálculo de las pensiones de la Seguridad Social. En ella se vuelve a reiterar que tal algoritmo es información pública y, además, relevante, en tanto que sirve para explicar cómo se adopta una decisión administrativa; pero se añade una afirmación que me parece fundamental: “Mientras no se instauren otros mecanismos que permitan alcanzar los fines señalados con garantías equivalentes –como podrían ser, por ejemplo, auditorías independientes u órganos de supervisión– el único recurso eficaz a tales efectos es el acceso al algoritmo propiamente dicho, a su código, para su fiscalización tanto por quienes se puedan sentir perjudicados por sus resultados como por la ciudadanía en general en aras de la observancia de los principios

⁸⁰ VESTRI, G. (2022): “El acceso a la información algorítmica a partir del caso bono social vs. fundación ciudadana Civio”, en *Revista General de Derecho Administrativo*, n.º. 61.

⁸¹ En esta misma línea se había pronunciado ya la Comisión de Garantías del Derecho de Acceso de Cataluña en sus Resoluciones (acumuladas) 123 y 124 de 2016, de 21 de septiembre, relativas al acceso a un algoritmo que determinaba la composición de los tribunales de corrección de pruebas de la PAU. En ellas se establece la obligación de facilitar al interesado el algoritmo en lenguaje matemático (si se posee) o, cuando menos, en lenguaje natural, a fin de permitirle conocer la lógica de funcionamiento del sistema.

éticos y de justicia”. Esta doctrina ha sido reiterada en otras reclamaciones posteriores.

Este planteamiento no ha sido seguido por la primera sentencia dictada en relación con esta cuestión en nuestro país –precisamente a propósito del caso BOSCO anteriormente mencionado–. La Sentencia del Juzgado Central de lo Contencioso-administrativo núm. 8, de 30 de diciembre de 2021 (Procedimiento 18/2019), parte, efectivamente, de una consideración diferente. Más allá de que confirma la invocabilidad del límite de la propiedad intelectual y en lo que interesa a estas páginas, se fundamenta en la premisa de que “hay que considerar que la denegación del acceso al código fuente de la aplicación informática no supone una vulneración del principio de legalidad, pues siempre, en último extremo, se podrá comprobar si el solicitante cumple los requisitos para que se le concede el bono social”. Junto con ello, entiende que la aplicación informática es una mera herramienta al servicio del órgano administrativo, manifestando que “no puede considerarse que el acto administrativo se dicte por una aplicación informática, sino por un órgano administrativo, y en caso de que el destinatario de dicho acto esté disconforme con el mismo, podrá impugnarlo en vía administrativa, y en vía judicial. Por tanto, la legalidad del acto administrativo no está justificada por la aplicación informática que instrumentalmente se utiliza en una fase del correspondiente procedimiento administrativo, sino por la normativa que regula la materia” (F. J. 3º).

En cualquier caso, para determinar si procede o no dar acceso al código fuente, es importante, en mi opinión, profundizar en la relación intrínseca existente entre transparencia externa y transparencia interna. Efectivamente, la motivación puede verse afectada por los límites a la publicidad activa sobre la existencia y uso de sistemas algorítmicos. Una amplia exigencia de difundir la información relativa al diseño del sistema, su finalidad de uso, su lógica de funcionamiento, los datos de los que se nutre, el margen de error testado y el nivel de precisión verificado o los resultados de las auditorías realizadas contribuye a la motivación de la decisión adoptada sobre la base del mismo por la Administración; por el contrario, aun cuando se realicen ingentes esfuerzos por motivar en el caso concreto la decisión, difícilmente en ausencia de la información anteriormente mencionada podrá aceptarse la misma como suficiente cuando se trate de sistemas

autónomos. No bastan las razones; resulta imprescindible saber que el sistema funciona de forma correcta.

No podemos olvidar que la motivación no sólo tiene consecuencias a nivel individual. Aunque, en principio, está referida al concreto procedimiento y a la específica decisión adoptada en el seno del mismo, en casos en los que este tipo de decisiones se adoptan en masa partiendo de un sistema de información programado, afectan a la colectividad⁸². Por ello, junto con las garantías para los interesados previstas en el sistema, deben contemplarse una serie de garantías de tutela de la colectividad en su conjunto⁸³; dicho sencillamente, el potencial impacto que los sistemas algorítmicos pueden tener en la práctica hace esencial ofrecer la capacidad de verificar su funcionamiento en aplicación de los principios de transparencia y participación⁸⁴.

Las mismas consideraciones pueden formularse desde la perspectiva del control en vía administrativa y en vía contencioso-administrativa: aun cuando resulte imposible “entrar en la mente” del decisor, en el caso de los sistemas algorítmicos sí resultará fundamental poder conocer cómo razona esa “mente programada”, más aún en los casos en los que opera con cierto nivel de autonomía.

En definitiva, transparencia y motivación son premisas, requisitos instrumentales para el control del cumplimiento del principio de legalidad⁸⁵. Proyectar las exigencias que de una y otra se derivan sobre el

⁸² G. COGLIANESE; D. LEHR afirman por ello que en la mayor parte de los casos será suficiente con mostrar que el sistema ha sido diseñado y opera para el cumplimiento de una finalidad legalmente prevista, que funciona correctamente y permite alcanzar los objetivos programados, y que se usa en el caso concreto para lograr esos objetivos, COGLIANESE, G. y LEHR, D. (2019): “Transparency and algorithmic governance”..., *op. cit.*, p. 40 y 47.

⁸³ PINOTTI, G. (2021): “Amministrazione digitale algoritmica e garanzie procedurali”..., *op. cit.*, p. 85.

⁸⁴ CITRON, D.K. (2008): “Open Code Governance”..., *op. cit.*, p. 357.

⁸⁵ En este sentido, A. SORIANO considera que la falta de transparencia de los sistemas algorítmicos dificulta el control general de legalidad de los programas empleados en la toma de decisiones automatizadas y limita la posibilidad de que los interesados impugnen los resultados, SORIANO, A. (2021): “Decisiones automatizadas: problemas y soluciones jurídicas. Más allá de la protección de datos”, en *Revista de Derecho Público: Teoría y Método*, vol. 3, p. 94.

momento del diseño de los sistemas algorítmicos es una de las modificaciones que deben realizarse.

Pero van más allá. No se trata de cuestiones puramente formales, sino que pueden tener derivaciones sustantivas⁸⁶ que afecten al fondo de la actuación administrativa. Como consecuencia de ello, los eventuales incumplimientos de las obligaciones en las que se manifiesta una y otra podrían derivar en nulidad de pleno derecho o, en su caso, en anulabilidad.

Finalmente, transparencia y motivación guardan relación directa, en este caso, con un doble elemento adicional de carácter tecnológico: la explicabilidad y la interpretabilidad.

De un lado, la explicabilidad se asocia con la lógica interna y mecanismos de operatividad de los sistemas algorítmicos⁸⁷; de otro, interpretabilidad significa, en relación con esta cuestión, capacidad de explicar o representar en términos comprensibles por un humano el funcionamiento de un sistema algorítmico y constituye una exigencia ineludible en aquellos supuestos en los que pueden generarse consecuencias desfavorables para los interesados ante eventuales resultados negativos o no esperados –no, por tanto, en aquellos casos de los que no se deriva consecuencia alguna o el sistema está lo suficientemente testado y evaluado como para confiar completamente en él⁸⁸–. Desde esta lógica, un modelo interpretable no siempre se corresponde con un modelo que los humanos son capaces de entender en su lógica interna.

Esta cuestión afecta directamente a las exigencias formales predicables respecto de los sistemas algorítmicos desde la perspectiva de la transparencia y su ulterior control. La interpretabilidad no siempre implica explicabilidad y a la inversa⁸⁹.

⁸⁶ CARLONI, E. (2020): “I principi della legalità algoritmica. Le decisioni automatizzate di fronte al giudice amministrativo”, *Diritto Amministrativo*, n° 2, p. 293.

⁸⁷ LINARDATOS, P.; PAPASTEFANOPOULOS, V.; y KOTSIANTIS, S. (2021): “Expainable AI: a review of machine learning interpretability methods”, *Entropy*, n° 23, pp. 2-3.

⁸⁸ DOSHI-VELAZ, F. y KIM, B. (2017): “Towards a rigorous science of interpretable machine learning”, disponible en <https://arxiv.org/abs/1702.08608> (Última visita: 30/12/2022).

⁸⁹ LINARDATOS, P.; PAPASTEFANOPOULOS, V.; y KOTSIANTIS, S. (2021): “Expainable AI...”, *op. cit.*, p. 3.

Realizar la distinción es importante porque permite ilustrar una última idea que difiere de la línea defendida por la mayor parte de la doctrina: no necesariamente ha de equipararse transparencia y motivación con apertura total del sistema algorítmico. Aquellas apuntan hacia la legitimidad del sistema, la rendición de cuentas y el control de la decisión adoptada; y para cumplir tales fines no siempre resultará preciso dar a conocer el interior del sistema, el código fuente⁹⁰. Lo relevante en términos jurídicos es difundir la existencia del sistema, los datos de los que se nutre, la información relativa a su operatividad técnica, las razones por las que se ha adoptado la decisión. Importa la transparencia formal y el razonamiento legal, no tanto los procesos técnicos internos seguidos para llegar al mismo –que, por otra parte, no siempre permitirán entender cómo funciona un sistema algorítmico–. En definitiva, es precisa una explicación en lenguaje natural sobre la lógica del funcionamiento, pero no necesariamente y en todo caso la lógica matemática o informática; lo relevante jurídicamente no es dar a conocer sistemáticamente la forma en la que opera el sistema desde el punto de vista tecnológico, sino la interacción humana con el *output* del algoritmo y los criterios de diseño de los *inputs* para mantener la exigencia de inteligibilidad de la decisión⁹¹.

Todo ello debería quedar integrado en una norma de Derecho positivo en aras del principio de seguridad jurídica, del principio democrático y del derecho de defensa.

5.- A MODO CONCLUSIÓN: ALGUNAS GARANTÍAS NECESARIAS

El uso de la IA en el ámbito de la Administración no es una imposición que venga predeterminada, sino que existe un amplio margen de decisión acerca de qué problemas desean resolverse, qué tipo de solución es más adecuada para ello y cómo implantarla en el caso

⁹⁰ PALMER OLSEN, H.; LIVINGSTON SLOSSER, J.: y TROELS HILDEBRANDT, T. (2019): “What's in the Box? The Legal Requirement of Explainability in Computationally Aided Decision-Making in Public Administration, *iCourts Working Paper Series* No. 162, p. 224.

⁹¹ *Ibidem*, p. 227.

concreto. Ciertamente, renunciar de forma completa a la IA no ha de considerarse una opción –sería tanto como renunciar a satisfacer el principio de eficacia, en su manifestación de deber de incorporar la mejor tecnología disponible, al desaprovechar el potencial que presenta–, pero tampoco ha de serlo aplicarla acríticamente como si no existieran múltiples vías, posibilidades y opciones para ello.

La clave de partida para identificar el espacio de la IA en el Derecho Administrativo debe ser siempre la misma: valerse de las innovaciones tecnológicas para resolver problemas jurídicos. O, dicho de otro modo, no se trata sencillamente de asegurar el cumplimiento de las exigencias derivadas del Estado de Derecho y el imperio de la Ley en el uso de sistemas algorítmicos, sino también de garantizar que la forma en que éstos se diseñan y operan integra sus principios nucleares⁹².

Si bien es cierto que todavía se necesita un conocimiento más profundo del poder de la IA y de sus implicaciones jurídicas, resulta posible afirmar que en el Derecho Administrativo hay un espacio para ella y, más concretamente, para los algoritmos que aprenden de manera autónoma. En un contexto de aumento exponencial de los datos, necesitamos herramientas para gestionar toda esta información a fin de mejorar el proceso de toma de decisiones.

Como punto de partida, debemos aprovechar todas las herramientas para optimizar los criterios de decisión. Partiendo de esta lógica, puede afirmarse que el principio de buena administración y, como concreción del mismo, el deber de actuar y resolver con la debida diligencia a la hora de ponderar todos los hechos, intereses y derechos relevantes para la toma de decisiones exige incorporar el uso de la IA. Efectivamente, la inteligencia artificial es un instrumento útil, que puede contribuir eficazmente a mejorar la acción administrativa. Pero, por el momento, es solo eso, una herramienta que nos permite transformar la información en patrones que nos ayudan a tomar decisiones. No creo que la plena comprensión del razonamiento de los algoritmos por parte de los operadores humanos –y, en particular, de los interesados– sea una cuestión dirimente. La opacidad algorítmica se puede transformar en transparencia algorítmica

⁹² HILDEBRANDT, M. (2018): “Algorithmic regulation and the rule of law”, *Philosophical Transactions*, nº 376, p. 9.

a través de la exigencia del cumplimiento de una serie de condiciones idóneas en el diseño⁹³. Es necesario, sin embargo, identificar un conjunto de garantías esenciales y criterios de legalidad que permitan mantener un grado mínimo de autonomía de la máquina y, al mismo tiempo, neutralizar sesgos⁹⁴, impedir desviaciones, detectar vulnerabilidades o corregir errores. También establecer límites claros, que, en todo caso, pasarán por excluir la aplicación de la IA allí donde la tarea que ha de llevarse a cabo sea propia y exclusiva de un ser humano. En definitiva, para mí la clave está en la idea del complemento y colaboración humano-máquina: allí donde la mente y la capacidad humana no puede llegar, resulta conveniente hacer uso de la inteligencia artificial⁹⁵, siempre bajo control humano y respetando los derechos y garantías de los afectados por la actuación administrativa; especialmente –y por ahí debería empezarse– donde la capacidad humana no aporta un valor añadido por aplicarse a tareas sencillas o rutinarias.

Por ello es fundamental tener presentes los riesgos reales. Aunque no es este el espacio adecuado para tratar en profundidad la cuestión relativa a los sesgos –en sentido negativo–, puede señalarse que un sesgo algorítmico se produce cuando un sistema, como consecuencia de deficiencias en los datos de entrenamiento, en la metodología seguida o en el diseño del modelo, genera resultados diferentes en función del colectivo al que pertenecen los sujetos, perjudicándolos

⁹³ COGLIANESE, C. y LEHR, D. (2016): “Regulating by Robot: Administrative Decision Making in the Machine-Learning Era”, en *Research Paper*, nº 17-8, *Institute for Law and Economics.*, p. 7.

⁹⁴ CAMPOS ACUÑA, C. (2019): “Inteligencia Artificial e innovación en la Administración Pública: (in)necesarias regulaciones para la garantía del servicio público”, *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, nº 3, p. 87, disponible en https://www.ivap.euskadi.eus/contenidos/informacion/rvgrp_ultimo_numero/es_def/Campos%2074_91.pdf (Última visita: 30/12/2022).

⁹⁵ Para justificar esta afirmación basta con tener presente un dato que ofrece CARDON: si se digitalizaran todas las comunicaciones y escritos generados desde los albores de la humanidad hasta 2003 con el fin de guardarlos en una memoria serían necesarios 5 millardos de gigabytes; en la actualidad ese volumen de datos se genera en dos días; CARDON, D. (2018): *Con qué sueñan los algoritmos...*, *op. cit.*, p. 18.

precisamente por su pertenencia al mismo⁹⁶. Pero no puede perderse de vista, en este sentido, que lo que comúnmente recibe la denominación de sesgo puede ser sencillamente una consecuencia natural y lógica de los datos de los que se sirve el sistema, que simplemente lleva a cabo un razonamiento estadístico con consecuencias probabilísticas. Cuando, objetivamente, los datos apuntan hacia una determinada solución, el sistema no ha incurrido en error ni en sesgo alguno, sino que, sencillamente, evidenciará lo que se deriva de los mismos⁹⁷. En estos casos resultará necesario ponderar debidamente las actuaciones y decisiones que han de adoptarse en el caso concreto para evitar que la categorización y la identificación de patrones perjudique a personas específicas que, aunque reúnen el perfil, no incurren en el caso concreto en el supuesto de hecho que justifica la consecuencia jurídica que exige el ordenamiento.

En cualquier caso, ello debe llevar a una reflexión más profunda en respuesta a una pregunta clave: ¿es viable, con la calidad de los datos actualmente existente, emprender grandes proyectos transformadores sobre la base de IA? ¿Resulta más conveniente centrar esfuerzos en mejorar su calidad o en establecer parámetros para que los datos generados verdaderamente puedan ser de calidad pro futuro? En definitiva, aunque el riesgo de los sesgos siempre estará presente, afrontar el reto que plantea requiere planificar –y regular– la gobernanza de los datos⁹⁸.

⁹⁶ MORENO REBATO, M. (2021): *Inteligencia Artificial (Umbralés éticos, Derecho y Administraciones Públicas)*..., op. cit., p. 53. Coincido con HUERGO LORA, A. (2022): “Gobernar con algoritmos, gobernar los algoritmos”, en *El Cronista del Estado Social y Democrático de Derecho*, n° 100, p. 85.

⁹⁷ M.L. GÓMEZ JIMÉNEZ, sobre esta cuestión, señala que cuando no sea posible liberarse de esos sesgos mediante la debida reprogramación del sistema, los actos adoptados sobre la base del mismo adolecerían de nulidad radical, GÓMEZ JIMÉNEZ, M.L. (2021): *Automatización procedimental y sesgo electrónico*..., op. cit., p. 149. En mi opinión, no necesariamente será así en todos los supuestos.

⁹⁸ Para un análisis muy completo sobre la prevención de la discriminación algorítmica puede verse SORIANO ARNANZ, A. (2021): *Data protection for the prevention of algorithmic discrimination*, Thomson Reuters-Aranzadi, Cizur Menor. Sobre la cuestión del sesgo de género, puede verse E. MENÉNDEZ SEBASTIÁN, E. (2021): “Buena administración, algoritmos y perspectiva de

En términos generales, se necesitan garantías y principios. Y los que tenemos en estos momentos permiten con carácter general dar una respuesta adecuada a las necesidades actualmente existentes, si bien pueden requerir, en relación con extremos concretos, de una interpretación adaptada a la concreta realidad de la IA. Aun siendo cierto que la Administración digital no consiste únicamente en la incorporación de herramientas tecnológicas en el ámbito del Derecho Administrativo, sino que conduce a una revisión de los criterios de aplicación para testar su funcionamiento ante esta nueva realidad, las categorías generales que la doctrina ha creado siguen siendo útiles. El principio de transparencia, el principio de publicidad, el principio de buena administración, el principio de legalidad, el principio de responsabilidad son principios generales que, tal vez con nuevas variantes, pueden seguir utilizándose en el ámbito del Derecho Administrativo. También las categorías relativas a los vicios de legalidad de la actuación administrativa continúan vigentes: errores materiales, irracionalidad de la elección de la solución adoptada por el sistema, exceso de poder en la programación, defecto de motivación, incompetencia, etc.⁹⁹

No obstante lo anterior, debe reconocerse la existencia de una gran diferencia: la Administración decide sobre la base de una inteligencia no humana, es decir, a través del empleo de algoritmos que aprenden de forma autónoma, que poseen la capacidad de gestionar información y toman decisiones basadas en conocimiento que los humanos no pueden generar por sí mismos. Este es el aspecto en el que, en mi opinión, debemos centrar nuestros esfuerzos como juristas y tecnólogos.

Por lo tanto, aunque se mantengan las categorías generales, su aplicación en relación con la actividad administrativa algorítmica deberá ajustarse en algunos casos. En relación con el concreto extremo

género”, en BONORINO, P.R.; VALCÁRCEL, P.; FERNÁNDEZ, R. (Coord.), *Nuevas normalidades: Inteligencia Artificial, Derecho y género*, Aranzadi, Cizur Menor, pp. 35-63.

⁹⁹ Para una visión general sobre la invalidez de los actos administrativos electrónicos, vid. BERNING PRIETO, A.D. (2019): *Validez e invalidez de los actos administrativos en soporte electrónico*, Aranzadi, Cizur Menor. En relación con la actuación administrativa automatizada, MARTÍN DELGADO, I. (2009): “Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada”..., *op. cit.*, pp. 381 y ss.

relativo a la transparencia de la actividad administrativa algorítmica, resulta preciso establecer claramente la obligación de publicar la existencia de sistemas algorítmicos desplegados en el sector público para la toma de decisiones administrativas con el fin de garantizar un adecuado escrutinio público de la actividad administrativa; junto con ello, han de reforzarse las exigencias de motivación mediante su conexión con la publicidad activa, en la línea apuntada en estas páginas, dando a conocer los diferentes extremos relacionados con el funcionamiento de los sistemas algorítmicos, que pueden sintetizarse en los siguientes: responsabilidad sobre el modelo, fines del mismo, datos empleados, tipo de algoritmos en los que se basa, lógica de funcionamiento, medidas de prevención de sesgos, evaluaciones previas y posteriores de impacto, y documentación técnica asociada al modelo¹⁰⁰. Y todo ello, mediante el establecimiento de reglas jurídicas claras y precisas, de carácter jurídico vinculante.

Más en general, vale la pena señalar algunas propuestas para el debate –que sin duda merecerían ser estudiadas más a fondo–, encaminadas a buscar un equilibrio entre Tecnología y Derecho: en definitiva, a “humanizar la maquina”¹⁰¹.

Como premisa de todas ellas y como síntesis de las reflexiones formuladas, pueden identificarse algunos elementos clave de los que ha de partir toda labor de reinterpretación y, en su caso, de construcción de nuevos principios o de creación de nuevas garantías: primacía del ser humano y control humano sobre el sistema algorítmico; transparencia y explicabilidad; aprobación previa del sistema sobre la base de evaluación de riesgos; auditabilidad de su funcionamiento. Y, siempre, posibilidad de impugnación directa del mismo¹⁰².

Efectivamente, el control humano sobre el algoritmo apunta la idea de la evaluación continua de su funcionamiento. La necesidad de aprobación previa de los sistemas algorítmicos se justifica asimismo sobre la base de que resulta preciso identificar los problemas que desean afrontarse en la organización a través del uso de técnicas y

¹⁰⁰ GUTIÉRREZ DAVID, M.E. (2021): “Administraciones inteligentes y acceso al código fuente...”, *op. cit.*, pp. 193 y 194.

¹⁰¹ FROSINI, V. (2019): *Cibernética, Derecho, Internet y Sociedad... op. cit.*, p. 70.

¹⁰² BOIX, A. (2020): “Los algoritmos son reglamentos...”, *op. cit.*, pp. 260 y ss.

herramientas de IA, elegir de entre ellas las que mejor se adapten al mismo y ponerlas en práctica con las suficientes garantías jurídicas y tecnológicas. La posibilidad de impugnar directamente el sistema algorítmico, en tanto que predetermina los actos finales, deberá ser admitida.

Desde una perspectiva más conceptual, la dimensión objetiva de los derechos fundamentales obliga a fijar la atención sobre la organización y el procedimiento como garantías de los mismos. Aunque la dimensión subjetiva de los derechos fundamentales no resulte transformada radicalmente por la tecnología, la generalización de la IA, también en el sector público –y teniendo en cuenta su impacto comunitario–, conduce a la reivindicación de nuevas exigencias en relación con la existencia de garantías orgánicas y de procedimientos de implantación de sistemas algorítmicos¹⁰³. Más en concreto, y a modo de síntesis de lo expuesto en este trabajo, pueden formularse tres propuestas concretas para el debate, que exigen regulación.

Primera.– En primer lugar, resulta conveniente la configuración de un nuevo principio: el “principio de mínima actividad algorítmica autónoma” de carácter decisorio. La funcionalidad y eficacia en nuestro sistema jurídico de los principios generales del Derecho está fuera de toda duda; actúan, en tanto que reflejo de valores sociales, como guía de otras fuentes del Derecho, como regla interpretativa de las mismas y como fuente subsidiaria¹⁰⁴. El riesgo derivado del uso de *big data* y, en particular, del tratamiento de los datos personales, el desconocimiento parcial de las pautas y criterios utilizados por los algoritmos, la incapacidad de entender completamente la motivación de la decisión final de predicción, refuerzan la idea de que la acción administrativa automatizada autónoma debe estar, por el momento, limitada a las acciones materiales o acciones formales que no impliquen margen de discrecionalidad política y, en cualquier caso, con imputación de responsabilidad (material y formal) al titular del órgano que utiliza la máquina. Su principal uso, cuando exista discrecionalidad política,

¹⁰³ Sobre la dimensión objetiva de los derechos fundamentales desde la perspectiva de la IA, PRESNO LINERA, M. (2022): *Derechos fundamentales e inteligencia artificial*, op. cit., pp. 107 y ss.

¹⁰⁴ ORTEGA ÁLVAREZ, L. (2002): “Funcionalidad y eficacia de los principios generales del Derecho”, *Justicia Administrativa*, nº 15, pp. 5 a 22.

será apoyar las decisiones, pero no sustituirlas¹⁰⁵. En mi opinión, con el estado actual de la tecnología, no cabe ejercer potestades discrecionales en sentido estricto de forma completamente automatizada a través de sistemas algorítmicos precisamente por ese margen “político” de actuación: la expresión de la voluntad político-administrativa, que implica actos de dirección, coordinación o similares, ha de llevarse a cabo por humanos¹⁰⁶. Cuestión distinta es que el órgano que tenga encomendada la competencia pueda servirse de sistemas de IA como apoyo para adoptar la decisión.

En cualquier caso, en la actualidad, la ausencia completa de ser humano en los procesos decisionales no será posible¹⁰⁷, porque el *machine*

¹⁰⁵ Para D. MARONGIU considera que en el ámbito del ejercicio de autoridad, el control humano sobre los resultados de la máquina debe ser total y permanente, y concluye que la IA podrá usarse solo en el área de la actividad administrativa de servicio público, MARONGIU, D. (2018): “Inteligencia artificial y administración pública”, en GARCÍA NOVOA, C.: SANTIAGO IGLESIAS, D.: *4ª Revolución Industrial: impacto de la automatización y la inteligencia artificial en la sociedad y la economía digital*, Thomson-Aranzadi, Cizur Menor, pp. 396-397. M. D’ANGELOSANTE, por su parte, precisa que la vaguedad de las condiciones que deben determinarse y evaluarse antes de la toma de decisiones y el carácter discrecional de la decisión son dos obstáculos a la automatización de la medida, D’ANGELOSANTE, M. (2016): “La consistenza del modello dell’amministrazione “invisibile” nell’età della tecnificazione: dalla formazione delle decisioni alla responsabilità per le decisioni”, en CIVITARESSE MATEUCI, S. (Coord): *A 150 anni dall’unificazione amministrativa italiana – La tecnificazione*, Firenze University Press, Florencia, p. 166.

¹⁰⁶ RAGANELLI, B. (2020): “Le decisioni pubbliche al vaglio degli algoritmi”, *Scritti in onore di Eugenio Picozza*. Editoriale scientifica, Napoli, p. 15. Para esta autora, existen tres modelos de diálogo entre inteligencia artificial y decisiones discrecionales: preclusivo, en el que no cabría hacer uso de sistemas algorítmicos por existir una amplia discrecionalidad; cooperativo, que considera la IA como herramienta útil de soporte de la decisión discrecional; y autorregulado, que utiliza la IA para autolimitarse en el ejercicio de potestades discrecionales, pre-determinando su acción futura y vinculándose a sus propias reglas (pp. 18 y ss.).

¹⁰⁷ J. PONCE entiende que el uso de la IA en el ámbito de las facultades discrecionales debe estar prohibido, y propone la necesidad de regulación de una “reserva de humanidad” para garantizar que ciertas decisiones sean adoptadas solo por humanos, PONCE, J. (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad...”, pp. 26-33. A. CERRILLO considera que es necesario reconocer la necesidad de garantizar la supervisión humana del uso de algoritmos para evitar que pueda tener efectos adversos, supervisión que puede concretarse en la necesidad de incorporar en los procesos decisionales un elemento de

learning tiene una naturaleza predictiva y no permite realizar interpretaciones causales¹⁰⁸. En definitiva, los sistemas algorítmicos han de ser empleados allí donde verdaderamente se pueda mejorar la acción de la Administración¹⁰⁹ no sólo desde la perspectiva de la eficacia y la eficiencia, sino también desde la óptica de la satisfacción de los derechos de los ciudadanos. Es la consecuencia lógica de la aplicación del principio de precaución al uso de IA en el sector público¹¹⁰.

En este sentido, resulta fundamental garantizar la calidad de los datos. En tanto que esta premisa no quede verificada, el citado principio resultará aplicable y no procederá emplear sistemas algorítmicos para la adopción de decisiones administrativas¹¹¹.

Así pues, solo cuando pueda determinarse con precisión y claridad el objetivo que se pretende lograr, exista suficiente cantidad de datos

participación humana a través de la revisión o la validación de la decisión adoptada por la máquina, CERRILLO, A. (2020): “El impacto de la inteligencia artificial en las Administraciones Públicas...”, p. 82. En el mismo sentido, MORENO REBATO manifiesta que la utilización de sistemas de IA de alto riesgo requiere de vigilancia humana; MORENO REBATO, M. (2021): *Inteligencia Artificial (Umbral ético, Derecho y Administraciones Públicas)*..., *op. cit.*, p. 53. M. D’ANGELOSANTE, por su parte, imagina un escenario en el cual la máquina proporciona autonomía y el funcionario recupera un papel de decisión solo en caso de disputas sobre la elección hecha por la máquina, D’ANGELOSANTE, M. (2016): *La consistenza del modello dell’amministrazione “invisibile”*, p. cit., p. 157. Igualmente considera que la ausencia de funcionario público podría violar el derecho del administrado a tener un interlocutor personalizado (p. 161).

¹⁰⁸ COGLIANESE, C. y LEHR, D. (2016): “Regulating by Robot...”, *op. cit.*, pp. 23-24.

¹⁰⁹ COGLIANESE, C. y LEHR, D. (2019): “Transparency and algorithmic governance...”, *op. cit.*, pp. 55-56.

¹¹⁰ Sobre las relaciones entre IA y principio de precaución, véase BOIX, A. (2020): “Los algoritmos son reglamentos...”, *op. cit.*, pp. 225 y ss. En el mismo sentido, PRESNO conecta este principio con el mandato de limitar el uso de la informática que la Constitución dirige al Legislador en el art. 18.4, PRESNO LINERA, M. (2022): *Derechos fundamentales e inteligencia artificial*, *op. cit.*, p. 42.

¹¹¹ CARLONI insiste en la idea de que la legalidad de la acción administrativa no queda garantizada únicamente por la cognoscibilidad y la comprensibilidad del algoritmo, su posibilidad de impugnación y control y la validación por parte de un funcionario; será preciso igualmente asegurar que el sistema no incurre en discriminación, para lo cual resulta imprescindible demostrar que los datos son de calidad, CARLONI, E. (2020): “I principi della legalità algoritmica...”, *op. cit.*, p. 289.

para actuar y posean la calidad mínima precisa y se haya testado con éxito el sistema algorítmico, valorando además sus potenciales riesgos para los derechos e intereses de los ciudadanos y dando la posibilidad a éstos de participar en su diseño y puesta en funcionamiento, podrá hacerse uso del mismo¹¹². Junto con ello, no puede olvidarse que ello será posible en ámbitos decisionales muy estructurados, parametrizables, y no donde los conceptos abstractos sean predominantes, puesto que no son fácilmente codificables. De este modo, resultaría conveniente incorporar estas cautelas en una norma con rango de ley, prohibiendo con carácter general el empleo de sistemas algorítmicos para la toma de decisiones administrativas cuando no se garanticen los extremos mencionados.

Segundo.- En segundo lugar, la elaboración de una regulación específica sobre el proceso de adopción de los programas informáticos y la transparencia de su funcionamiento que concrete y refuerce los principios de transparencia, imparcialidad y participación a la hora de configurar el sistema y su proceso de actuación. Efectivamente, aunque estos principios resultan actualmente de aplicación, deben establecerse reglas específicas para hacerlos valer ante la puesta en práctica de sistemas algorítmicos. Por señalar un ejemplo, no puede prescindirse de la lógica de la participación tanto en su diseño¹¹³ como en su aplicación, permitiendo al interesado tratar de influir en la decisión final soportada en un sistema algorítmico; la despersonalización que se deriva de la automatización de procesos decisorios no ha de ser óbice para que los ciudadanos tengan un referente al cual poder dirigirse para plantear dudas o expresar discrepancias, más allá de los medios de impugnación. Asimismo, es necesario e imprescindible garantizar la objetividad de los datos utilizados y la imparcialidad del procedimiento de transformación de esos datos en predicciones. Necesitamos de una nueva regulación de la apertura de los algoritmos¹¹⁴;

¹¹² COGLIANESE, C. y LAI, A. (2022): “Algorithm vs. algorithm”..., *op. cit.*, pp. 1324-1337.

¹¹³ PINOTTI, G. (2021): “Amministrazione digitale algoritmica e garanzie procedurali”..., *op. cit.*, p. 89.

¹¹⁴ K. MILLER afirma que la legitimidad el carácter razonable de las decisiones basadas en algoritmos dependen de la transparencia del sistema de toma de decisiones; MILLER, K. (2016): “The application of Administrative Law Principles to technology-assisted decision-making”, en *AIAL Forum*, 86, p. 31.

en concreto, para poder conocer –en los casos en los que sea necesario y establezca como tal– los valores del código fuente –tanto los valores matemáticos, como los valores sustantivos, esto es, su finalidad–, que se configuran como condiciones de legalidad. En definitiva, la decisión de utilizar algoritmos en el contexto de un procedimiento administrativo o para el desarrollo de una actividad administrativa debe ser pública. Por todo ello resulta imprescindible la opción por la aprobación previa de estos sistemas¹¹⁵.

¹¹⁵ La utilidad práctica de la misma puede apreciarse mejor en los casos concretos. Muy interesante resulta el análisis de la experiencia andaluza de uso de un sistema de automatización robótica de procesos para la tramitación y resolución de miles de ayudas a autónomos llevada a cabo en aplicación del Decreto autonómico 622/2019, de 27 de diciembre, de Administración electrónica, simplificación de procedimientos y racionalización administrativa de la Junta de Andalucía, cuyo art. 40 añade a las garantías del 41 LRJSP la necesidad de aprobación de las actividades que se realicen mediante actuación administrativa automatizada. Para una visión crítica, puede verse BENÍTEZ PALMA, E. (2020): “La transformación digital del control externo del gasto público”, en *Auditoría Pública*, n° 76, pp. 19-30, disponible en <https://asocex.es/wp-content/uploads/2020/11/Revista-Auditoria-Publica-nº-76-pag-19-a-30.pdf> (Última visita: 30/12/2022). VALERO es quien ha sostenido en nuestro sistema con más insistencia la necesidad de aprobación previa de los sistemas, sobre la base de lo previsto en el art. 41 LRSJP, al considerarla condición inexcusable para la defensa de los derechos e intereses legítimos de los ciudadanos. Junto con ello, entiende precisa la creación de órganos colegiados *ad hoc* dotados de autonomía funcional para la resolución de los recursos que se pudieran interponer frente a los actos administrativos basados en IA y frente a la puesta en funcionamiento de los sistemas algorítmicos, VALERO, J. (2019): “Las garantías jurídicas de la inteligencia artificial ...”, *op. cit.*, pp. 87 y 91. En el mismo sentido, como consecuencia de la consideración de los algoritmos como reglamentos, BOIX, A. (2020): “Los algoritmos son reglamentos...”, *op. cit.*, pp. 250 y ss. CARLONI considera, por su parte, que, en el caso de procedimientos totalmente automatizados resulta de aplicación el principio de tipicidad, siendo necesaria la existencia de una norma previa habilitadora mientras que, por el contrario, en los procedimientos semiatuomatizados bastaría con aplicar los principios generales –autonomía y discrecionalidad organizativa–; CARLONI, E. (2020): “I principi della legalità algoritmica...”, *op. cit.*, p. 295. Finalmente, A.G. OROFINO afirmó tempranamente que las reglas de programación empleadas por la Administración deben ser formalmente aprobadas a través de una resolución administrativa, OROFINO, A.G. (2005): “La automazione amministrativa: imputazione e responsabilità”, *Giornale di Diritto Amministrativo*, no 12, pp. 1308-1309.

En definitiva, el énfasis ha de ponerse no sólo en el por qué se ha adoptado esa decisión, sino también en el cómo.

Tercero.- En tercer lugar, la existencia de un organismo o autoridad de supervisión especializado y de carácter independiente que tenga la función de aprobar el sistema algorítmico y visar el concreto modo en el que opera, así como garantizar su correcto funcionamiento durante el ciclo de vida del mismo¹¹⁶. La normativa actualmente existente –art. 41 LRJSP–, al prever para el caso de actuaciones administrativas automatizadas la necesidad de determinar el órgano u órganos competentes para la supervisión y control de calidad y, en su caso, la auditoría del sistema de información y de su código fuente, está indirectamente exigiendo que estas funciones se realicen en la práctica¹¹⁷. Pero la especialización y la independencia son elementos determinantes en el caso de la IA, más aún teniendo en cuenta que, en

Desde la perspectiva de la ciencia política, partidario de la aprobación previa por parte de una comisión multidisciplinar que valore, en particular, si los algoritmos responden a la ética y valores públicos, RAMÍO, C. (2022): *Burocracia inteligente...*, *op. cit.*, pp. 193-194.

- ¹¹⁶ A. CERRILLO ha sostenido la necesidad de impulsar la creación de agencias independientes con funciones de supervisión de los algoritmos utilizados por las Administraciones, que podrían contar con la colaboración de entidades de

auditoría, CERRILLO, A. (2019): “El impacto de la inteligencia artificial en el derecho administrativo ¿nuevos conceptos para nuevas realidades técnicas?”, *Revista General de Derecho Administrativo*, nº 50, p. 27. En esta misma tendencia se muestra MORENO REBATO, M. (2021): *Inteligencia Artificial (Umbral éticos, Derecho y Administraciones Públicas)*..., *op. cit.*, p. 81, quien manifiesta que a esta entidad se le podría otorgar la función de evaluar, a través del código fuente del sistema y de los datos y el resto de la documentación, el cumplimiento de los requisitos técnicos y jurídicos establecidos, así como el impacto en los derechos de los ciudadanos.

- ¹¹⁷ GAMERO CASADO, E (2021): “Compliance (o cumplimiento normativo) de desarrollos de Inteligencia Artificial...”, *op. cit.*, p. 3.

J. PONCE considera las auditorías de algoritmos una alternativa a los supuestos en los que la transparencia total –el acceso al código fuente– no es posible o deseable, PONCE, J. (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad...”, *op. cit.*, p. 46. Por su parte, G. VESTRI se muestra partidario de que la actividad de auditoría y control sea ejercida por empresas externas de forma independiente y confidencial, VESTRI, G. (2021): “La inteligencia artificial ante el desafío de la transparencia algorítmica. Una aproximación desde la perspectiva jurídico-administrativa”, *Revista Aragonesa de Administración Pública*, nº 56, p. 391.

términos generales, las Administraciones de nuestro país carecen de los medios adecuados para supervisar los sistemas que hacen uso de IA¹¹⁸. Constituiría, además, un claro apoyo al control jurisdiccional ordinario efectuado por los Jueces y Tribunales del orden contencioso-administrativo¹¹⁹. El cambio conceptual que implica el uso de sistemas algorítmicos se concreta en la idea de que el control ha de focalizarse en la programación y no solo en la adopción de la decisión¹²⁰.

Efectivamente, como ha podido apreciarse a lo largo de estas páginas, suele distinguirse, con diferentes denominaciones –algoritmos deterministas y no deterministas, condicionales y no condicionales, *code-driven* y *data driven*...– entre aquellos algoritmos que actúan siguiendo fielmente unas pautas de programación previas completamente acotadas y aquellos otros que, por el contrario, aprenden durante su desarrollo sobre la base de los datos que usan. En todo caso, en uno y otro supuesto, desde la perspectiva jurídica resulta fundamental prestar atención a un doble extremo: en primer lugar, a cómo las normas jurídicas a cuya aplicación sirven (ya sea como apoyo a la decisión, ya decidiendo directamente) han sido traducidas a lenguaje informático; en segundo lugar, a los datos utilizados como referencia por el sistema para operar y lograr el objetivo para el cual ha sido diseñado. En coherencia con ello, existiría una doble vía de impugnación: directa, atacando la decisión adoptada por ser contraria al ordenamiento jurídico; e indirecta, por una deficiente interpretación del ordenamiento jurídico a la hora de diseñar el sistema algorítmico

¹¹⁸ BOIX PALOP, A. (2022): “Transparencia en la utilización de inteligencia artificial...” *op. cit.*, p. 104.

¹¹⁹ Como mantiene AUBY, los jueces y tribunales del contencioso-administrativo no necesariamente estarán mejor preparados que los ciudadanos para comprender los procesos tecnológicos usados por la Administración y, por tanto, su función de control de la legalidad ejercida a través de técnicas clásicas puede verse afectada por el uso de IA en la toma de decisiones administrativas. Vid. AUBY J.B. (2018): “Algorithmes et Smart Cities: Données Juridiques”, *Revue Générale du Droit*, n° 29878, disponible en <https://www.revuegeneraledudroit.eu/blog/2018/12/19/algorithmes-et-smart-cities-donnees-juridiques/> (Última visita: 30/12/2022).

¹²⁰ GÓMEZ JIMÉNEZ, M. L. (2021): *Automatización procedimental y sesgo electrónico...*, *op. cit.*, p. 140.

o por deficiencias en los datos de los que se nutre y en la forma en que ha sido entrenado sobre la base de los mismos¹²¹.

La técnica de la certificación –acreditación de que un sistema de IA es transparente, responsable y equitativo¹²²– puede ayudar en el cumplimiento de este objetivo. Junto con ello, la designación de personas dentro de la organización responsables de garantizar el cumplimiento normativo en el diseño y uso de estos sistemas también es necesaria. Asimismo, la existencia de un registro de algoritmos y sistemas de IA utilizados por las Administraciones Públicas puede ser de utilidad¹²³, también como herramienta para fomentar la transparencia frente a los ciudadanos respecto de la existencia de los mismos. Tales garantías, sin embargo, resultarían insuficientes desde la perspectiva de la legitimidad de la acción administrativa sin una conexión directa con los ciudadanos. En consecuencia, ha de insistirse en sumar a los órganos de evaluación y a los sistemas de certificación un mecanismo adecuado de rendición de cuentas¹²⁴.

Dicho sencillamente, desde el momento en el que las decisiones adoptadas por medio de sistemas algorítmicos tienen su explicación en la forma en la que los mismos hayan sido programados y operen

¹²¹ HILDEBRANDT, M. (2018): “Algorithmic regulation and the rule of law”..., *op. cit.*, p. 3.

¹²² Vid. Directrices Éticas para una AI fiable, *op. cit.*, p. 28.

¹²³ Así lo mantienen SORIANO, A. (2021): “Decisiones automatizadas: problemas y soluciones jurídicas...”, *op. cit.*, p. 115; y O. CORTÉS, para quien, además, sería necesario poder realizar inspecciones periódicas para verificar cómo están actuando los algoritmos registrados, CORTÉS, O. (2020): “Algoritmos y algunos retos jurídico-institucionales para su aplicación en la Administración Pública”, *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, nº 18, p. 59. La conexión entre transparencia y registro público de algoritmos es evidente y no necesita de particular justificación. El informe elaborado por la Comisión Europea AI Wacht. *European Landscape on the Use of Artificial Intelligence by the Public Sector* (disponible en <https://publications.jrc.ec.europa.eu/repository/handle/JRC129301>, Última visita: 30/12/2022) evidencia que los Estados miembros de la Unión Europea están lejos de este ideal. En él se analizan un total de 686 casos de uso de IA por parte del sector público, poniendo de manifiesto las dificultades para su detección, a excepción de aquellos supuestos en los que existe un repositorio nacional, como es el caso de Holanda (pp. 65 y 36).

¹²⁴ PINTO FONTANILLO, J.A. (2021): *El Derecho ante los retos de la Inteligencia Artificial. Marco ético y jurídico*, Edisofer, Madrid, p. 91.

en la práctica en su tarea de aplicar la norma, para poder controlar aquellas será necesario poder conocer y evaluar estos extremos¹²⁵.

Junto con todo ello, no puede olvidarse que resulta imprescindible igualmente pensar en una modalidad diferente de selección, organización y gestión de los recursos humanos¹²⁶. La Administración no puede continuar siendo “esclava” del sector privado en todo lo que respecta al uso de las tecnologías. Esto no quiere decir que todo el desarrollo tecnológico en el ámbito administrativo deba ser público –la colaboración del sector privado es imprescindible en este ámbito–, pero sí que debe reforzarse la capacidad de crear algoritmos públicos.

Finalmente, la ética ha de estar absolutamente presente en el debate y la reflexión sobre la regulación del uso de los algoritmos y la programación de los sistemas algorítmicos, pero sin entenderse en ningún caso como sustitutiva del Derecho.

En definitiva, antes de optar por la idea de Administración inteligente como nuevo paradigma del Derecho Administrativo, debemos profundizar en la idea de Administración racional, entendida como la Administración desprovista de irracionalidad mediante el uso de algoritmos y la adecuada gestión de *big data*, que permite tomar decisiones más lógicas y racionales allí donde resulte procedente.

La Administración Pública actúa en ejecución de la Ley sobre la base de la información y la tecnología disponibles; la automatización y, en particular, la IA permiten el acceso a una cantidad mayor de información, así como un mejor procesamiento de los datos y, por ello, presentan una capacidad de generación de conocimiento que supera la humana. El Derecho no puede reducirse a la dimensión única de la

¹²⁵ COBBE, J. (2019): “Administrative Law and the Machines of Government...”, *op. cit.*, p. 8.

¹²⁶ Sobre este tema, con un enfoque amplio, GALINDO GALDÉS, R. (2020): “Automatización, inteligencia artificial y empleados públicos”, en CERRILLO I MARTÍNEZ, A. y PEGUERA POCH, M.: *Retos jurídicos de la inteligencia artificial*, Thomson-Reuters Aranzadi, Cizur Menor, pp. 93 y ss. Desde la perspectiva de la ciencia política, resultan muy interesantes los planteamientos de RAMIÓ, C. (2019): *Inteligencia Artificial y Administración Pública...*, *op. cit.*, *in totum*; y, más recientemente, (2022) *Burocracia inteligente...* *op. cit.*, pp. 68 y ss.

racionalidad¹²⁷, aunque la racionalidad operativa, concebida sobre la organización y el procedimiento administrativo, contribuye a mejorar la actuación de la Administración en términos de eficacia y objetividad¹²⁸. Los algoritmos no son, por sí mismos, fuente de autoridad, sino herramientas al servicio de la Administración. Los seres humanos somos algo más que datos. Una Administración nunca será “inteligente” si no cumple los principios y respeta las garantías. Esta es la clave del equilibrio entre los algoritmos y la acción administrativa. Y, a día de hoy, ese equilibrio pasa por el empleo de los sistemas algorítmicos para extraer información sobre la base de datos y ponerla al servicio de los humanos en los procesos de toma de decisiones; hemos de ser los humanos, por tanto, quienes, en el sector público, llevemos a cabo el paso de la mera computación a la decisión con efectos externos en el mundo real.

Si la historia del Derecho Administrativo es la historia de la lucha entre poder y libertad¹²⁹, aunque con la tecnología –y, en particular, con la IA–, podamos ampliar los límites del conocimiento, ello no equivaldrá necesariamente y en todo caso a ampliar los límites de la libertad¹³⁰. Para lograrlo habrá que revisar categorías, instituciones jurídicas, conceptos y contextos y valorar convenientemente las eventuales tensiones que se plantean entre los nuevos retos derivados del uso de IA en la Administración Pública y los principios y derechos de los ciudadanos. Es lo que se ha pretendido –de forma limitada en cuanto al ámbito de aplicación– en este trabajo.

¹²⁷ FROSINI, V. (2019): *Cibernética, Derecho, Internet y Sociedad*, op. cit., p. 38.

¹²⁸ En este sentido, BARONA VILAR manifiesta que tras los artefactos tecnológicos ha emergido una nueva ideología vital que guarda relación directa con el control de los individuos y que, por tanto, requiere un importante esfuerzo por mantener los valores y los derechos humanos, en riesgo de ser arrebatados o cedidos voluntariamente, BARONA VILLAR, S. (2021): *Algoritmización del Derecho y de la Justicia. De la Inteligencia Artificial a la Smart Justice*, Tirant lo Blanch, Valencia, p. 17.

¹²⁹ E. GARCÍA DE ENTERRÍA, “La lucha contra las inmunidades del poder en el Derecho Administrativo (poderes discrecionales, poderes de Gobierno, poderes normativos)”, *Revista de Administración Pública*, nº 38, 1962, pp. 159-208.

¹³⁰ J.M. LASALLE, *Ciberleviatán...*, op. cit., pp. 78.

BIBLIOGRAFÍA CITADA

- AUBY J.B. (2018): "Algorithmes et Smart Cities: Données Juridiques", *Revue Générale du Droit*, n° 29878, disponible en <https://www.revuegeneraledudroit.eu/blog/2018/12/19/algorithmes-et-smart-cities-donnees-juridiques/> (Última visita: 30/12/2022).
- BALAGUER CALLEJÓN, F. (2022): *La Constitución del algoritmo*, Fundación Manuel Jiménez Abad, Zaragoza.
- BARRIO ANDRÉS, M. (2022), "Inteligencia artificial: origen, concepto, mito y realidad", en *El Cronista del Estado Social y Democrático de Derecho*, n° 100, pp. 14-21.
- BARONA VILLAR, S. (2021): *Algoritmización del Derecho y de la Justicia. De la Inteligencia Artificial a la Smart Justice*, Tirant lo Blanch, Valencia.
- BENÍTEZ PALMA, E. (2020): "La transformación digital del control externo del gasto público", en *Auditoría Pública*, n° 76, pp. 19-30, disponible en <https://asocex.es/wp-content/uploads/2020/11/Revista-Auditoria-Publica-n-76-pag-19-a-30.pdf> (Última visita: 30/12/2022).
- BERNING PRIETO, A.D. (2019): *Validez e invalidez de los actos administrativos en soporte electrónico*, Aranzadi, Cizur Menor.
- BOIX A. (2020): "Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la Administración para la adopción de decisiones", *Revista de Derecho Público: Teoría y Método*, Vol. I, pp. 223-270.
- (2022): "Transparencia en la utilización de inteligencia artificial por parte de la Administración", en *El Cronista del Estado Social y Democrático de Derecho*, n° 100, pp. 90-105.
- CAMPOS ACUÑA, C. (2019): "Inteligencia Artificial e innovación en la Administración Pública: (in)necesarias regulaciones para la garantía del servicio público", *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, n° 3, pp. 74-91.
- CANALS, D. (2019): "El proceso normativo ante el avance tecnológico y de la transformación digital (inteligencia artificial, redes sociales y datos masivos)", en *Revista General de Derecho Administrativo*, n° 50.
- CARLONI, E. (2020): "I principi della legalità algoritmica. Le decisioni automatizzate di fronte al giudice amministrativo", *Diritto Amministrativo*, n° 2-2020, pp. 272-304.
- CARULLO, G. (2021): "Decisione amministrativa e intelligenza artificiale", *Diritto dell'informazione e dell'informatica*, n° 3, 2021, pp. 431-461.
- CERRILLO, A. (2019): "El impacto de la inteligencia artificial en el derecho administrativo ¿ El impacto de la inteligencia artificial en el derecho administrativo?", *Revista General de Derecho Administrativo*, n° 50.

- (2020): “El impacto de la inteligencia artificial en las Administraciones Públicas: estado de la cuestión y una agenda”, en A. CERRILLO I MARTÍNEZ; M. PEGUERA POCH: *Retos jurídicos de la inteligencia artificial*, Thomson Reuters Aranzadi, Cizur Menor, pp. 75-92.
- CITRON, D.K. (2008): “Open Code Governance”, *University of Chicago Legal Forum*, n° 1, p. 366-367. <https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1430&context=uclf>
- COBBE, J. (2019): “Administrative Law and the Machines of Government: Judicial Review of Automated Public-Sector Decision-Making”, *Legal Studies*, vol. 39, y disponible en el link <http://dx.doi.org/10.2139/ssrn.3226913>, (Última visita: 30/12/2022).
- COGLIANESE, C. y LAI, A. (2022): “Algorithm vs. algorithm”, *Duke Law Journal*, n° 71, pp. 1281-1340.
- COGLIANESE, C. y LEHR, D. (2016): “Regulating by Robot: Administrative Decision Making in the Machine-Learning Era”, en *Research Paper*, n° 17-8, *Institute for Law and Economics*.
- COGLIANESE, G. y LEHR, D. (2019): “Transparency and algorithmic governance”, *Administrative Law Review*, n° 71, pp. 9-15
- CORTÉS, O. (2020): “Algoritmos y algunos retos jurídico-institucionales para su aplicación en la Administración Pública”, *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, n° 18, pp. 54-63.
- COTINO, L. (2021): “Hacia la Transparencia 4.0: el uso de la Inteligencia Artificial y *big data* para la lucha contra el fraude y la corrupción y las (muchas) exigencias constitucionales”, en RAMIÓ C. (Coord.), *Repensando la Administración digital y la innovación pública*, Instituto Nacional de Administración Pública, Madrid, 2021, pp. 169-196.
- (2022): “Transparencia y explicabilidad de la inteligencia artificial y “compañía” (comunicación, interpretabilidad, auditabilidad, testabilidad, comprobabilidad, simulabilidad...). Para qué, para quién y cuánta”, en COTINO, L. y CASTELLANOS, J. (Eds.): *Transparencia y explicabilidad de la inteligencia artificial*, Tirant Lo Blanch, Valencia, 2022, pp. 25-70.
- D’ANGELOSANTE, M. (2016): “La consistenza del modello dell’amministrazione “invisibile” nell’età della tecnificazione: dalla formazione delle decisioni alla responsabilità per le decisioni”, en CIVITARESE MATEUCI, S. (Coord): *A 150 anni dall’unificazione amministrativa italiana – La tecnificazione*, Firenze University Press, Florencia.
- DE LA SIERRA, S. (2022): “La matematización de la realidad y del Derecho Público”, en *Ibericonect*, 14 de marzo de 2022, disponible en <https://www.ibericonnect.blog/2022/03/la-matematizacion-de-la-realidad-y-del-derecho-publico/> (Última visita: 30/12/2022).

- DE LA CUEVA, J. (2019): “La importancia del código fuente”, en CAPI-LLA RONCERO, F.S. (Dir.), *Derecho digital: Retos y cuestiones actuales*, Thomson Reuters Aranzadi, Cizur Menor, pp. 109-127.
- DÍAZ GONZÁLEZ, G. M. (2020): “Algoritmos y actuación policial: la policía predictiva”, en HUERGO LORA, A. (Dir.): *La regulación de los algoritmos*, Thomson-Aranzadi, Cizur Menor, pp. 181-234.
- DOSHI-VELAZ, F. y KIM, B. (2017): “Towards a rigorous science of interpretable machine learning”, disponible en <https://arxiv.org/abs/1702.08608> (Última visita: 30/12/2022).
- DUNI, G. (1978): “L'utilizzabilità delle tecniche elettroniche nell'emanazione degli atti e dei procedimenti amministrativi. Spunto per una teoría dell'atto emanato in forma elettronica, *Rivista amministrativa della Repubblica Italiana*, n° 129, pp. 407 y ss.
- GARCÍA DE ENTERRÍA, E. (1962): “La lucha contra las inmunidades del poder en el Derecho Administrativo (poderes discrecionales, poderes de Gobierno, poderes normativos)”, *Revista de Administración Pública*, n° 38, 1962, pp. 159-208.
- GÓMEZ JIMÉNEZ, M.L. (2021): *Automatización procedimental y sesgo electrónico: el procedimiento administrativo electrónico desde la Inteligencia Artificial*, Thomson-Reuters Aranzadi, Cizur Menor.
- FERNÁNDEZ GIL, A. (2021): Introducción a la obra *Inteligencia Artificial (Umbrales éticos, Derecho y Administraciones Públicas)*, de MORENO REBATO, M.; Aranzadi, Cizur Menor.
- FERNÁNDEZ, T.-R. (1971): “Jurisprudencia y computadores”, *Revista de Administración Pública*, n° 64, pp. 327-336.
- FROSINI, V. (1984): “Informática y Administración Pública”, *Revista de Administración Pública*, n° 105, pp. 447-458.
- (2019): *Cibernética, Derecho, Internet y Sociedad*, Ediciones Olejnik, Santiago de Chile.
- FRY, H. (2019): *Hola mundo. Cómo seguir siendo humanos en la era de los algoritmos*, Blackiebooks, Barcelona. pp. 291-293.
- GALETTA, D.U. (2020): “Digitalizzazione e diritto ad una buona amministrazione (il procedimento amministrativo, fra Diritto UE e tecnologie ICT)”, en CAVALLO PERIN, R.; GALETTA, D.U. (ed), *Il Diritto dell'Amministrazione Pubblica digitale*, Giappichelli, Turín, pp. 85-86.
- GALINDO GALDÉS, R. (2020): “Automatización, inteligencia artificial y empleados públicos”, en CERRILLO I MARTÍNEZ, A. y PEGUERA POCH, M.: *Retos jurídicos de la inteligencia artificial*, Thomson-Reuters Aranzadi, Cizur Menor.
- GAMERO CASADO, E. (2021): “Compliance (o cumplimiento normativo) de desarrollos de Inteligencia Artificial para la toma de decisiones administrativas”, en *Diario La Ley*, n° 50, 19 de abril de 2021.

- GOERLICH PESET, J. M. (2021): “Decisiones administrativas automatizadas en materia social: algoritmos en la gestión de la Seguridad Social y en el procedimiento sancionador”, *Labor*, vol. 2, n° 2, pp. 22-42.
- GUTIÉRREZ DAVID, M.E. (2021): “Administraciones inteligentes y acceso al código fuente de los algoritmos públicos. Conjuro de riesgos de cajas negras decisionales”, *Derecom*, n° 30, pp. 143-229.
- HARLOW, C. y RAWLINGS, R. (2019): “Proceduralism and automation: challenges to the values of Administrative Law”, *Law, Society, Economy Working Papers*, n° 3, p. 2.
- HILDEBRANDT, M. (2018): “Algorithmic regulation and the rule of law”, *Philosophical Transactions*, n° 376.
- HOGAN-DORAN, D. (2017): “Computer says “no”: automation, algorithms and artificial intelligence in Government decision-making”, *The Judicial Review: Selected Conference Papers: Journal of the Judicial Commission of New South Wales*, n° 13 (3), pp. 345-382.
- HUERGO LORA, A. (2020): “Una aproximación a los algoritmos desde el Derecho Administrativo”, en HUERGO LORA, A. (Dir.): *La regulación de los algoritmos*, Thomson-Aranzadi, Madrid, pp. 23-87.
- (2021): “Administraciones Públicas e inteligencia artificial: ¿más o menos discrecionalidad?”, *El Cronista del Estado Social y Democrático de Derecho*, n° 96-97, pp. 78-95.
- (2022): “Gobernar con algoritmos, gobernar los algoritmos”, en *El Cronista del Estado Social y Democrático de Derecho*, n° 100, pp. 80-89.
- LASALLE, J.M. (2019): *Ciberleviatán. El colapso de la democracia liberal frente a la revolución digital*, Arpa, Barcelona. pp. 261-264.
- LESSIG, L. (1999): *Code and other laws of cyberspace*, Basic Books, New York.
- LINARDATOS, P.; PAPASTEFANOPOULOS, V.; y KOTSIANTIS, S. (2021): “Explainable AI: a review of machine learning interpretability methods”, *Entropy*, n° 23 (18).
- MARONGIU, D. (2018): “Inteligencia artificial y administración pública”, en GARCÍA NOVOA, C.: SANTIAGO IGLESIAS, D.: *4ª Revolución Industrial: impacto de la automatización y la inteligencia artificial en la sociedad y la economía digital*, Thomson-Aranzadi, Cizur Menor, pp. 396-397.
- MILLER, K. (2016): “The application of Administrative Law Principles to technology-assisted decision-making”, en *AIAL Forum*, 86, pp. 20-34.
- MARTÍN DELGADO, I. (2009): “Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada”, en *Revista de Administración Pública*, n° 189, pp. 353-386.

- MASUCCI, A. (1993): *L'atto amministrativo informatico. Primi lineamenti di una ricostruzione*, Jovene editore, Napoli.
- MEDINA GUERRERO, M. (2022): “El derecho a conocer los algoritmos utilizados en la toma de decisiones. Aproximación desde la perspectiva del derecho fundamental a la protección de datos personales”, en *Teoría y Realidad Constitucional*, n° 49, pp. 141-171.
- MENÉNDEZ SEBASTIÁN, E. (2021): *De la función consultiva clásica a la buena administración. Evolución en el Estado Social y Democrático de Derecho*, Marcial Pons, Madrid, pp. 43-432.
- (2021): “Buena administración, algoritmos y perspectiva de género”, en BONORINO, P.R.; VALCÁRCEL, P.; FERNÁNDEZ, R. (Coord.), *Nuevas normalidades: Inteligencia Artificial, Derecho y género*, Aranzadi, Cizur Menor, pp. 35-63.
- MORENO REBATO, M. (2021): *Inteligencia Artificial (Umbrales éticos, Derecho y Administraciones Públicas)*, Aranzadi, Cizur Menor.
- MUÑOZ MACHADO, S. (2022): “Prólogo”, en *El Cronista del Estado Social y Democrático de Derecho*, n° 100, pp. 5-13.
- OROFINO, A.G. (2005): “La automazione amministrativa: imputazione e responsabilità”, *Giornale di Diritto Amministrativo*, n° 12, pp. 1308-1309.
- (2020): “The Implementation of the Transparency Principle in the Development of Electronic Administration”, *European Review of Digital Administration & Law*, vol. 1, pp. 123-142, disponible en <https://www.erdalreview.eu/free-download/978882553896012.pdf> (Última visita: 30/12/2022).
- ORTEGA ÁLVAREZ, L. (2002): “Funcionalidad y eficacia de los principios generales del Derecho”, *Justicia Administrativa*, n° 15, pp. 5-22.
- OSÉS, G. (2020): “Algoritmos con derechos”, *Diario 16*, 8 de diciembre de 2020, disponible en <https://diario16.com/algoritmos-con-derechos/> (Última visita: 30/12/2022).
- PALMER OLSEN, H.; LIVINGSTON SLOSSER, J., y TROELS HILDEBRANDT, T. (2019): “What's in the Box? The Legal Requirement of Explainability in Computationally Aided Decision-Making in Public Administration”, *iCourts Working Paper Series* No. 162.
- PALMIOTTO ETTORRE, F. (2022): “The Right to Contest Automated Decisions”, *The Digital Constitutionalist*, disponible en <https://digi-con.org/the-right-to-contest-automated-decisions/> (Última visita: 30/12/2022)
- PINOTTI, G. (2021): “Amministrazione digitale algoritmica e garanzie procedurali”, *Labour & Law Issues*, vol. 7, n° 1, pp. 77-95.
- PINTO FONTANILLO, J.A. (2021): *El Derecho ante los retos de la Inteligencia Artificial. Marco ético y jurídico*, Edisofer, Madrid.

- PRESNO LINERA, M. (2022): *Derechos fundamentales e inteligencia artificial*, Marcial Pons, Madrid.
- PONCE SOLÉ, J. (2001): *Deber de buena administración y derecho al procedimiento administrativo debido. Las bases constitucionales del procedimiento administrativo y del ejercicio de la discrecionalidad*, Lex Nova, Valladolid.
- (2018): “La prevención de riesgos de mala administración y corrupción, la inteligencia artificial y el derecho a una buena administración”, *Revista Internacional de Transparencia e Integridad*, nº 6, pp. 1-19.
 - (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad; algoritmos y procedimiento administrativo debido tecnológico”, *Revista General de Derecho Administrativo*, nº 50.
 - (2019): *La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida*, Cuadernos de la Cátedra de Democracia y Derechos Humanos, nº 15, Universidad de Alcalá de Henares-Defensor del Pueblo, Madrid.
 - (2022): “Reserva de humanidad y supervisión humana de la Inteligencia artificial”, en *El Cronista del Estado Social y Democrático de Derecho*, nº 100, pp. 58-67.
- RAGANELLI, B. (2020): “Le decisioni pubbliche al vaglio degli algoritmi”, *Scritti in onore di Eugenio Picozza*. Editoriale scientifica, Napoli.
- RANCHORDÁS, S. (2022): “Empathy in the Digital Administrative State”, en *Duke Law Journal*, nº 71, pp. 1341-1389.
- RAMÍO, C. (2019): *Inteligencia Artificial y Administración Pública: Robots y humanos compartiendo el servicio público*, Catarata, Madrid.
- (2022): *Burocracia in inteligente. Guía para transformar la Administración pública*, Catarata, Madrid.
- RUSSELL, S. y NORVING, P. (2010): *Artificial Intelligence, a modern approach*, 3ª edición, Prentice Hall, New Jersey.
- SOLUM, L.B. (2019): “Artificially Intelligent Law”, en *BioLaw Journal –Rivista di BioDiritto*, nº 1, pp. 53-62.
- SORIANO ARNANZ, A. (2021): “Decisiones automatizadas: problemas y soluciones jurídicas. Más allá de la protección de datos”, en *Revista de Derecho Público: Teoría y Método*, nº.1 3, 2021, pp. 85-127.
- (2021): *Data protection for the prevention of algorithmic discrimination*, Thomson Reuters-Aranzadi, Cizur Menor.
- SURDEN, H. (2019) “Artificial Intelligence and Law: An Overview”, *Georgia State University Law Review*, nº 35, pp. 1306-1307.
- TODOLÍ SIGNES, A. (2020): “Retos legales del uso del *big data* en la selección de sujetos a investigar por la Inspección de Trabajo y de la Seguridad Social”, *Revista Galega de Administración Pública*, nº 59, pp. 313-337.

- TURING, A.M. (1950): “Computing Machinery and Intelligence”, *Mind* New Series, Vol. 59, n°. 236, pp. 433-460.
- URBINA MENDOZA, E.J. (2020): “El Derecho público del algoritmo. Reflexiones sobre la transición de la modernidad jurídica crítico/lineal a la cuántica/fractal”, en *Revista de Derecho Público*, n° 161/162, p.11-39.
- VALERO, J (2019): “Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración”, *Revista Catalana de Dret Public*, n° 58.
- VESTRI, G. (2021): “La inteligencia artificial ante el desafío de la transparencia algorítmica. Una aproximación desde la perspectiva jurídico-administrativa”, *Revista Aragonesa de Administración Pública*, n° 56, pp. 368-398.
- (2022): “El acceso a la información algorítmica a partir del caso bono social vs. fundación ciudadana Civio”, en *Revista General de Derecho Administrativo*, n°. 61.
- VILLAR PALASÍ, J.L. (1978): “Aspectos jurídicos y políticos de la Telemática”, *Revista Española de Derecho Administrativo*, n° 19, pp. 501-514.
- WHALEY, A. (2018): “Regulatory technology: replacing Law with computer code”, *Law, Society, Economy Working Papers*, n° 14, pp. 1-28.
- WU, T. (2019): “Will artificial intelligence eat the law? The rise of hybrid social-ordering systems”, *Columbia Law Review*, Vol. 119, n° 7, pp. 2001-2028.
- YEUNG, K. (2018): “Algorithmic Regulation: a critical Interrogation”, en *Regulation & Governance*, 12, pp. 505–523.

Capítulo V

Seres humanos e inteligencia artificial: discrecionalidad artificial, reserva de humanidad y supervisión humana¹

JULI PONCE SOLÉ²

Catedrático de Derecho Administrativo.

Universidad de Barcelona

“La compasión sin técnica es un caos; y la técnica sin compasión es una amenaza”

Frase de Karl Llewellyn, uno de los veinte juristas académicos más citados del siglo XX en EE. UU. y miembro del movimiento jurídico realista norteamericano, citada por uno de sus discípulos: CRAMTON, R. C. (1987) “Beyond the ordinary religion”, *Journal of Legal Education*, 1987, vol. 37, no 4, p. 510 (la traducción es nuestra).

¹ El presente estudio es resultado del proyecto de investigación PID2020-115774RB-I00, “Personalización de los servicios públicos, sesgos e inteligencia artificial: Hacia la consolidación de los derechos digitales en las administraciones públicas”, financiado por el Ministerio de Ciencia e Innovación español, en el que el autor es coinvestigador principal.

Algunas ideas contenidas en este estudio fueron expuestas inicialmente en las jornadas que tuvieron lugar los días 28 y 29 de junio de 2022 en Sevilla con el título *Inteligencia artificial en el sector público: retos, límites y medios*, impulsadas por la Universidad Pablo Olavide. En la crónica por escrito de la mencionada jornada aparecida en el blog del aludido proyecto de investigación dichas ideas fueron esbozadas: <https://rednmr.wordpress.com/2022/07/14/inteligencia-artificial-en-el-sector-publico-retos-limites-y-medios/> (última visita 28/07/2022) Esas reflexiones iniciales fueron desarrolladas para su publicación en el número 100 de la revista *El Cronista del Estado Social y Democrático de Derecho*, aceptando así una amable invitación al respecto. Sobre la base de todo lo expuesto, la presente contribución ha desarrollado, modificado y ampliado la reflexión, a los específicos efectos de esta publicación.

² ORCID: 0000-0002-1977-5063

Resumen: Este análisis considera la relación entre seres humanos e inteligencia artificial en referencia a dos aspectos: la llamada reserva de humanidad y la supervisión humana del funcionamiento de la inteligencia artificial. Por lo que se refiere a la reserva de humanidad, el estudio analiza el elemento de la empatía como relevante para la decisión jurídica de excluir de ciertas decisiones automatizadas a la inteligencia artificial, que podría actuar sólo como apoyo al decisor humano. Se analizan las regulaciones de la UE y de Estados miembros que realizan dicha reserva, bien sea en general, bien sea en relación con el ejercicio de “discrecionalidad artificial”. En cuanto a la supervisión humana, se analizan también regulaciones de la UE y españolas, que la exigen con determinadas características. El estudio se cierra con las conclusiones y el planteamiento de algunas cuestiones para seguir investigando, como, por ejemplo, determinar si la supervisión humana es realmente efectiva en el control de la inteligencia artificial.

Abstract: This analysis considers the relationship between humans and artificial intelligence with reference to two aspects: the so-called reservoir of humanity and human oversight of the functioning of artificial intelligence or human in the loop. Regarding the reservoir of humanity, the study analyses the element of empathy as relevant to the legal decision to exclude artificial intelligence from certain automated decisions, which could only act in support of the human decision-maker. It analyses EU and Member State regulations that make such a reservation, either in general or in relation to the exercise of ‘artificial discretion’. With regard to human supervision, it also analyses the EU and Spanish regulations that require it with certain characteristics. The study closes with some conclusions and raises some questions for further research, such as, for example, whether human supervision is really effective in the control of artificial intelligence.

Palabras clave: inteligencia artificial, reserva de humanidad, discrecionalidad, supervisión humana.

Keywords: artificial intelligence, reservoir of humanity, discretion, human in the loop.

SUMARIO.–1. INTRODUCCIÓN: OBJETO DE ESTE ANÁLISIS.–2. HUMANOS, EMPATÍA, INTELIGENCIA ARTIFICIAL, DISCRECIONALIDAD.–3. RESERVA DE HUMANIDAD Y “DISCRECIONALIDAD ARTIFICIAL”.– 3.1 Reserva de humanidad general. 3.2 Reserva de humanidad referida al ejercicio de la discrecionalidad.–4. SUPERVISIÓN HUMANA DE LA IA: *THE HUMAN IN THE LOOP*.–5. CONCLUSIONES Y TEMAS PARA CONSIDERAR EN EL FUTURO. –6 BIBLIOGRAFÍA

1.- INTRODUCCIÓN: OBJETO DE ESTE ANÁLISIS

Se cuenta, apócrifamente, que cuando Einstein y Marilyn Monroe se conocieron en 1949 ella le hizo la siguiente pregunta: “¿Qué dice, profesor, deberíamos casarnos y tener un hijo juntos? ¿Se imagina un bebe con mi belleza y su inteligencia?”. Einstein sonrió y le contestó, seriamente: “Desafortunadamente, me temo que el experimento salga a la inversa y terminemos con un hijo con mi belleza y con su inteligencia”.

Este análisis va a versar sobre otra unión, en este caso la de las capacidades humanas con la Inteligencia Artificial, con la esperanza de que de dicha mezcla se pueda extraer lo mejor, no lo peor, de cada bando y de que el Derecho ayude a ese parto feliz.

Este estudio, pues, versa sobre el papel del factor humano en relación con la inteligencia artificial—tanto respecto, por ejemplo, a la traslación de los sesgos cognitivos humanos a los algoritmos, como respecto al diseño humano de la gobernanza de la inteligencia artificial, entre otras cuestiones.

Se trata de un tema relevante. Lo es porque en esa relación surgen diversos elementos cruciales para la regulación de la inteligencia artificial, por un lado. Pero también lo es, en nuestra opinión, porque la investigación sobre las relaciones entre los seres humanos y la inteligencia artificial nos lleva, inevitablemente, a reflexionar sobre qué significa ser humano y qué nos caracteriza como tales, en general y en relación específica con nuestro papel en el funcionamiento de las Administraciones Públicas, y cuál debe ser nuestro papel como humanos respecto de la IA.

En estas reflexiones que siguen, de todos los temas posibles en esa *interfaz* seres humanos-inteligencia artificial voy a seleccionar uno en concreto: el papel de los humanos respecto a la limitación y supervisión del funcionamiento de la IA en la práctica, recuperando y ampliando alguna reflexión previa ya efectuada.

Efectivamente, en estudios anteriores hemos tenido ocasión de destacar como creíamos que era necesario limitar, en principio, el uso de la IA en relación con el ejercicio de potestades administrativas discrecionales (esto es la existencia de lo que podríamos denominar “discrecionalidad artificial”), sobre la base de que ésta carece de la empatía necesaria para adoptar decisiones que puedan afectar a los ciudadanos de manera adecuada³. La cuestión de si existen ámbitos en los que sería mejor reservar las decisiones a humanos y excluir a la IA de su adopción no es nueva, aunque ha sido poco estudiada en general y menos aún en el ámbito jurídico.

³ PONCE SOLÉ, J. (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico”, *Revista General de Derecho Administrativo*, núm. 1, enero.

En ese sentido, es destacable, desde una perspectiva general no jurídica, el análisis de Weizenbaum, científico y programador del MIT (*Massachusetts Institute of Technology*), de origen alemán, que, como es sabido, diseñó en 1966 el primer *chatbot* de la historia, a quien bautizó como Eliza, en honor del personaje de la obra *Pigmalión* de George Bernard Shaw. Eliza fue diseñada y programada para actuar como un terapeuta rogeriano, entrando en comunicación con las personas mediante preguntas y respuestas, que se realizaban mediante un teclado. Weizenbaum descubrió que las personas, incluidas su secretaria, pasaban a confiar en el programa, al que humanizaban y en el que confiaban, dando lugar al llamado *efecto Eliza*, del que, por cierto, en 2022 hemos tenido un nuevo ejemplo con la IA LaMDA desarrollada por Google, un sistema mucho más sofisticado⁴.

En fecha tan temprana como 1976 Weizenbaum publicó un relevante libro, traducido al español, en el que alertaba sobre el peligro de atribuir tareas que exigían auténtica empatía humana a sistemas algorítmicos, aludiendo entre ellas a algunas tales como la de juez u oficial de policía⁵.

Retomaremos aquí esta cuestión de la reserva a humanos de ciertas decisiones y añadiremos además otra: la de la supervisión, control o vigilancia humana del funcionamiento del sistema, lo que en la literatura específica se conoce desde hace tiempo como *human in the loop*, o el humano en el sistema, en aquellos supuestos en que se acepte el uso de IA.

Ambas intervenciones humanas creemos que son distintas. En la primera, son los humanos, a través de los mecanismos propios del Estado Democrático y de Derecho los que deciden, en un espacio temporal y territorial determinado, prohibir el uso de la IA en ciertos ám-

⁴ LaMDA es el acrónimo en inglés de Modelo de Lenguaje para Aplicaciones de Diálogo. En 2022 un ingeniero de Google, Blake Lemoine, padeció el efecto Eliza, al afirmar que LaMDA era un ser consciente y sintiente, al que incluso proveyó de ayuda legal mediante un abogado. Puede consultarse “The Google engineer who thinks the company’s AI has come to life”, *The Washington Post*, 11 de junio de 2022.

⁵ WEIZENBAUM, J. (1976): *Computer power and human reason: From judgment to calculation*, W. H. Freeman & Co, traducido al español un año más tarde con el título *La frontera entre el ordenador y la mente*, Ed. Pirámide, S.A. Madrid. 1977.

bitos, por diversas razones, como veremos. En cambio, en el segundo caso, aceptado dicho uso, el papel humano ahora será el de garantizar el correcto funcionamiento del sistema para evitar, en lo que aquí nos interesa, la violación de derechos o intereses legítimos de los ciudadanos afectados por la IA que use una Administración.

En ambos casos, el Derecho, el Derecho público, el Derecho Administrativo, tienen que hacer frente a una situación que hasta hace relativamente poco era inédita: que quien toma la decisión sometida al ordenamiento jurídico en vigor no sea un humano, un gestor público, un burócrata, sino un sistema algorítmico, en su caso con uso de IA (esto es sistemas algorítmicos que hacen uso de *Big Data* con aprendizaje automático o profundo, *machine learning*, *deep learning*), una máquina, pues. Decisión de la máquina que puede servir de apoyo para la toma de la decisión final administrativa por parte de un ser humano (resolución, contrato o convenio, reglamento, plan: decisión semiautomatizada) o bien, prescindiendo totalmente de éste, para adoptarla directamente (decisión automatizada).

Frente a esa intervención de las máquinas, el Derecho, como decimos, puede decidir, como veremos, prohibirla en ciertos casos, debido a determinadas razones, o aceptarla, con supervisión humana (como veremos es el caso europeo) o sin ella.

Una pregunta relevante entonces puede ser la siguiente: ¿por qué le interesaría al Derecho prohibir la toma de decisiones automatizadas o imponer una supervisión humana respecto al uso de la IA?

2.- HUMANOS, EMPATÍA, INTELIGENCIA ARTIFICIAL, DISCRECIONALIDAD

Las ventajas de la IA parecen indudables: una capacidad de procesamiento mucho mayor que los humanos (cuya memoria es mucho más limitada y se ve afectada por la fatiga), la capacidad de realizar predicciones mediante correlaciones, mayor eficacia, eficiencia y, por tanto, buena administración⁶, la posibilidad de evitar los sesgos cognitivos de

⁶ COGLIANESE, C. y LAI, A. (2022), “Algorithm vs. Algorithm”, *Faculty Scholarship at Penn Law*. 2795.

los humanos y de acabar con el ruido, esto es la existencia de decisiones distintas no queridas en condiciones similares⁷, y la capacidad de generar nuevas ocupaciones laborales relacionadas.

Ahora bien, en todo caso no hay que olvidar los problemas que también puede generar la IA: desde la substitución de humanos en tareas privada y públicas con el consiguiente desempleo⁸, pasando por el hecho de que las correlaciones pueden dar lugar a errores y que dichas correlaciones tienen una tendencia *conservadora*, al primar el *status quo*, puesto que se extraen predicciones de lo que ocurrió en el pasado, pero el pasado, y las personas, pueden cambiar. Los mencionados errores pueden existir también en la programación (*bugs*), que pueden tener un impacto a gran escala, afectando a muchas más personas que una decisión adoptada por un humano.

Por otro lado, no es cierto que la IA esté libre de sesgos (pues sesgos existen en los datos manejados para entrenar a la IA y para que esta decida o estadísticos). Ni siquiera de los sesgos cognitivos humanos, de los que es *infectada* por los programadores humanos⁹, que sufren de sesgos como el de *disponibilidad*, respecto al uso de ciertos datos empleados para entrenar a la máquina, o de *confirmación*, por ejemplo, al seleccionar dichos datos.

Además, la IA produce un sesgo cognitivo humano específico y bien estudiado: el *sesgo de automatización*, por el que, en general, los humanos tendemos a confiar en lo que nos dicen los sistemas empleados y rara vez lo cuestionamos, lo que puede dar lugar a supuestos extremos como las llamadas *muertes por GPS*: personas pérdidas en el desierto que mueren por seguir a rajatabla las equivocadas indicaciones del navegador¹⁰.

⁷ KAHNEMAN, D., SIBONY, O., SUNSTEIN, C. (2021): *Ruido: Un fallo en el juicio humano*, Debate.

⁸ PASTOR, A. y NOGALES, P. (2019): “El futuro del trabajo en la administración pública ¿estamos preparados?” *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, n.º. extra 3, págs. 34-51.

⁹ PONCE SOLÉ, J. (2022), *Acicates* (Nudges), *bueno gobierno y buena administración. Aportaciones de las ciencias conductuales*, nudging y sectores público y privado, Marcial Pons, Madrid.

¹⁰ Véase al respecto, y en general, el interesante libro de BRIDLE, J. (2020): *La nueva edad oscura: la tecnología y el fin del futuro*, Debate.

Además, como es sabido, los sistemas de aprendizaje automático y profundo presentan el problema de las llamadas *cajas negras*. La complejidad de estos sistemas de toma de decisiones hace difícil explicar los pasos seguidos hasta la toma de la decisión final, lo que plantea cuestiones jurídicamente relevantes, relativas al control judicial del seguimiento del debido proceso, en términos de interdicción de la arbitrariedad (art. 9.3 CE), es decir la prohibición de decisiones irracionales o no motivadas, y también de buena administración, esto es la exigencia de decisiones fruto del respeto de la obligación jurídica de debido cuidado o debida diligencia en la toma de la decisión final, que debe haber considerado todos los factores relevantes y descartado los irrelevantes.

En ocasiones se alega que la IA no presenta una caja negra mayor que el cerebro de un humano que toma decisiones, por ejemplo, un gestor público, desde un punto de vista jurídico. Pero esto no es realmente así: la jurisprudencia a nivel internacional, europeo y español ha hecho un enorme esfuerzo en las últimas décadas para desvelar la *caja negra humana*, exigiendo el cumplimiento de las obligaciones jurídicas aplicables al proceso cognitivo humano de recopilación de información y ponderación de ésta, mediante el análisis atento del expediente administrativo, de la motivación ofrecida para la decisión final y de la motivación de ésta. Debe notarse que se trata de un avance notable, que no puede pasar desapercibido, fruto del *deber jurídico de debida diligencia o debido cuidado asociado al derecho a una buena administración* reconocido en el art. 41 de la Carta de Derechos Fundamentales de la Unión Europea, en la CE¹¹, de acuerdo con la jurisprudencia del Tribunal Supremo, en diversos Estatutos de Autonomía y en la legislación de buen gobierno, por ejemplo¹².

Está aún por ver si estas mismas técnicas podrán dar iguales frutos en relación con el proceso de toma de decisiones de la IA y sus cajas negras.

¹¹ HOFMANN, HCH (2020): “The Duty of Care in EU Public Law—A Principle Between Discretion and Proportionality” *Review of European Administrative Law*, vol. 13, no 2, p. 87-112.

¹² PONCE SOLÉ, J. (2019): *La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida*, UAH-Defensor del Pueblo.

Ahora bien, siendo cierto todo ello, la limitación humana del uso de la IA se basa en algo mucho más profundo aún. La doctrina jurídica y normativa que se ocupan del tema mencionan los conceptos de dignidad humana y posible violación de derechos, como en este segundo caso hace el art. 14 de la propuesta de Reglamento de la Unión Europea de Inteligencia Artificial (en adelante, RIA), en proceso avanzado de aprobación, para justificar la “desconfianza” humana en la IA (debido a “los riesgos para la salud, la seguridad o los derechos fundamentales que pueden surgir cuando un sistema de IA de alto riesgo se utiliza conforme a su finalidad prevista o cuando se le da un uso indebido razonablemente previsible, en particular cuando dichos riesgos persisten a pesar de aplicar otros requisitos...”).

Pero entonces, cabe formular una pregunta similar a la anterior, aunque cambiando la perspectiva: ¿por qué un sistema algorítmico en su caso con IA puede afectar a la dignidad de la persona o violar sus derechos de un modo específico en que no lo puede hacer un humano?

Creemos que la respuesta es la siguiente: las máquinas carecen de empatía (aunque puedan imitarla). Por lo tanto, no pueden adoptar buenas decisiones en aquellos casos en que exista un margen de valoración administrativo (cuestión distinta son las potestades puramente regladas) y, por tanto, no pueden respetar el derecho a una buena administración de los ciudadanos.

Efectivamente, las máquinas carecen de empatía, un factor relevante en las relaciones entre seres humanos, que se vincula con su dignidad y su personalidad (art. 10.2 CE). Atienza se ha preguntado recientemente si máquinas con IA pueden ser o llegar a ser agentes morales, a los que habría que reconocer algún grado de dignidad. Su respuesta es negativa, “puesto que no tienen la capacidad de sentir placer y dolor”, aunque “pueden quizás comportarse como si la tuvieran, pero eso es otra cosa”¹³.

Podemos definir la empatía ahora muy sencillamente como un componente cognitivo humano que permite comprender a los demás, incluyendo sus dificultades, vulnerabilidades y problemas, y adaptar nuestro comportamiento de forma correspondiente. El Diccionario de la Lengua Española de la Real Academia Española define a la empatía

¹³ ATIENZA, M. (2022): *Sobre la dignidad humana*, Trotta, Madrid.

como el “sentimiento de identificación con algo o alguien” y como la “capacidad de identificarse con alguien y compartir sus sentimientos”. Asimismo, incluye en una de las entradas de la palabra humanidad la siguiente: “sensibilidad, compasión por las desgracias de otras personas”. A su vez, la palabra humano/a es definida en una de sus entradas como “comprensivo, sensible a los infortunios ajenos”.

Obviamente, no todos los humanos tienen el mismo nivel de empatía, pero casi todos tienen la capacidad no sólo de tenerla sino de entrenarla. Decimos casi todos porque, precisamente, un rasgo que caracteriza a los psicópatas, que serían un 1% de la población, sería la falta de esta (lo que no significa que no la puedan imitar, para manipular a otros seres humanos en su favor). Estamos frente a una enfermedad grave y peligrosa, pese a intentos de configurarla como una ventaja evolutiva (!)¹⁴, en vez de como lo que es: un trastorno grave de la personalidad, denominado por el DSM-5 trastorno de la personalidad antisocial (el DSM-5 es el *Diagnostic and statistical Manual of Mental Disorders*, editado por la *American Psychiatric Association*).

Por otro lado, no cabe confundir la empatía con la simpatía o con la compasión: la primera incluye sentimientos similares a los que siente el otro, pero no sentimientos por lo que siente esa otra persona. La empatía es un primer paso que puede o no conducir a la simpatía y la compasión o bien a un contagio emocional que puede ser de angustia orientado no al otro, sino a uno mismo. En consecuencia, no cabe confundir empatía con una simple emoción que va a conducir en el ámbito jurídico a la adopción de decisiones parciales y no objetivas: bien al contrario, va a ser esa empatía, esa capacidad de considerar perspectivas distintas la que nos llevará a la imparcialidad y a la objetividad ponderando todos los factores.

Pues bien, los sistemas algorítmicos y la inteligencia artificial no tienen empatía. Los medios de comunicación acostumbran a utilizar un lenguaje metafórico, *humanizando* a las máquinas, y así, por ejemplo, han dado noticia de *Norman*, un sistema, que, entrenado con

¹⁴ PULLMAN LE, REFAIE N, LALUMIÈRE ML, KRUPP D. (2021): “Is Psychopathy a Mental Disorder or an Adaptation? Evidence From a Meta-Analysis of the Association Between Psychopathy and Handedness”. *Evolutionary Psychology*, October.

datos específicamente negativos, al cual se le ha denominado el primer sistema de IA psicópata¹⁵.

No somos partidarios de usar ese lenguaje metafórico con las máquinas, pero si seguimos ahora el juego, en realidad los medios de comunicación creemos que erraron completamente: *Norman* no era el primer sistema de IA psicópata, por la sencilla razón de que *todos ellos han sido, son y en principio, luego diremos algo más al respecto, serán psicópatas*. Todos carecen de empatía, no pueden comprender ni ponerse en el lugar de un ser humano, porque las máquinas no enferman, envejecen o mueren; no sufren y por tanto no pueden empatizar con un ser humano, si bien pueden imitar la empatía (como los humanos psicópatas), lo cual si cabe es todavía más escalofriante y puede conducir a manipulaciones de los humanos...

Pasando de estas reflexiones generales al ámbito del Derecho administrativo y de la gestión pública, hay que destacar que, no tanto en España, pero sí en otros países, la literatura ha insistido en la importancia de la empatía, un valor del servicio público y un elemento a considerar en la elaboración y aplicación del Derecho¹⁶, que cristaliza, por ejemplo, en el principio de la equidad jurídica (art. 3 del Código Civil español). De hecho, la empatía está ínsita en el principio y derecho de buena administración ya mencionado, por cuanto, como es sabido, la buena administración exige la debida diligencia y debido cuidado en la toma en consideración de todos los factores relevantes antes de adoptar una decisión administrativa y, sin duda, la situación concreta de la persona que se relaciona jurídicamente con la Administración (por ejemplo, su vulnerabilidad por pobreza, discapacidad, dificultades vitales, errores humanos...) es un factor relevante a tomar en cuenta.

En consecuencia, esa falta de empatía de las máquinas, ese “secreto de la filantropía” humana, al que se refiere Machado en su conocido poema *Retrato*, es la que las hace inadecuadas para adoptar decisiones automatizadas con margen de valoración que afectan a seres

¹⁵ La BBC, por ejemplo, habla de algoritmo psicópata... Véase <https://www.bbc.com/mundo/noticias-44482471> (última visita: 28/07/2022).

¹⁶ RANCHORDAS, S. (2022): “Empathy in the digital administrative state”, *Duke Law Journal, Forthcoming*, *University of Groningen Faculty of Law Research Paper*, 2021, no 13.

humanos¹⁷. Para adoptar buenas decisiones hace falta racionalidad, pero no sólo, como ha destacado el neurocientífico portugués que investiga en los EE. UU. Antonio Damasio, por ejemplo, en su libro *El error de Descartes*¹⁸.

En el ámbito jurídico, como en otros ámbitos profesionales, persiste la creencia, heredada de la Ilustración, de que una buena decisión no debe incorporar ni emociones ni sentimientos, sino que ha de usar, simplemente, la *fría racionalidad*. Sin embargo, los estudios neurocientíficos de las últimas décadas nos demuestran que ese es un concepto burdo y anticuado: somos seres sintientes y emocionales que, por ello y gracias a ello, toman decisiones racionales. Parece como si la naturaleza hubiese construido el aparato de la racionalidad desde y con el cuerpo.

Para terminar estas reflexiones, cabe plantearse si podrían adquirir empatía real en el futuro las máquinas y por tanto llegar a replantearnos algunas de las reflexiones que estamos efectuando aquí. Al respecto, en un interesante artículo en coautoría con Man, aparecido en 2019 en *Nature*¹⁹ y en su último libro *Sentir y Sentir*²⁰, Damasio no cierra la puerta a ello: propone trabajar en la línea de otorgar a las máquinas un sentido de supervivencia, simulando en ellas una propiedad biológica de los humanos, la de la homeostasis, esto es, la capacidad de los organismos de permanecer en unas condiciones aceptables para la vida (por ejemplo, temperatura). Si fuéramos capaces de dotar a las máquinas de un sentido de vulnerabilidad y de un comportamiento auto conservador, mediante unos sensores que les enseñen qué factores ocupan un papel en su propia supervivencia (cables conectados, cantidad adecuada de electricidad...), se sostiene, ello podría conducir a generar sentimientos (inquietud, satisfacción...) y de ahí, quizás, llegar a la empatía.

¹⁷ “Converso con el hombre que siempre va conmigo —quien habla solo espera hablar a Dios un día—; mi soliloquio es plática con ese buen amigo que me enseñó el secreto de la filantropía”

MACHADO, A. (1912): *Mi infancia son recuerdos de un patio de Sevilla* y “Retrato”. *Campos de Castilla*, 1912.

¹⁸ DAMASIO, A. (1994): *El error de Descartes*. Andrés Bello.

¹⁹ MAN, K.; DAMASIO, A. (2019): “Homeostasis and soft robotics in the design of feeling machines”, *Nature Machine Intelligence*, vol. 1, no 10, p. 446-452

²⁰ DAMASIO, A. (2021), *Sentir y Sentir*, Destino.

En todo caso, se trata de una hipótesis. El presente estudio se basa en la situación existente en los momentos de redactar estas líneas, la cual no parece que vaya a cambiar ni a corto ni a medio plazo.

En definitiva, la adopción de buenas decisiones en el ámbito administrativo exige respetar el derecho a una buena administración. Éste implica la obligación jurídica de tomar en cuenta los factores relevantes antes de decidir. Entre ellos se encuentra el poder conocer la situación de las personas que se relacionan con la Administración, entenderla y obrar en consecuencia, aplicando la equidad si fuera preciso, en garantía del procedimiento administrativo debido, de la dignidad de la persona y del libre desarrollo de su personalidad. Sin embargo, nada de todo eso pueden hacer hoy en día los sistemas algorítmicos y la IA.

Esta constatación conduce, entonces, al uso de dos técnicas reguladoras (la reserva de humanidad y la supervisión humana), que debemos entender basadas en el principio de precaución aplicado a la AI²¹, para limitarla y vigilarla en aquellos ámbitos que una sociedad considere especialmente sensibles, dadas las limitaciones intrínsecas de la tecnología ya expuestas.

A continuación, vamos a realizar un análisis de las dos técnicas mencionadas.

3.- RESERVA DE HUMANIDAD Y “DISCRECIONALIDAD ARTIFICIAL”

En cuanto al concepto de reserva de humanidad, se trataría de una decisión humana respecto de ciertos ámbitos, en los que los seres humanos, a través de los órganos legitimados para ello, excluyen la aplicación de la IA porque se considera inadecuado el uso de ésta.

La idea de reserva de humanidad es comparable, en otro nivel, a la idea de reserva del ejercicio de potestades para los funcionarios, prevista actualmente, como es sabido, en el art. 9.2 del Real Decreto

²¹ SOM, C., HILTY, L.M and KÖHLER, A.R. (2009): “The Precautionary Principle as a Framework for a Sustainable Information Society”, *Journal of Business Ethics*, 85:493–505

Legislativo 5/2015, de 30 de octubre, por el que se aprueba el texto refundido de la Ley del Estatuto Básico del Empleado Público. En este contexto, se considera por el legislador que las potestades administrativas sólo deben ser ejercidas por empleados públicos que, debido a su relación estatutaria con la Administración, la cual protege su imparcialidad y objetividad y les ofrece resistencia a presiones políticas o de grupos de presión privados, están en la situación adecuada para el buen servicio a los intereses generales.

Pues bien, en el ámbito de la IA, un razonamiento similar creemos que se encuentra implícitamente tras las normas jurídicas ya vigentes o en proceso de elaboración que requieren que sólo sean humanos los que tomen ciertas decisiones (impidiendo pues, decisiones completamente automatizadas) debido a diversas razones, entre las que se encuentra esa capacidad empática de la que carecen las máquinas.

Estas normas se articulan en torno a dos ideas: o bien la reserva total o bien la reserva limitada sólo a decisiones discrecionales.

3.1.- Reserva de humanidad general

Así, el art. 22 del RGPD configura, de hecho, una reserva de humanidad como principio, que puede ser excepcionado en diversos casos, como es sabido, entre los que se encuentra que una norma específicamente prevea el uso de la IA:

“Artículo 22

Decisiones individuales automatizadas, incluida la elaboración de perfiles

1. Todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar.

2. El apartado 1 no se aplicará si la decisión:

a) es necesaria para la celebración o la ejecución de un contrato entre el interesado y un responsable del tratamiento;

b) está autorizada por el Derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento y que establezca asimismo medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, o

c) se basa en el consentimiento explícito del interesado.

3. En los casos a que se refiere el apartado 2, letras a) y c), el responsable del tratamiento adoptará las medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, como mínimo el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión.

4. Las decisiones a que se refiere el apartado 2 no se basarán en las categorías especiales de datos personales contempladas en el artículo 9, apartado 1, salvo que se aplique el artículo 9, apartado 2, letra a) o g), y se hayan tomado medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado”.

Esa norma jurídica nacional que puede prever el uso de la IA y, por tanto, excepcionar la reserva de humanidad inicialmente establecida por el art. 22, creemos que en el ámbito español debería ser una norma con rango de ley, por aplicación del art. 18.4, el cual establece una reserva de ley constitucional:

“La ley limitará el uso de la informática para garantizar el honor y la intimidad personal y familiar de los ciudadanos y el pleno ejercicio de sus derechos”

Podríamos, decir entonces, parafraseando la conocida expresión acuñada en la Revolución norteamericana: *No automation Without Representation* en el caso español. En consecuencia, el empleo por las AAPP de IA sin una norma con rango de ley que así lo permita y regule su uso sería inconstitucional por violar dicha reserva de ley.

En la misma línea, aunque sin referencia específica a norma con rango de ley, como exige el art. 18 CE, se sitúa el art. XVIII.6.d de la Carta española de derechos digitales (“Que la adopción de decisiones discrecionales quede reservada a personas, salvo que normativamente se prevea la adopción de decisiones automatizadas con garantías adecuadas”).

Cuestión distinta a lo dicho es si se considera que el actual art. 41 (art. 41 Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, LRJSP), con su muy escasa regulación, cubre dicha reserva de ley:

“Artículo 41. Actuación administrativa automatizada.

2. En caso de actuación administrativa automatizada deberá establecerse previamente el órgano u órganos competentes, según los casos, para la definición de las especificaciones, programación, mantenimiento,

supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente. Asimismo, se indicará el órgano que debe ser considerado responsable a efectos de impugnación.”

De nuevo aquí, una analogía con el régimen de los funcionarios públicos puede ayudar a aclarar la cuestión. Como es sabido, la regulación del estatuto funcional está constitucionalmente reservada a la ley en virtud del artículo 103.3 de la Constitución. No obstante, como ha recordado, por ejemplo, la sentencia del Tribunal Constitucional 1/2003 de 16 de enero (reiterando la doctrina establecida en sentencias anteriores), esta reserva de Ley tiene un alcance relativo, pues no impide la colaboración de las normas reglamentarias y, en su caso, de otro tipo de fuentes normativas, aunque ésta por definición deba estar limitada en la ordenación de la materia. De manera parecida a como acontece con otros ámbitos materiales reservados por la Constitución a la Ley, es posible en esta materia una intervención auxiliar o complementaria del reglamento, siempre que estas remisiones legales, nunca en blanco, restrinjan efectivamente el ejercicio de la potestad reglamentaria a un complemento de la regulación legal que sea indispensable por motivos técnicos o para optimizar el cumplimiento de las finalidades propuestas por la Constitución o por la propia Ley, de modo que no se llegue a una total abdicación por parte del legislador de su facultad para establecer reglas limitativas, transfiriendo esa facultad al titular de la potestad reglamentaria, sin fijar ni siquiera cuáles son los fines u objetivos que la reglamentación ha de perseguir.

Pues bien, nos parece que la regulación legal de las decisiones automatizadas administrativas, en su caso con empleo de IA, en la LRJ-SP, no cumple actualmente esas exigencias.

3.2.- Reserva de humanidad referida al ejercicio de la discrecionalidad

Por otro lado, con un enfoque más reducido, afectando sólo a potestades discrecionales, desde otra lógica, diversas leyes nacionales actualmente también prohíben el uso de IA salvo que se trate del ejercicio de potestades regladas.

Así, la legislación alemana (“Un acto administrativo puede dictarse de forma totalmente automática, siempre que así lo permita la

ley y que no exista ni discrecionalidad ni margen de apreciación”²², *Verwaltungsverfahrensgesetz*, VwVfG, § 35a *Vollständig automatisierter Erlass eines Verwaltungsaktes*) o catalana (Ley 26/2010, art. 44.2: “Sólo son susceptibles de actuación administrativa automatizada los actos que puedan adoptarse con una programación basada en criterios y parámetros objetivos”).

La lógica subyacente a este tipo de prohibiciones, creemos, debe hallarse en la ya mencionada ausencia de empatía de las máquinas: el ejercicio de discrecionalidad exige empatía, ponderación de todos los elementos relevantes en la toma de la decisión (obligación de debida diligencia o debido cuidado propio del derecho a una buena administración) y aplicación, en su caso de la equidad. Sólo los humanos pueden decidir de este modo. Cuestión diferente es la aplicación automatizada de potestades regladas sin ningún tipo de margen de apreciación.

Dicho esto, queda pendiente una mayor matización de qué entendemos por discrecionalidad que deba extraerse de la IA, lo que no podemos afrontar ahora en detalle en estas páginas. Baste adelantar que frente a una exclusión sin matices, quizás debiera avanzarse en una precisión mayor, sobre la base de la existencia de distintos tipos de discrecionalidad humana, si bien entendemos que tal extracción debe ir relacionada con la existencia, en definitiva, de un margen de apreciación, de una necesaria elección basada en elementos no jurídicamente predeterminados que incorporan una valoración que debe basarse en la empatía y la sabiduría humanas, por decirlo en los términos ya referidos de Weizenbaum²³.

Está por determinar si podría llegar a ser aceptable lo que ha sido denominado por Young, Bullock y Lecy como “discrecionalidad artificial”, esto es, el ejercicio por máquinas de ciertas funciones administrativas que incorporan lo que ellos denominan una “baja discrecionalidad” (usando el ejemplo de otorgamiento de licencias). Estos

²² La traducción es nuestra.

²³ Sobre esta cuestión, tempranamente, el clásico trabajo de MARTÍN DELGADO, I (2009): “Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada”, *Revista de administración pública*, N° 180, págs. 353-386, distinguiendo entre una discrecionalidad política y otra técnica, y dentro de la segunda, una parametrizable y otra no parametrizable.

autores, en cambio, sugieren el uso de la IA sólo como apoyo (dando lugar, pues a decisiones semiautomatizada) cuando nos movemos en conceptos de discrecionalidad “media” (ej. procesos de selección de personal) o “alta” (ej. respuestas a crisis)²⁴. Si bien su taxonomía puede ser discutible y necesitada de mayores precisiones, tiene el interés de intentar distinguir tipos de discrecionalidad humana y discutir pros y contras de su substitución por “discrecionalidad artificial”

De igual modo, el proyecto de RIA, en su art. 5, toma la decisión de excluir diversos supuestos del uso de la IA, entre ellos, salvo autorización específica, la identificación biométrica. Transcribimos los apartados esenciales de este precepto:

“TÍTULO II

PRÁCTICAS DE INTELIGENCIA ARTIFICIAL PROHIBIDAS

Artículo 5

1. Estarán prohibidas las siguientes prácticas de inteligencia artificial:

a) La introducción en el mercado, la puesta en servicio o la utilización de un sistema de IA que se sirva de técnicas subliminales que trasciendan la conciencia de una persona para alterar de manera sustancial su comportamiento de un modo que provoque o sea probable que provoque perjuicios físicos o psicológicos a esa persona o a otra.

b) La introducción en el mercado, la puesta en servicio o la utilización de un sistema de IA que aproveche alguna de las vulnerabilidades de un grupo específico de personas debido a su edad o discapacidad física o mental para alterar de manera sustancial el comportamiento de una persona que pertenezca a dicho grupo de un modo que provoque o sea probable que provoque perjuicios físicos o psicológicos a esa persona o a otra.

c) La introducción en el mercado, la puesta en servicio o la utilización de sistemas de IA por parte de las autoridades públicas o en su representación con el fin de evaluar o clasificar la fiabilidad de personas físicas durante un período determinado de tiempo atendiendo a su conducta social o a características personales o de su personalidad conocidas o predichas, de forma que la clasificación social resultante provoque una o varias de las situaciones siguientes:

²⁴ YOUNG, M. M.; BULLOCK, J. B.; LECY, J. D (2019): “Artificial discretion as a tool of governance: a framework for understanding the impact of artificial intelligence on public administration”, *Perspectives on Public Management and Governance*, vol. 2, no 4, p. 301-313.

i) un trato perjudicial o desfavorable hacia determinadas personas físicas o colectivos enteros en contextos sociales que no guarden relación con los contextos donde se generaron o recabaron los datos originalmente;

ii) un trato perjudicial o desfavorable hacia determinadas personas físicas o colectivos enteros que es injustificado o desproporcionado con respecto a su comportamiento social o la gravedad de este.

d) El uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de aplicación de la ley, salvo y en la medida en que dicho uso sea estrictamente necesario para alcanzar uno o varios de los objetivos siguientes:

i) la búsqueda selectiva de posibles víctimas concretas de un delito, incluidos menores desaparecidos;

ii) la prevención de una amenaza específica, importante e inminente para la vida o la seguridad física de las personas físicas o de un atentado terrorista;

iii) la detección, la localización, la identificación o el enjuiciamiento de la persona que ha cometido o se sospecha que ha cometido alguno de los delitos mencionados en el artículo 2, apartado 2, de la Decisión Marco 2002/584/JAI del Consejo 62, para el que la normativa en vigor en el Estado miembro implicado imponga una pena o una medida de seguridad privativas de libertad cuya duración máxima sea al menos de tres años, según determine el Derecho de dicho Estado miembro.

2. El uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de aplicación de la ley para conseguir cualquiera de los objetivos mencionados en el apartado 1, letra d), tendrá en cuenta los siguientes aspectos:

a) la naturaleza de la situación que dé lugar al posible uso, y en particular la gravedad, probabilidad y magnitud del perjuicio que se produciría de no utilizarse el sistema;

b) las consecuencias que utilizar el sistema tendría para los derechos y las libertades de las personas implicadas, y en particular la gravedad, probabilidad y magnitud de dichas consecuencias.

Además, el uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de aplicación de la ley para cualquiera de los objetivos mencionados en el apartado 1, letra d), cumplirá salvaguardias y condiciones necesarias y proporcionadas en relación con el uso, en particular en lo que respecta a las limitaciones temporales, geográficas y personales.

Con respecto al apartado 1, letra d), y el apartado 2, cualquier uso concreto de un sistema de identificación biométrica remota «en tiempo real» en un espacio de acceso público con fines de aplicación de la ley

estará supeditado a la concesión de una autorización previa por parte de una autoridad judicial o una autoridad administrativa independiente del Estado miembro donde vaya a utilizarse dicho sistema, que la otorgarán previa solicitud motivada y de conformidad con las normas detalladas del Derecho interno mencionadas en el apartado 4. No obstante, en una situación de urgencia debidamente justificada, se podrá empezar a utilizar el sistema antes de obtener la autorización correspondiente, que podrá solicitarse durante el uso o después de este.

..."

En el ámbito de los consumidores, un proyecto de ley actualmente en discusión en el Congreso español acepta el uso de inteligencia artificial en las relaciones de las empresas con los consumidores, hasta el momento en que éstos ejerzan un derecho a relacionarse con un humano obteniendo atención personalizada, momento en el que la IA debe retirarse y la relación entre empresa y consumidor queda reservada sólo a humanos (Proyecto de ley por la que se regulan los servicios de atención a las personas consumidoras y usuarias que actúan en calidad de clientela).

4.- SUPERVISIÓN HUMANA DE LA IA: *THE HUMAN IN THE LOOP*

Una segunda técnica reguladora para controlar la IA, menos drástica, es permitir su uso, si bien sometido a supervisión o vigilancia humana.

En esta línea se sitúan el ya mencionado art. 22 RGPD (en caso de que se excepcione la reserva de humanidad), el art. 41 LRJSP o el art. 96 Ley General Tributaria²⁵, ambos con una muy parca alusión a dicha supervisión que levanta dudas sobre el cumplimiento del art. 18

²⁵ “Artículo 96. Utilización de tecnologías informáticas y telemáticas.

1. La Administración tributaria promoverá la utilización de las técnicas y medios electrónicos, informáticos y telemáticos necesarios para el desarrollo de su actividad y el ejercicio de sus competencias, con las limitaciones que la Constitución y las leyes establezcan.

2. Cuando sea compatible con los medios técnicos de que disponga la Administración tributaria, los ciudadanos podrán relacionarse con ella para ejercer sus derechos y cumplir con sus obligaciones a través de técnicas y medios electróni-

CE, así como también la Carta española de derechos digitales en su art. XXV.2 y 3 (cuando no se aplique la reserva de humanidad para la discrecionalidad, esto es en potestades regladas o en discrecionales donde exista norma permitiendo el uso de IA):

“XXV Derechos ante la inteligencia artificial

2. En el desarrollo y ciclo de vida de los sistemas de inteligencia artificial:

a) Se deberá garantizar el derecho a la no discriminación cualquiera que fuera su origen, causa o naturaleza, en relación con las decisiones, uso de datos y procesos basados en inteligencia artificial.

b) Se establecerán condiciones de transparencia, auditabilidad, explicabilidad, trazabilidad, supervisión humana y gobernanza. En todo caso, la información facilitada deberá ser accesible y comprensible.

c) Deberán garantizarse la accesibilidad, usabilidad y fiabilidad.

3. Las personas tienen derecho a solicitar una supervisión e intervención humana y a impugnar las decisiones automatizadas tomadas por sistemas de inteligencia artificial que produzcan efectos en su esfera personal y patrimonial”

cos, informáticos o telemáticos con las garantías y requisitos previstos en cada procedimiento.

3. Los procedimientos y actuaciones en los que se utilicen técnicas y medios electrónicos, informáticos y telemáticos garantizarán la identificación de la Administración tributaria actuante y el ejercicio de su competencia. Además, cuando la Administración tributaria actúe de forma automatizada se garantizará la identificación de los órganos competentes para la programación y supervisión del sistema de información y de los órganos competentes para resolver los recursos que puedan interponerse.

4. Los programas y aplicaciones electrónicos, informáticos y telemáticos que vayan a ser utilizados por la Administración tributaria para el ejercicio de sus potestades habrán de ser previamente aprobados por ésta en la forma que se determine reglamentariamente.

5. Los documentos emitidos, cualquiera que sea su soporte, por medios electrónicos, informáticos o telemáticos por la Administración tributaria, o los que ésta emita como copias de originales almacenados por estos mismos medios, así como las imágenes electrónicas de los documentos originales o sus copias, tendrán la misma validez y eficacia que los documentos originales, siempre que quede garantizada su autenticidad, integridad y conservación y, en su caso, la recepción por el interesado, así como el cumplimiento de las garantías y requisitos exigidos por la normativa aplicable.”

Como *lege ferenda*, destaca el ya citado art. 14 del RIA, que se refiere precisamente a esta cuestión, bajo el rótulo “Vigilancia humana” y respecto sólo a los usos de IA calificados como de alto riesgo (los detallados en el Anexo III), no existiendo la misma en otros supuestos. Caso de ser aprobado en su versión actual, este precepto se ocuparía de diversos aspectos de la supervisión humana. A continuación, reproducimos este precepto y lo comentamos brevemente:

“Artículo 14

Vigilancia humana

1. Los sistemas de IA de alto riesgo se diseñarán y desarrollarán de modo que puedan ser vigilados de manera efectiva por personas físicas durante el período que estén en uso, lo que incluye dotarlos de una herramienta de interfaz humano-máquina adecuada, entre otras cosas.

2. El objetivo de la vigilancia humana será prevenir o reducir al mínimo los riesgos para la salud, la seguridad o los derechos fundamentales que pueden surgir cuando un sistema de IA de alto riesgo se utiliza conforme a su finalidad prevista o cuando se le da un uso indebido razonablemente previsible, en particular cuando dichos riesgos persisten a pesar de aplicar otros requisitos establecidos en el presente capítulo.

3. La vigilancia humana se garantizará de una de las siguientes maneras o de ambas:

a) el proveedor definirá las medidas de vigilancia humana y, cuando sea técnicamente viable, las integrará en el sistema de IA de alto riesgo antes de su introducción en el mercado o puesta en servicio;

b) el proveedor definirá las medidas de vigilancia humana, que serán adecuadas para que las lleve a cabo el usuario, antes de la introducción del sistema de IA de alto riesgo en el mercado o de su puesta en servicio.

4. Las medidas mencionadas en el apartado 3 permitirán que las personas a quienes se encomiende la vigilancia humana puedan, en función de las circunstancias:

a) entender por completo las capacidades y limitaciones del sistema de IA de alto riesgo y controlar debidamente su funcionamiento, de modo que puedan detectar indicios de anomalías, problemas de funcionamiento y comportamientos inesperados y ponerles solución lo antes posible;

b) ser conscientes de la posible tendencia a confiar automáticamente o en exceso en la información de salida generada por un sistema de IA de alto riesgo («sesgo de automatización»), en particular con aquellos sistemas que se utilizan para aportar información o recomendaciones con el fin de que personas físicas adopten una decisión;

c)interpretar correctamente la información de salida del sistema de IA de alto riesgo, teniendo en cuenta en particular las características del sistema y las herramientas y los métodos de interpretación disponibles;

d)decidir, en cualquier situación concreta, no utilizar el sistema de IA de alto riesgo o desestimar, invalidar o revertir la información de salida que este genere;

e)intervenir en el funcionamiento del sistema de IA de alto riesgo o interrumpir el sistema accionando un botón específicamente destinado a tal fin o mediante un procedimiento similar.

5.En el caso de los sistemas de IA mencionados en el punto 1, letra a), del anexo III, las medidas que figuran en el apartado 3 garantizarán además que el usuario no actúe ni tome ninguna decisión sobre la base de la identificación generada por el sistema, salvo que un mínimo de dos personas físicas la hayan verificado y confirmado.”

A continuación, vamos a analizar este precepto sobre la base de comprobar qué se dice en él sobre quién supervisa y cuándo y cómo se supervisa la IA:

- Quien supervisa:

- La supervisión debe llevarse a cabo tanto por el proveedor como por el usuario, que de conformidad con el art. 3 son:
 - ““Proveedor”: toda persona física o jurídica, autoridad pública, agencia u organismo de otra índole que desarrolle un sistema de IA o para el que se haya desarrollado un sistema de IA con vistas a introducirlo en el mercado o ponerlo en servicio con su propio nombre o marca comercial, ya sea de manera remunerada o gratuita”;
 - “Usuario”: toda persona física o jurídica, autoridad pública, agencia u organismo de otra índole que utilice un sistema de IA bajo su propia autoridad, salvo cuando su uso se enmarque en una actividad personal de carácter no profesional”.
- El ser humano que vigila a la IA debe tener determinadas capacidades para poder realizar bien su labor, pues debe en las palabras del art. 14:
 - o “entender por completo las capacidades y limitaciones del sistema de IA de alto riesgo y controlar debidamente su funcionamiento, de modo que puedan detectar indi-

cios de anomalías, problemas de funcionamiento y comportamientos inesperados y ponerles solución lo antes posible”;

- “ser conscientes de la posible tendencia a confiar automáticamente o en exceso en la información de salida generada por un sistema de IA de alto riesgo («sesgo de automatización»), en particular con aquellos sistemas que se utilizan para aportar información o recomendaciones con el fin de que personas físicas adopten una decisión”;
- “interpretar correctamente la información de salida del sistema de IA de alto riesgo, teniendo en cuenta en particular las características del sistema y las herramientas y los métodos de interpretación disponibles”

Por lo tanto, no basta con que haya un humano que vigile. Debe ser un humano cualificado y dotado de conocimientos y habilidades específicas para ello

- Para el caso de la identificación biométrica autorizada (y sólo para ésta, no para otros casos que pueden ser igual de sensibles, por cierto), se aplica la técnica de los *cuatro ojos*: una doble supervisión humana (lo que suscita dudas en caso de divergencia en la opinión de ambos supervisores, supuesto no regulado).
- Cuándo se supervisa: el precepto quiere supervisión humana en todo el ciclo de funcionamiento de los sistemas de IA, desde su concepción hasta su aplicación. Dicha supervisión incluye “dotarlos de una herramienta de interfaz humano-máquina adecuada, entre otras cosas”.
- Cómo se supervisa:
 - Conforme a lo que desde hace tiempo (así, recuérdense las *guidelines* de *Article 29 Working Party* en relación con el RGPD²⁶) se señala, la supervisión debe ser significativa

²⁶ La traducción es nuestra: “Para que se considere intervención humana, el controlador debe asegurarse de que cualquier supervisión de la decisión sea significativa, y no sólo un gesto simbólico. Debe ser llevada a cabo por alguien que

(*meaningful*) esto es, capaz de interferir (“decidir, en cualquier situación concreta, no utilizar el sistema de IA de alto riesgo o desestimar, invalidar o revertir la información de salida que este genere”) y, en su caso, detener el funcionamiento de la IA, mediante lo que podríamos definir como el *botón del pánico*: la capacidad de activar un mecanismo de parada del sistema que funciona deficientemente.

- No se establece por el RIA un estándar determinado de diligencia en la supervisión de la IA. En lo que se refiere a la Administración, ese estándar debe hallarse en la obligación de buena administración de diligencia debida o debido cuidado, un concepto jurídico indeterminado que habrá que precisarse caso a caso, si bien en otro momento hemos propuesto la fórmula $c=p \times D$, una sencilla adaptación de la conocida fórmula del juez estadounidense Learned Hand, donde el coste del nivel de diligencia administrativa (c) ha de ser, como mínimo, igual a la probabilidad de daño (p) multiplicada por la gravedad de éste (D). Por supuesto, la falta de supervisión administrativa que genere daños podrá dar lugar a la correspondiente responsabilidad administrativa²⁷.

5.- CONCLUSIONES Y TEMAS PARA CONSIDERAR EN EL FUTURO

Ante todo, hay que partir de una primera conclusión: que exista IA disponible técnicamente, ahora o en el futuro, para ser aplicada en cualquier ámbito de funcionamiento administrativo no significa que sea necesario y conveniente aplicarla siempre.

Del mismo modo que debe existir una necesaria evaluación de la necesidad o no de contratar bienes y servicios por parte de las AAPP

tenga la autoridad y competencia para cambiar la decisión. Como parte del análisis, debe considerar todos los datos de entrada y salida disponibles”.

²⁷ PONCE SOLÉ, J. (2021): *Manual de Fonaments del Dret Administratiu i de la Gestió Pública*, Tirant Lo Blanch, València, 2ª edición.

(art. 28 de la Ley 9/2017, de 8 de noviembre, de Contratos del Sector Público vigente, LCSP), la cual debe acreditar la necesidad, idoneidad y eficiencia de la contratación, debería de existir, antes incluso de pasar a evaluar ningún sistema de IA, una previa ponderación administrativa de si en un ámbito determinado es necesario, idóneo y eficiente el uso de sistemas de IA, considerando costes, beneficios y riesgos. Evaluación que debiera ser exigida legalmente. El punto de partida, entendemos debiera ser el equivalente al art. 30 de la mencionada LCSP adaptándolo a la IA: el desarrollo de las funciones públicas se desarrollará por humanos, salvo que, previa evaluación ponderada justificada en el expediente se decida que es necesario, idóneo y eficiente el uso de la IA.

Los límites vistos de la IA y sus diferencias con los humanos pueden aconsejar prohibir normativamente su uso en ciertos ámbitos, prohibición que entendemos debería ser aplicando una opción por defecto (*opt ut*) en favor del no uso de la IA, salvo que una norma con rango de ley, que establezca suficientes garantías, la permita, en aplicación de la reserva de ley existente en la CE y del principio de precaución mencionado.

Aceptado el uso de la IA, ahora ya en el marco de la evaluación previa de idoneidad expuesta, la Administración debiera considerar, caso a caso la necesidad, idoneidad y eficiencia del uso de IA. Superado este segundo paso, y en fase ya de evaluación concreta del específico uso de la IA, éste debería acompañarse de mecanismos de supervisión o vigilancia humana que sean efectivos.

Esa vigilancia humana no se satisface de cualquier manera. El humano que supervise en el seno de una Administración debe tener los conocimientos y habilidades precisas, entre ellas, la capacidad de identificar su propio sesgo cognitivo de automatización y ser capaz de revertirlo.

Ello implica la necesidad de una formación y entrenamientos específicos de los empleados públicos que van a vigilar los sistemas de IA en materia de *des sesgamiento* cognitivo, así como formación respecto al funcionamiento de estos. Ante lo que nos podemos plantear si ese tipo de formación puede ser realmente efectiva para la evitación de dicho sesgo, un debate abierto que parece irse decantando por la respuesta afirmativa siempre que tal formación esté convenientemente

diseñada²⁸. Además, la intervención humana debe ser efectiva y significativa, en el sentido de ser capaz de interferir con el funcionamiento y llegar a detener éste en caso necesario.

Aún y con todo, existen voces que cuestionan que la supervisión humana de la IA sea una solución realmente efectiva, que pueda dar respuesta adecuada a las necesidades existentes, debido a la complejidad y rapidez de la IA y las limitaciones humanas en la práctica²⁹. Se trata de un argumento importante, que requerirá ulteriores análisis. Puesto que, de ser cierto en todos o algunos casos, podría convertir la regulación de la supervisión humana en un mecanismo psicológicamente tranquilizador para la sociedad, una especie de *placebo regulador*, pero, quizás, inane³⁰.

Si ello fuera así, si la supervisión humana no permitiera realmente controlar el uso de la IA y evitar los riesgos asociados a la misma, entonces cabría plantearse, caso a caso, si lo que debe hacerse, en realidad, es prohibir el uso de la IA. Desde la conocida perspectiva del principio de proporcionalidad, si no existe otra alternativa igualmente eficaz a la reserva de humanidad, entonces ésta cumpliría el requisito de necesidad.

Por otro lado, la futura aprobación del RIA plantea algunas cuestiones. Una es la posible contradicción entre este y el RGPD en la materia que nos ocupa. El 22 RGPD, como ya sabemos, establece una reserva de humanidad generalizada, excepcionable por norma jurídica y sustituible en tal supuesto por la supervisión humana, en todo caso. Por su parte, el art. 5 del del RIA, establece la prohibición de IA sólo en ciertos casos y el mencionado art. 14, en cuanto a la supervisión humana, se refiere sólo a supuestos de alto riesgo (en los otros

²⁸ HAUSER, C. (2020): "From Preaching to Behavioral Change: Fostering Ethics and Compliance Learning in the Workplace," *Journal of Business Ethics*, Springer, vol. 162(4), pp. 835-855.

²⁹ Muy ilustrativo al respecto es el punto de vista de LEINS, K y KASPERSEN, A. (2021): "7 Myths of Using the Term "Human on the Loop": "Just What Do You Think You Are Doing, Dave?", consultable en: <https://www.bbc.com/mundo/noticias-44482471> (última visita: 28/07/2022).

³⁰ KOULU, R. (2020) "Proceduralizing control and discretion: Human oversight in artificial intelligence policy" *Maastricht Journal of European and Comparative Law*, vol. 27, no 6, p. 720-735.

usos de IA que no implican alto riesgo, el RIA no prevé ni reserva de humanidad ni supervisión humana).

Otra cuestión relacionada con el RIA será la de su impacto en el ordenamiento español. La regulación en éste es muy escasa y dispersa (normativa de protección de datos, LRJSP y LGT, como dijimos, ley 18/2011, de 5 de julio, reguladora del uso de las tecnologías de la información y la comunicación en la Administración de Justicia, art. 42 y anexo definiendo la actuación judicial automatizada...), lo que obliga a los gestores públicos y a los ciudadanos a realizar una dificultosa interpretación sistemática. No es infrecuente escuchar o leer sobre la preocupación (sincera o *inducida* por grupos de presión privados) de que la regulación de la IA puede suponer un freno a la innovación.

Si bien ello puede ser cierto (recuérdese con el surgimiento del automóvil la *Red Flag Act* inglesa de finales del siglo XIX, que exigía la existencia de un ser humano marchando con una bandera roja frente al vehículo para controlar su—entonces muy limitada de por sí—velocidad...), también lo es que el Derecho del siglo XXI conoce de técnicas que permiten incluso promover dicha innovación, como las *sandboxes* (previstas en el RIA), por ejemplo³¹. Asimismo, téngase en cuenta que la innovación puede ser frenada a su vez por la inseguridad jurídica, por lagunas en el ordenamiento jurídico que no ofrezcan seguridad a los emprendedores en este campo.

En esa línea, el RIA ofrecerá un suelo europeo, pero cabe reflexionar sobre si necesariamente un techo español. En la Exposición de Motivos del RIA se indica que se ha elegido la figura del Reglamento (art. 288 TFUE) “por la necesidad de aplicar uniformemente las nuevas normas, tales como la definición de IA, la prohibición de determinadas prácticas perjudiciales que la IA permitiría y la clasificación de determinados sistemas de IA”. Todo ello sin perjuicio de que “las disposiciones del Reglamento no son excesivamente prescriptivas y permiten que los Estados miembros actúen a distintos niveles en relación con aquellos elementos que no socavan los objetivos de la iniciativa, en particular en lo que respecta a la organización interna del sistema

³¹ RANCHORDAS, S. (2021): “Experimental Regulations for AI: Sandboxes for Morals and Mores”, *University of Groningen Faculty of Law Research Paper* No. 7/2021

de vigilancia del mercado y la adopción de medidas para promover la innovación”.

Así, cabría plantearse si el regulador español podría superar esa anomia reguladora y esa dispersión actual en el futuro, estableciendo, en su caso, opciones que, respetando el RIA, puedan completar ésta. Por ejemplo, extendiendo la doble supervisión humana a ciertos supuestos que se consideren igual de sensibles que la identificación biométrica, o estableciendo mayores reservas de humanidad o extendiendo la supervisión humana a más supuestos de los previstos actualmente.

Por otro lado, la supervisión o vigilancia humana de la IA en el ámbito administrativo, si es tan relevante como se sostiene, debería configurarse como un auténtico derecho subjetivo generador de correlativas obligaciones jurídicas para la Administración. Obligaciones cuanto menos de medios, sometidas al estándar de debido cuidado propio de la buena administración. Al respecto, es de interés la sentencia del Tribunal Supremo de 15 de febrero de 2022 núm. de Recurso: 7359/2020, referida a la protección de datos, pero perfectamente extensible a la supervisión humana de la IA:

“En las obligaciones de resultado existe un compromiso consistente en el cumplimiento de un determinado objetivo, asegurando el logro o resultado propuesto, en este caso garantizar la seguridad de los datos personales y la inexistencia de filtraciones o quiebras de seguridad. En las obligaciones de medios el compromiso que se adquiere es el de adoptar los medios técnicos y organizativos, así como desplegar una actividad diligente en su implantación y utilización que tienda a conseguir el resultado esperado con medios que razonablemente puedan calificarse de idóneos y suficientes para su consecución, por ello se las denomina obligaciones “de diligencia “ o “de comportamiento””

En caso de reconocerse tal derecho, probablemente debería reflexionarse sobre la necesidad de incluir éste entre los derechos de los interesados en sus relaciones jurídicas con las Administraciones Públicas de la Ley de Procedimiento Administrativo Común. Asimismo, ¿debería incluir un derecho a la atención humana personalizada similar al que se quiere instaurar para los consumidores? ¿o es razonable pensar que los usuarios de servicios públicos y los consumidores ostenten derechos distintos, menores los de los primeros, y, por

tanto, las empresas y las AAPP obligaciones jurídicas diferentes en este campo?

Lo cierto es que la que ha sido denominada por algunos como la primera regulación de la IA en Derecho español, contenida en el art. 23 de la Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación (BOE de 13 de julio), no aborda las cuestiones aquí planteadas, limitándose a hacer una genérica llamada a la “rendición de cuentas”:

“Artículo 23. Inteligencia Artificial y mecanismos de toma de decisión automatizados.

1. En el marco de la Estrategia Nacional de Inteligencia Artificial, de la Carta de Derechos Digitales y de las iniciativas europeas en torno a la Inteligencia Artificial, las administraciones públicas favorecerán la puesta en marcha de mecanismos para que los algoritmos involucrados en la toma de decisiones que se utilicen en las administraciones públicas tengan en cuenta criterios de minimización de sesgos, transparencia y rendición de cuentas, siempre que sea factible técnicamente. En estos mecanismos se incluirán su diseño y datos de entrenamiento, y abordarán su potencial impacto discriminatorio. Para lograr este fin, se promoverá la realización de evaluaciones de impacto que determinen el posible sesgo discriminatorio.

2. Las administraciones públicas, en el marco de sus competencias en el ámbito de los algoritmos involucrados en procesos de toma de decisiones, priorizarán la transparencia en el diseño y la implementación y la capacidad de interpretación de las decisiones adoptadas por los mismos.

3. Las administraciones públicas y las empresas promoverán el uso de una Inteligencia Artificial ética, confiable y respetuosa con los derechos fundamentales, siguiendo especialmente las recomendaciones de la Unión Europea en este sentido.

4. Se promoverá un sello de calidad de los algoritmos.”

En todo caso, de cara a futuras regulaciones, deberá hacerse un esfuerzo para aunar lo mejor de lo humano con lo mejor de la IA, para evitar así lo contrario: conjugar lo peor de ambos mundos.

Los reguladores deberán hacer un esfuerzo para que no ocurra en el ámbito de la IA el pronóstico de Einstein cuando conoció a Marilyn Monroe, con el que iniciamos este estudio. Esperemos que tal cosa no ocurra en las relaciones entre humanos e IA y acabemos teniendo en nuestras sociedades la (inexistente) empatía y compasión de la IA

junto con los límites y errores cognitivos humanos transferidos a la IA y obstaculizando su correcto funcionamiento.

BIBLIOGRAFÍA CITADA

- ATIENZA, M. (2022): *Sobre la dignidad humana*, Trotta, Madrid.
- BRIDLE, J. (2020): *La nueva edad oscura: la tecnología y el fin del futuro*, Debate.
- COGLIANESE, C. y LAI, A. (2022), “Algorithm vs. Algorithm”, *Faculty Scholarship at Penn Law*. 2795.
- CRAMTON, R. C. (1987) “Beyond the ordinary religion”, *Journal of Legal Education*, 1987, vol. 37, no 4,
- DAMASIO, A. (1994): *El error de Descartes*. Andrés Bello.
- DAMASIO, A. (2021), *Sentir y Sentir*, Destino.
- HAUSER, C. (2020): “From Preaching to Behavioral Change: Fostering Ethics and Compliance Learning in the Workplace,” *Journal of Business Ethics*, Springer, vol. 162(4), pp. 835-855.
- HOFMANN, HCH (2020): “The Duty of Care in EU Public Law—A Principle Between Discretion and Proportionality” *Review of European Administrative Law*, vol. 13, no 2, p. 87-112.
- KAHNEMAN, D., SIBONY, O., SUNSTEIN, C. (2021): *Ruido: Un fallo en el juicio humano*, Debate.
- KOULU, R. (2020) “Proceduralizing control and discretion: Human oversight in artificial intelligence policy” *Maastricht Journal of European and Comparative Law*, vol. 27, no 6, p. 720-735.
- LEINS, K y KASPERSEN, A. (2021): “7 Myths of Using the Term “Human on the Loop”: “Just What Do You Think You Are Doing, Dave?””, consultable en: <https://www.carnegiecouncil.org/media/article/7-myths-of-using-the-term-human-on-the-loop> (última consulta: 28/07/2022)
- MACHADO, A. (1912): Mi infancia son recuerdos de un patio de Sevilla” y “Retrato”. *Campos de Castilla*, 1912.
- MAN, K.; DAMASIO, A. (2019): “Homeostasis and soft robotics in the design of feeling machines”, *Nature Machine Intelligence*, vol. 1, no 10, p. 446-452
- MARTÍN DELGADO, I (2009): “Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada”, *Revista de administración pública*, N° 180, págs. 353-386
- PASTOR, A. y NOGALES, P. (2019): “El futuro del trabajo en la administración pública ¿estamos preparados?” *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, n°. extra 3, págs. 34-51.

- PONCE SOLÉ, J. (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico”, *Revista General de Derecho Administrativo*, núm. 1, enero.
- PONCE SOLÉ, J. (2019): *La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida*, UAH-Defensor del Pueblo.
- PONCE SOLÉ, J. (2021): *Manual de Fonaments del Dret Administratiu i de la Gestió Pública*, Tirant Lo Blanch, València, 2ª edición.
- PONCE SOLÉ, J. (2022), *Acicates (Nudges), bueno gobierno y buena administración. Aportaciones de las ciencias conductuales, nudging y sectores público y privado*, Marcial Pons, Madrid.
- PULLMAN LE, REFAIE N, LALUMIÈRE ML, KRUPP D. (2021):” Is Psychopathy a Mental Disorder or an Adaptation? Evidence From a Meta-Analysis of the Association Between Psychopathy and Handedness”. *Evolutionary Psychology*, October.
- RANCHORDAS, S. (2021): “Experimental Regulations for AI: Sandboxes for Morals and Mores”, *University of Groningen Faculty of Law Research Paper* No. 7/2021
- RANCHORDAS, S. (2022): “Empathy in the digital administrative state”, *Duke Law Journal, Forthcoming*, *University of Groningen Faculty of Law Research Paper*, 2021, no 13.
- SOM, C., HILTY, L.M and KÖHLER, A.R. (2009): “The Precautionary Principle as a Framework for a Sustainable Information Society”, *Journal of Business Ethics*, 85:493–505
- WEIZENBAUM, J. (1976): *Computer power and human reason: From judgment to calculation*, W. H. Freeman & Co
- YOUNG, M. M.; BULLOCK, J. B.; LECY, j. D (2019): “Artificial discretion as a tool of governance: a framework for understanding the impact of artificial intelligence on public administration”, *Perspectives on Public Management and Governance*, vol. 2, no 4, p. 301-313.

Capítulo VI

Responsabilidad administrativa por el uso de la Inteligencia Artificial

RUBÉN MARTÍNEZ GUTIÉRREZ¹

Profesor Titular de Derecho Administrativo

Delegado de Protección de Datos

Universidad de Alicante

Resumen: El presente trabajo establece la necesaria aplicación de la teoría general de la responsabilidad administrativa ante actuaciones y resoluciones de la Administración, organismos y entidades del sector público mediante sistemas de Inteligencia Artificial que generen daños. Para ello se determinarán los elementos que configuran la aplicación de la técnica de la responsabilidad y los supuestos más característicos de uso de la IA por las Administraciones.

Abstract: This paper establishes the necessary application of the general theory of administrative liability in the face of actions and resolutions of the Administration, bodies and entities of the public sector by means of Artificial Intelligence systems that generate damage. To this end, the paper will be determined the elements that make up the application of the liability technique and the most characteristic cases of the use of AI by administrations.

Palabras-clave: Responsabilidad administrativa, Inteligencia Artificial, algoritmo

Key-words: Administrative liability, Artificial Intelligence, algorithm

SUMARIO.- 1.- PLANTEAMIENTO GENERAL.- 2.- ASPECTOS CONCEPTUALES DE LA RESPONSABILIDAD ADMINISTRATIVA POR EL USO DE LA IA.- 2.1.- Concepto de Inteligencia Artificial a los efectos del presente estudio.- 2.2.- Concepto legal de responsabilidad administrativa.- 3.- BREVE REFERENCIA A LAS ESPECIALIDADES DEL PROCEDIMIENTO PARA EXIGIR RESPONSABILIDAD ADMINISTRATIVA.- 3.1.- Iniciación del procedimiento.- 3.1.1.- Inicio de oficio. Una modalidad de inicio con buen encaje en supuestos de daño por IA.- 3.1.2.- Inicio a solicitud de las personas interesadas.- 3.2.- Instrucción del procedimiento.- 3.2.1.- Los informes y dictámenes en los procedimientos de responsabilidad patrimonial.- 3.2.2.- Las especialidades del trámite de audiencia en materia de responsabilidad patrimonial.- 3.3.- Terminación del procedimiento.- 3.3.1.- Los órganos competentes para resolver procedimientos de responsabilidad patrimonial.- 3.3.2.- La resolución de los procedimientos de responsabilidad patrimonial. Especialidades.- 4.- EL PARÁMETRO DE CALIDAD Y LOS ELEMENTOS PARA LA EXIGIBILIDAD DE LA RESPONSABILIDAD ADMINISTRATIVA POR IA.- 4.1.- Elementos para la exigibilidad de la responsabilidad administrativa.- 4.2.- Paráme-

¹ ORCID: 0000-0001-6667-1120.

tro de la calidad en el uso de tecnologías de IA en el sector público.- 5.- SUPUESTOS DE POSIBLE APLICACIÓN DE LA RESPONSABILIDAD ADMINISTRATIVA POR INTELIGENCIA ARTIFICIAL.- 6.- CONCLUSIONES.- BIBLIOGRAFÍA CITADA

1.- PLANTEAMIENTO GENERAL

La posible exigencia de responsabilidad administrativa por el uso de sistemas de Inteligencia Artificial es una cuestión que tiene una enorme relevancia respecto de las garantías y el control en el uso de estas tecnologías, teniendo en cuenta que su uso puede tener efectos perjudiciales para las personas destinatarias de las actuaciones y resoluciones efectuadas mediante dichas técnicas. Por el momento, y teniendo en cuenta que el uso de los sistemas de IA en nuestras Administraciones, organismos y entidades del sector público no está completamente generalizado, no se están planteando demasiados supuestos específicos en los que se aprecie la posible exigencia de responsabilidad administrativa. La normal evolución de la implantación y utilización de la Inteligencia Artificial generará en la actividad de nuestras Administraciones supuestos en los que se generen daños a las personas que éstas no tengan el deber jurídico de soportar y, en consecuencia, en los que proceda la aplicación de la responsabilidad administrativa si se cumplen los parámetros, requisitos y exigencias que la teoría general determina como necesarios para la procedencia de esta técnica de resarcimiento.

Con el objetivo de demostrar esta realidad, en el presente trabajo de investigación comenzaremos analizando el marco conceptual de referencia que vamos a considerar para la posible exigencia de responsabilidad. Una vez establecida esta premisa previa, nos referiremos al marco normativo y a las especialidades del procedimiento para exigir responsabilidad administrativa. Con posterioridad, analizaremos los elementos normativos que se precisan para que proceda el reconocimiento de la responsabilidad administrativa, así como también lo que hemos venido a denominar el parámetro de calidad en el desarrollo de actuaciones administrativas mediante sistemas de Inteligencia Artificial. Finalmente, se analizarán los principales supuestos de uso de la IA por las Administraciones que pueden generar responsabilidad administrativa, y se realizarán las oportunas conclusiones para cerrar este Capítulo.

2.- ASPECTOS CONCEPTUALES DE LA RESPONSABILIDAD ADMINISTRATIVA POR EL USO DE LA IA

2.1.- *Concepto de Inteligencia Artificial a los efectos del presente estudio*

Los conceptos de algoritmos, automatización e inteligencia artificial se suelen utilizar para definir ampliamente la IA, ya que uno de los primeros problemas que nos encontramos al aproximarnos a este concepto es la dificultad de disponer de una definición clara, tal y como ha puesto de manifiesto CERRILLO i MARTÍNEZ². Según este autor podemos señalar que “la inteligencia artificial persigue emular las facultades intelectuales humanas en máquinas para que estas puedan realizar tareas propias de los seres humanos. En el estadio actual de desarrollo de la inteligencia artificial únicamente consigue realizar tareas específicas como traducir textos, conducir vehículos sin conductor o reconocer imágenes (es lo que se conoce como inteligencia artificial limitada) siendo aún un mito, para muchos inalcanzable, conseguir que los ordenadores puedan tener una conducta tan avanzada como una persona respecto a un amplio conjunto de tareas cognitivas (es decir, inteligencia artificial general)”³. La inteligencia artificial se basa en el uso de datos, que tras su procesamiento mediante algoritmos permite obtener resultados concretos a problemas planteados⁴. Ahora bien, será necesario establecer una adecuada programación de los algoritmos para que el procesamiento de datos nos lleve a resultados que puedan ser satisfactorios desde el punto de vista del concepto más amplio de IA, a la vez que puedan ser asumidos por los destinatarios de los mismos entendiéndose dentro del

² En este mismo sentido se ha manifestado CERRILLO i MARTÍNEZ, A. (2019): “El impacto de la inteligencia artificial en el Derecho Administrativo ¿Nuevos conceptos para nuevas realidades técnicas?”, en *Revista General de Derecho Administrativo*, número 50, p. 3.

³ CERRILLO i MARTÍNEZ, A. (2019), p. 3.

⁴ PONCE SOLE, J. (2019): “Inteligencia artificial, Derecho Administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico”, en *Revista General de Derecho Administrativo*, número 50, p. 7-8.

principio de buena Administración⁵. El objetivo sería conseguir evolucionar las tecnologías de *big data* para alcanzar una verdadera IA⁶. Es necesario controlar los algoritmos y considerarlos como normas jurídicas sometidas a los principios básicos del Derecho⁷, ya que en un momento inicial las experiencias de aplicación de la IA han generado controversias⁸. De esta manera, y a los efectos del presente estudio, cuando hablamos de IA, algoritmos, automatización o *big data* nos encontramos ante conceptos convergentes que a los efectos de este trabajo se englobarán todos ellos dentro del concepto más amplio de Inteligencia Artificial, en la línea apuntada por COTINO HUESO⁹.

Este posicionamiento tiene además su reflejo en la normativa de *soft law* y documentos comunitarios¹⁰, en especial la “Carta Ética Europea sobre el uso de inteligencia artificial en los sistemas judiciales” de 2018¹¹, la Comunicación de la Comisión Europea COM (2019) 168 final “Generar confianza en la inteligencia artificial centrada en el ser humano”, las Conclusiones del Consejo (2020/C 342 I/01) “Acce-

⁵ VALERO TORRIJOS, J. (2020): “The Legal Guarantees of Artificial Intelligence in Administrative Activity: Reflections and Contributions from the Viewpoint of Spanish Administrative Law and Good Administration Requirements”, en *European Review of Digital Administration & Law – Erdal*, Volume 1, Issue 1-2, june-december, p. 60.

⁶ VALERO TORRIJOS, J. (2019): “Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena Administración”, en *Revista Catalana de Dret Públic*, Núm. 58, p. 3.

⁷ A mayor abundamiento sobre esta idea véase BOIX PALOP, A. (2020): “Algorithms as Regulations: Considering Algorithms, when Used by the Public Administration for Decision-making, as Legal Norms in order to Guarantee the proper adoption of Administrative Decisions”, en *European Review of Digital Administration & Law – Erdal*, Volume 1, Issue 1-2, june-december, p. 75 y ss.

⁸ SARASIBAR IRIARTE, M. (2019): “La Cuarta Revolución Industrial: el Derecho Administrativo ante la inteligencia artificial”, en *Revista Vasca de Administración Pública*, Núm. 115, septiembre-diciembre, p. 380.

⁹ COTINO HUESO, L. (2019): “Derecho y garantías ante el uso público y privado de inteligencia artificial, robótica y big data”, en BAUZÁ REILLY, M. (Dir.): *El Derecho de las TIC en Iberoamérica*, FIADI – La Ley, Montevideo, p. 917 y ss.

¹⁰ De especial interés para este análisis resulta el trabajo de GAMERO CASADO, E. (2021): “El enfoque europeo de Inteligencia Artificial”, en *Revista de Derecho Administrativo – CDA*, Núm. 20, p. 268 y ss.

¹¹ Adoptada por la Comisión Europea para la Eficacia de la Justicia (CEPEJ) del Consejo de Europa los días 3 y 4 de diciembre de 2018.

so a la justicia: aprovechar las oportunidades de la digitalización”¹² o la “Resolución del Parlamento Europeo, de 20 de enero de 2021, sobre inteligencia artificial: cuestiones de interpretación y de aplicación del Derecho internacional en la medida en que la UE se ve afectada en los ámbitos de los usos civil y militar, así como de la autoridad del Estado fuera del ámbito de la justicia penal”¹³. La UE es consciente de que en los últimos años la “moda” de la IA ha conllevado el surgimiento de múltiples conceptos y definiciones, por lo que mediante la Resolución de 2021 “considera que es necesario adoptar un marco jurídico europeo común con definiciones armonizadas y principios éticos comunes”, solicitando que se adopten las siguientes definiciones en esta materia:

- “Sistema de IA”: todo sistema basado en programas informáticos o incorporado en dispositivos físicos que muestra un comportamiento que simula la inteligencia, entre otras cosas, mediante la recopilación y el tratamiento de datos, el análisis y la interpretación de su entorno y la adopción de medidas, con cierto grado de autonomía, para lograr objetivos específicos”. Se trata de una definición que resulta interesante, ya que según la misma un sistema de IA no tiene porqué ser totalmente autónomo o ajeno a la supervisión humana, de forma que permite la supervisión *a posteriori* de las propuestas de decisión adoptadas por el sistema de IA, y esta realidad parece adecuada para las decisiones en las que existe un cierto grado o margen de discrecionalidad¹⁴.

¹² Publicadas en el DOUE de 14 de octubre de 2020, en cuyo tenor literal se señala expresamente que en materia de IA “ya se está investigando y preparando en la UE la implantación de sistemas de inteligencia artificial en el sector de la justicia y de que la aplicación práctica de dichos sistemas es inminente en algunos Estados miembros”.

¹³ Disponible en la web oficial: https://www.europarl.europa.eu/doceo/document/A-9-2021-0001_ES.html#title1 Fecha de consulta: 15.10.2022.

¹⁴ A mayor abundamiento véase: MARTÍNEZ GUTIÉRREZ, R. (2021): “Inteligencia artificial, algoritmos y automatización en la Justicia. Propuestas para su efectiva implantación”, en *Práctica de Tribunales: Revista de Derecho Procesal Civil y Mercantil*, Núm. 149, y HUERGO LORA A. (2021): “Administraciones Públicas e Inteligencia Artificial: ¿más o menos discrecionalidad?”, en *El Cronista del Estado Social y Democrático de Derecho*, Núms. 96-97.

- Sistema de IA “autónomo”: entendido como “todo sistema de IA que funciona interpretando determinados datos de entrada y utilizando un conjunto de instrucciones predeterminadas, sin limitarse a ellas, a pesar de que el comportamiento del sistema esté limitado y orientado a cumplir el objetivo que se le haya asignado y otras decisiones pertinentes de diseño tomadas por su desarrollador”. En este caso debemos entender que la supervisión humana del algoritmo que posibilita la decisión basada en el procesamiento y tratamiento de datos se produce *a priori*, en su programación. Quizá por ello, los sistemas de IA “autónomos” podrían tener encaje en lo que nuestro ordenamiento legal ha categorizado como actuación administrativa automatizada, con cobertura jurídica específica en los artículos 41 y 42 de la Ley 40/2015, de Régimen Jurídico del Sector Público, donde se define la misma como “cualquier acto o actuación realizada íntegramente a través de medios electrónicos por una Administración Pública en el marco de un procedimiento administrativo y en la que no haya intervenido de forma directa un empleado público”.

En definitiva, y a los efectos del presente trabajo, el concepto de IA que vamos a emplear será el más amplio posible, incluyendo los conceptos de IA, algoritmos, automatización o *big data* en el sector público, dado que además, y como se verá en el siguiente apartado, es la opción que tiene mejor encaje en el concepto de responsabilidad administrativa por el uso de estas tecnologías que necesariamente se debería predicar de cualquiera de las actividades señaladas.

2.2.- Concepto legal de responsabilidad administrativa

El principio de responsabilidad administrativa se materializa en la normativa en una doble dirección, en primer lugar, con el propósito de subsanar la deficiente actuación de las Administraciones públicas en su actividad (sea o no de carácter prestacional), y, en segundo lugar, resarcir a los administrados de los perjuicios que cause el deficiente funcionamiento de las Administraciones en cualquiera de las actividades que éstas realizan. La responsabilidad administrativa clásica (objetiva-patrimonial) se regula en nuestro ordenamiento en los artículos 106.2 de la Constitución Española y en los artículos 32 y siguientes

de la Ley 40/2015, de Régimen Jurídico del Sector Público, en cuanto a los principios y fundamentos de exigibilidad de la responsabilidad, y a lo largo del articulado de la Ley 39/2015, en cuanto al procedimiento de reclamación de responsabilidad patrimonial. Especialmente interesantes para el análisis de la responsabilidad patrimonial desde su perspectiva conceptual y sus efectos son las previsiones del artículo 106.2 de la CE y del 32 y siguientes de la LRJSP. El indicado precepto constitucional señala que “los particulares, en los términos establecidos por la ley, tendrán derecho a ser indemnizados por toda lesión que sufran en cualquiera de sus bienes y derechos, salvo en los casos de fuerza mayor, siempre que la lesión sea consecuencia del funcionamiento de los servicios públicos”. Esta previsión viene a ser concretada por el artículo 32 de la LRJSP, que regula los “Principios de la responsabilidad” estableciendo que:

“1. Los particulares tendrán derecho a ser indemnizados por las Administraciones Públicas correspondientes, de toda lesión que sufran en cualquiera de sus bienes y derechos, siempre que la lesión sea consecuencia del funcionamiento normal o anormal de los servicios públicos salvo en los casos de fuerza mayor o de daños que el particular tenga el deber jurídico de soportar de acuerdo con la Ley.

La anulación en vía administrativa o por el orden jurisdiccional contencioso administrativo de los actos o disposiciones administrativas no presupone, por sí misma, derecho a la indemnización.

2. En todo caso, el daño alegado habrá de ser efectivo, evaluable económicamente e individualizado con relación a una persona o grupo de personas.

3. Asimismo, los particulares tendrán derecho a ser indemnizados por las Administraciones Públicas de toda lesión que sufran en sus bienes y derechos como consecuencia de la aplicación de actos legislativos de naturaleza no expropiatoria de derechos que no tengan el deber jurídico de soportar cuando así se establezca en los propios actos legislativos y en los términos que en ellos se especifiquen.

(...)”

Sobre esta base legal, que prácticamente se ha mantenido invariable desde el punto de vista conceptual en nuestra legislación, la doctrina¹⁵

¹⁵ Para un examen general de esta materia véase GONZÁLEZ PÉREZ, J. (2006): *Responsabilidad patrimonial de las Administraciones públicas*, 4ª Ed., Thomson-

y la jurisprudencia¹⁶ han ido perfilando los caracteres de la responsabilidad y los requisitos exigibles para hacerla efectiva. Así, la responsabilidad patrimonial de la Administración ha sido configurada en nuestro ordenamiento como institución de naturaleza objetiva, de modo que cualquier consecuencia dañosa derivada del funcionamiento de los servicios públicos debe ser en principio indemnizada, porque de lo contrario, se produciría un sacrificio individual en favor de una actividad de interés público que debe ser soportada por la comunidad. Resulta indiferente que la actuación administrativa dañosa sea consecuencia de una actividad normal o anormal de la Administración en la prestación de los servicios públicos, bastando para declararla que como consecuencia directa de la misma se haya producido un daño efectivo (nexo causal), evaluable económicamente e individualizado, pues los preceptos constitucionales y legales que componen el régimen jurídico aplicable extienden la obligación de indemnizar tanto a los casos de funcionamiento normal como a los de funcionamiento anormal de los servicios públicos. Por tanto, podemos afirmar que la responsabilidad patrimonial comprende la obligación que incumbe a la Administración de resarcir a los particulares de los perjuicios económicos que les irroque y no estén obligados jurídicamente a soportar, que en palabras de MARTÍN MATEO y Díez SÁNCHEZ implica “la subsanación de los efectos de las lesiones patrimoniales imputables a la Administración previa determinación de la relación de causalidad, en actividad sujeta al Derecho público o privado”¹⁷.

Como se habrá podido comprobar, la responsabilidad administrativa se puede producir por cualquier actividad de la Administración y

Civitas, Cizur Menor (Navarra), GARCÍA DE ENTERRÍA, E. y FERNÁNDEZ RODRÍGUEZ, T. R. (2002): *Curso de Derecho Administrativo*, Tomo II, 8ª Ed., Civitas, Madrid, pp. 358 a 438, SANTAMARÍA PASTOR, J. A. (2009): *Principios de Derecho Administrativo General*, Tomo II, Madrid, Iustel, pp. 481 a 517, y, MARTÍN MATEO, R. y Díez SÁNCHEZ, J. J. (2012): *Manual de Derecho Administrativo*, 29ª Ed., Thomson Reuters – Aranzadi, Cizur menor (Navarra), p. 453 y ss.

¹⁶ Destacan las SSTs, Sala de lo Contencioso-Administrativo, de 14 de mayo, 4 de junio, 2 de julio, 27 de septiembre, 7 y 19 de noviembre de 1994, y de 28 de febrero y 1 de abril de 1995. Especialmente interesante resulta, debido a su claridad, la Sentencia de la Audiencia Nacional, Sala de lo Contencioso-Administrativo, Sección 8ª, de 18 de junio de 2002.

¹⁷ MARTÍN MATEO, R. y Díez SÁNCHEZ, J. J. (2012), p. 458.

de las entidades del sector público, y por supuesto, también procederá, como se verá en los siguientes apartados, cuando en este ámbito se utilicen tecnologías de IA, algoritmos, automatización o *big data* que provoquen daños en las personas que no tengan el deber jurídico de soportar, cuya casuística solo se apunta ahora de manera incipiente, pero que debido al incremento del uso de estas técnicas en nuestro sector público en los próximos años tendrá a buen seguro una presencia usual y habitual, o al menos, al mismo nivel que en los supuestos de responsabilidad administrativa tradicional.

3.- BREVE REFERENCIA A LAS ESPECIALIDADES DEL PROCEDIMIENTO PARA EXIGIR RESPONSABILIDAD ADMINISTRATIVA

3.1.- Iniciación del procedimiento

3.1.1.- Inicio de oficio. Una modalidad de inicio con buen encaje en supuestos de daño por IA

En el artículo 65 de la LPAC se concentran las “especialidades en el inicio de oficio de los procedimientos de responsabilidad patrimonial” a las que debemos adicionar las específicas para la modalidad de inicio del procedimiento de oficio mediante petición razonada de otros órganos. Ciertamente, la iniciación de procedimientos de responsabilidad patrimonial de oficio por las Administraciones es un supuesto que debemos calificar de excepcional, siendo la regla general la presentación por las personas interesadas de reclamaciones de responsabilidad patrimonial. En cuanto a las especialidades propiamente dichas, el contenido del artículo 65 es redundante respecto a cuestiones genéricas que se desprenden o bien de otros preceptos de la LPAC, o de los propios principios y fundamentos de la responsabilidad patrimonial, que como ya hemos señalado, se regulan en los artículos 32 y siguientes de la LRJSP. Así, y en esta línea, el apartado 1º del artículo 65 establece que “cuando las Administraciones Públicas decidan iniciar de oficio un procedimiento de responsabilidad patrimonial será necesario que no haya prescrito el derecho a la reclamación del interesado al que se refiere el artículo 67”, y que es de un año. Hubiese sido

deseable que, al menos, el legislador se hubiese referido a los posibles supuestos de inicio de oficio de estos procedimientos, dado su carácter excepcional.

Respecto al acuerdo de iniciación de oficio, el apartado 2º del artículo 65 señala que el mismo “se notificará a los particulares presuntamente lesionados, concediéndoles un plazo de diez días para que aporten cuantas alegaciones, documentos o información estimen conveniente a su derecho y propongan cuantas pruebas sean pertinentes para el reconocimiento del mismo. El procedimiento iniciado se instruirá aunque los particulares presuntamente lesionados no se personen en el plazo establecido”, ahora bien, debemos entender que su participación es importante para una mejor acreditación de los hechos y daños. En el caso de los posibles daños ocasionados por el uso de las tecnologías de Inteligencia Artificial, la iniciación de oficio de los procedimientos de responsabilidad patrimonial puede adquirir una importancia mayor que en el resto, dado que en este tipo de supuestos las propias Administraciones son conscientes de los daños que la tecnología puede ocasionar, por ejemplo, los daños provocados por un dron pilotado mediante técnicas de IA que genera daños en personas y bienes que la Administración identifica en ese mismo momento. Lógicamente en este tipo de supuestos susceptibles de generar responsabilidad administrativa la Administración debería iniciar el procedimiento de oficio sin ser necesaria la iniciación del procedimiento mediante la reclamación de la persona interesada.

3.1.2.- Inicio a solicitud de las personas interesadas

La LPAC ha querido regular de manera específica las solicitudes de iniciación en los procedimientos de responsabilidad patrimonial en su artículo 67, es decir, las solicitudes de reclamación de responsabilidad patrimonial. Respecto a esta modalidad de inicio del procedimiento, el citado artículo establece que “los interesados sólo podrán solicitar el inicio de un procedimiento de responsabilidad patrimonial, cuando no haya prescrito su derecho a reclamar. El derecho a reclamar prescribirá al año de producido el hecho o el acto que motive la indemnización o se manifieste su efecto lesivo. En caso de daños de carácter físico o psíquico a las personas, el plazo empezará a computarse desde la curación o la determinación del alcance de las secuelas”. De otro

lado, el apartado 2º del artículo 67, señala que en las solicitudes de reclamación de responsabilidad se estará al contenido general de las solicitudes regulado en el artículo 66 de la LPAC, a cuyo contenido se deberán adicionar las siguientes informaciones y contenidos:

- (1) “especificar las lesiones producidas”,
- (2) señalar “la presunta relación de causalidad entre éstas y el funcionamiento del servicio público”,
- (3) incorporar “la evaluación económica de la responsabilidad patrimonial, si fuera posible”,
- (4) indicar “el momento en que la lesión efectivamente se produjo”, y,
- (5) acompañar a la solicitud/reclamación “de cuantas alegaciones, documentos e informaciones se estimen oportunos y de la proposición de prueba, concretando los medios de que pretenda valerse el reclamante”, es decir, lo que podemos denominar la documentación aneja y la prueba de la solicitud.

Obviamente, ante daños producidos por el uso de tecnologías de IA, cualquier persona que los sufra puede iniciar un procedimiento de reclamación de responsabilidad antes del vencimiento del plazo de prescripción de un año. Ahora bien, en este tipo de supuestos resultará muy interesante y sería muy adecuado que la propia Administración instase el inicio de oficio de los procedimientos cuando tenga posibilidad de hacerlo por haber detectado el mal funcionamiento de las tecnologías de IA que hayan provocado daños indemnizables tal y como hemos señalado en el apartado anterior.

3.2.- Instrucción del procedimiento

3.2.1.- Los informes y dictámenes en los procedimientos de responsabilidad patrimonial

El artículo 81 de la LPAC regula la especialidad de la “solicitud de informes y dictámenes en los procedimientos de responsabilidad patrimonial”, y en su apartado 1º se establece el carácter preceptivo de los informes de los procedimientos de responsabilidad patrimonial por el servicio cuyo funcionamiento “haya ocasionado la presunta

lesión indemnizable”, no pudiendo exceder de diez días el plazo de emisión de los informes. Asimismo, según el apartado 2º del artículo 81 “cuando las indemnizaciones reclamadas sean de cuantía igual o superior a 50.000 euros o a la que se establezca en la correspondiente legislación autonómica, así como en aquellos casos que disponga la Ley Orgánica 3/1980, de 22 de abril, del Consejo de Estado, será preceptivo solicitar dictamen del Consejo de Estado o, en su caso, del órgano consultivo de la Comunidad Autónoma”, es decir, se establece para estos supuestos la exigencia de dictamen obligatorio, y “a estos efectos, el órgano instructor, en el plazo de diez días a contar desde la finalización del trámite de audiencia, remitirá al órgano competente para solicitar el dictamen una propuesta de resolución, que se ajustará a lo previsto en el artículo 91, o, en su caso, la propuesta de acuerdo por el que se podría terminar convencionalmente el procedimiento”. En este sentido, el órgano instructor del procedimiento deberá acometer la realización del trámite de audiencia a la persona o personas interesadas evidentemente antes de la remisión de la propuesta de resolución o de acuerdo al Consejo Consultivo que corresponda, propuesta que deberá remitirse en el plazo de diez días desde la finalización del trámite de audiencia, todo ello para no vulnerar el derecho a la tutela efectiva en el procedimiento máxime si el informe del Consejo Consultivo es esencial para la resolución final (si el trámite de audiencia se realizase después ya no podría valorar el Consejo los argumentos, razones y pruebas de las personas interesadas y contraponerlos a los de la Administración). En cuanto al plazo para emitir el dictamen por el Consejo Consultivo, éste será “de dos meses” y deberá pronunciarse sobre una serie de aspectos que vienen a coincidir con los elementos para valorar si existe o no responsabilidad, a saber: (1) “la existencia o no de relación de causalidad entre el funcionamiento del servicio público y la lesión producida” y, (2) “en su caso, sobre la valoración del daño causado y la cuantía y modo de la indemnización de acuerdo con los criterios establecidos en esta Ley”. A medida que se vayan generalizando los supuestos de utilización de tecnologías de IA que generen responsabilidad, los Consejos Consultivos tendrán que especializarse también en el conocimiento de este tipo de casos, aunque como se verá en el epígrafe 4 de este trabajo la base fundamental de aplicación de la responsabilidad será la teoría general clásica de esta institución.

3.2.2.- Las especialidades del trámite de audiencia en materia de responsabilidad patrimonial

El apartado 5º del artículo 82 regula las especialidades del trámite de audiencia en “los procedimientos de responsabilidad patrimonial a los que se refiere el artículo 32.9 de la Ley de Régimen Jurídico del Sector Público”, que son aquellos en los que se produzcan daños a terceros causados en el marco de una ejecución de un contrato “cuando sean consecuencia de una orden inmediata y directa de la Administración o de los vicios del proyecto elaborado por ella misma”. En estos supuestos el legislador establece que también será necesaria y preceptiva la audiencia al contratista “en todo caso”, y para ello se le notificarán “cuantas actuaciones se realicen en el procedimiento, al efecto de que se persone en el mismo, exponga lo que a su derecho convenga y proponga cuantos medios de prueba estime necesarios”. Es decir, que en estos supuestos, la audiencia será doble, de un lado a la persona o personas afectadas, y de otro lado, al contratista. Cabe decir en este punto, que es muy posible que en un futuro no muy lejano haya Administraciones que procedan al empleo de tecnologías de IA mediante contratación pública, dado el alto grado de especialización de este ámbito, y en estos supuestos se deberán observar las garantías de audiencia al contratista a las que nos hemos referido, y a la vez dirimir una cuestión que quizá no sea tan sencilla, como es saber si un daño producido por el uso de tecnologías de IA derivadas de la ejecución de un contrato han sido o no “consecuencia de una orden inmediata y directa de la Administración o de los vicios del proyecto elaborado por ella misma”, ya que será necesario para ello a buen seguro la aportación de pruebas periciales con un alto grado de especialización técnica en esta materia, cuestión que no será sencilla.

3.3.- Terminación del procedimiento

3.3.1.- Los órganos competentes para resolver procedimientos de responsabilidad patrimonial

La determinación de la competencia para la resolución de los procedimientos de responsabilidad patrimonial se regula en el artículo 92 de la LPAC, y en base al mismo se determinan las siguientes reglas:

- En el ámbito de la Administración General del Estado:
- Los procedimientos de responsabilidad patrimonial se resolverán por el Ministro respectivo con carácter general.
- El Consejo de Ministros será competente cuando una Ley así lo determine y también en materia de responsabilidad del Estado legislador regulada en el artículo 32.3 de la LRJSP.
- En el ámbito autonómico y local, los procedimientos de responsabilidad patrimonial se resolverán por los órganos correspondientes de las Comunidades Autónomas, es decir, haciéndose una extrapolación con el apartado anterior para la AGE, o de las Entidades que integran la Administración Local.
- En el caso de las Entidades de Derecho Público, según lo que determinen sus normas de régimen jurídico respecto a los órganos competentes para la resolución de los procedimientos de responsabilidad patrimonial, aplicándose solo de forma supletoria lo previsto en el artículo 92.

3.3.2.- La resolución de los procedimientos de responsabilidad patrimonial. Especialidades

En cuanto a las especialidades de la resolución de los procedimientos de responsabilidad patrimonial, el legislador determina tres adiciones al régimen general en el artículo 91 de la LPAC:

- Una vez recibido el dictamen para las indemnizaciones reclamadas de cuantía igual o superior a 50.000 euros (o en la cuantía que se establezca en la correspondiente legislación autonómica) y también de aquellos supuestos establecidos en la Ley Orgánica del Consejo de Estado (supuestos del artículo 81.2) o, cuando dicho dictamen no sea preceptivo, y una vez finalizado el trámite de audiencia del procedimiento, “el órgano competente resolverá o someterá la propuesta de acuerdo para su formalización por el interesado y por el órgano administrativo competente para suscribirlo”, en una suerte de “propuesta de terminación convencional”.
- Cuando no proceda el sistema de terminación convencional o no sea suscrito por las partes, el órgano competente resolverá

de conformidad con lo previsto con carácter general en el artículo 88 para la resolución de los procedimientos. Asimismo y adicionalmente, “será necesario que la resolución se pronuncie sobre la existencia o no de la relación de causalidad entre el funcionamiento del servicio público y la lesión producida y, en su caso, sobre la valoración del daño causado, la cuantía y el modo de la indemnización, cuando proceda, de acuerdo con los criterios que para calcularla y abonarla se establecen en el artículo 34 de la Ley de Régimen Jurídico del Sector Público”.

- Se insta en el apartado 3º del artículo 91 de la LPAC una regla de terminación desestimatoria del procedimiento por silencio, según la cual “transcurridos seis meses desde que se inició el procedimiento sin que haya recaído y se notifique resolución expresa o, en su caso, se haya formalizado el acuerdo, podrá entenderse que la resolución es contraria a la indemnización del particular”.

4.- EL PARÁMETRO DE CALIDAD Y LOS ELEMENTOS PARA LA EXIGIBILIDAD DE LA RESPONSABILIDAD ADMINISTRATIVA POR IA

4.1.- Elementos para la exigibilidad de la responsabilidad administrativa

Del análisis de la normativa, de la jurisprudencia y de la doctrina que en buena medida hemos analizado en la primera parte de este trabajo, podemos fijar cuatro elementos para la exigibilidad de la responsabilidad administrativa con carácter general, que por supuesto también se aplicarán cuando se produzcan daños como consecuencia del funcionamiento de las tecnologías de Inteligencia Artificial en el sector público. Así, estos cuatro elementos para la exigibilidad de la responsabilidad a los que nos vamos a referir constituyen un test para conocer si una reclamación de responsabilidad puede prosperar o no.

En primer lugar deberá comprobarse la existencia de un hecho imputable a la Administración. La actuación productora de la lesión de-

be proceder de la Administración como titular de esa actividad, y en materia de IA se pueden presentar supuestos en los que se produzca un daño que constituyen hechos imputables a la Administración que esté utilizando estas tecnologías, especialmente si en la actividad desarrollada con estas técnicas no se respetan los elementos exigidos por el parámetro de calidad que analizaremos en el siguiente apartado.

En segundo lugar, es necesario que el hecho imputable a la Administración haya producido un daño antijurídico que se traduzca en un menoscabo de sus intereses patrimoniales, evaluable económicamente e individualizado. La actividad que la Administración desarrolle mediante las tecnologías de Inteligencia Artificial debe generar un daño patrimonial a la persona que ésta no tenga el deber jurídico de soportar (daño antijurídico). La cuantificación del daño es un problema eminentemente práctico, y siempre que pueda efectuarse con demostración de que el daño patrimonial es real y efectivo e individualizado en una persona o grupo de personas podrá exigirse la indemnización correspondiente a dicha cuantía.

En tercer lugar, debe existir relación de causalidad directa y efectiva (nexo causal) entre el hecho que se imputa a la Administración y el daño producido. En los últimos tiempos se ha atenuado el rigor de la exigencia del nexo causal, admitiéndose que la relación de causalidad aparezca bajo formas mediatas, indirectas o concurrentes, lo que obligará a tener en cuenta todas las circunstancias y resultados a fin de determinar cuáles son las causas reales de la lesión y, en consecuencia, la responsabilidad (total o moderada) que se deriva de tal actuación dañosa¹⁸. El anormal funcionamiento del servicio público genera en estos supuestos una lesión patrimonial al administrado que obviamente debe ser susceptible de resarcimiento por la vía de la responsabilidad.

Y, en cuarto lugar, la ausencia de fuerza mayor, ya que el caso fortuito sí es resarcible. La fuerza mayor se concibe como factor de exclusión de la responsabilidad patrimonial y se presenta como “aquel suceso que está fuera del círculo de actuación del obligado, que no hubiera podido preverse o que previsto fuera inevitable, que haya causado un daño material y directo que exceda visiblemente los acci-

¹⁸ GONZÁLEZ PÉREZ, J. (2006), pp. 509 a 512.

dentes propios del curso normal de la vida por la importancia y trascendencia de su manifestación”¹⁹. Se puede identificar la fuerza mayor con “aquellos hechos que, aun siendo previsibles, sean, sin embargo, inevitables, insuperables e irresistibles, siempre que la causa que los motive sea extraña e independiente del sujeto obligado” (STS, Sala de lo Contencioso-Administrativo, de 23 de mayo de 1986). Las notas características de la fuerza mayor son la imprevisión e inevitabilidad del daño y el origen externo del mismo respecto de la Administración. Por el contrario, el caso fortuito sí origina responsabilidad patrimonial, y puede entenderse como tal aquel “evento interno, intrínseco, inscrito en el funcionamiento de los servicios públicos, producido por la misma naturaleza, por la misma consistencia de sus elementos, con causa desconocida”. En el caso fortuito “hay indeterminación e interioridad; indeterminación porque la causa productora del daño es desconocida (o por decirlo con palabras de la doctrina francesa: “falta del servicio que se ignora”); interioridad, además, del evento en relación con la organización en cuyo seno se produjo el daño, y ello porque está directamente conectado al funcionamiento mismo de la organización”²⁰. Sus elementos característicos son la previsión y la evitabilidad del daño, y su origen interno a la organización administrativa. Por la propia naturaleza de la actuación administrativa desarrollada por tecnologías de IA, que en la mayoría de supuestos requieren una planificación o programación previa, la generación de daños raramente vendrá derivada de supuestos de fuerza mayor, siendo como mucho el caso fortuito lo que podría presentarse en este tipo de supuestos.

¹⁹ Citando las Sentencias del Tribunal Supremo de 15 de febrero de 1968 y de 23 de octubre de 1969 véase GARCÍA DE ENTERRÍA, E. y FERNÁNDEZ RODRÍGUEZ, T. R. (2002), p.402 y, GONZÁLEZ PÉREZ, J. (2006), p. 499.

²⁰ Citando varias Sentencias del Tribunal Supremo véase GARCÍA DE ENTERRÍA, E. y FERNÁNDEZ RODRÍGUEZ, T. R. (2002), p. 401 y 402, y, GONZÁLEZ PÉREZ, J. (2006), p. 496.

4.2.- *Parámetro de la calidad en el uso de tecnologías de IA en el sector público*

El parámetro de calidad en la articulación de tecnologías de Inteligencia Artificial en el sector público es un elemento importante para saber si es posible evitar la responsabilidad administrativa en este ámbito. Con este concepto nos referimos a una serie de elementos que constituyen una garantía de diseño adecuado de las tecnologías de IA, que gracias al respeto de los mismos puede evitarse que aparezcan los supuestos que permiten exigir responsabilidad patrimonial. A buen seguro, con la evolución y maduración de las tecnologías de IA y su impacto creciente en el sector público podrán aparecer nuevos elementos que se vayan integrando en el parámetro de calidad, aunque por el momento podemos identificar tres.

En primer lugar, el cumplimiento de los principios generales y de los derechos y garantías en la programación de los sistemas. Se trata de un parámetro importante que ya comienza a tener incluso reflejo en la legislación administrativa y respecto de los cuales podemos destacar una serie de principios y garantías, siendo conscientes de que existen muchos más que también son de aplicación. El primero de los principios que debe tenerse en consideración es el principio de transparencia en el diseño del *software* cuya importancia es capital y al que nos vamos a referir en específicamente por su importancia en el siguiente párrafo. De otro lado, hablamos también aquí de que las tecnologías de IA deben estar diseñadas con el objetivo de evitar desigualdades y sesgos, es decir, cumpliendo con el principio de igualdad, cuestión que ha sido analizada en uno de los capítulos de esta obra colectiva. Asimismo, interesa destacar aquí el cumplimiento del principio de personalización y proactividad, incorporado en la letra f) del artículo 2 del Real Decreto 203/2021, de 30 de marzo, por el que se aprueba el Reglamento de actuación y funcionamiento del sector público por medios electrónicos, y que se define como “la capacidad de las Administraciones Públicas para que, partiendo del conocimiento adquirido del usuario final del servicio, proporcione servicios precumplimentados y se anticipe a las posibles necesidades de los mismos”. Tal y como señalé en su momento, “se trata de un principio absolutamente necesario en la evolución normal del modelo de Administración electrónica en pleno Siglo XXI, y que se desprende claramente

de algunos de los preceptos de la LPAC como los artículos 66 y siguientes cuando se refieren a las presentaciones en registro electrónico de modelos normalizados” asistidos por tecnologías de IA, y que por ello asocian “la normativa reglamentaria también a cuestiones como la innovación y la inteligencia artificial²¹”, aunque debemos lamentar “la nula regulación de estos aspectos en el texto del articulado” del Real Decreto 203/2021, “tratándose de una carencia que a buen seguro deberá subsanarse en un futuro no muy lejano” para incorporar adecuadamente las exigencias de la IA en el sector público²². A parte del cumplimiento de los principios que hemos señalado y de otros, también se debe tener en cuenta el necesario respeto de las garantías en la programación de los sistemas, como por ejemplo, el respeto del derecho fundamental a la protección de datos personales a través del diseño de las tecnologías de IA en base al principio de minimización de datos y el conocido como privacidad desde el diseño. Lógicamente el cumplimiento de estos requisitos que se desprenden del ordenamiento jurídico – administrativo pueden ayudar a evitar supuestos de responsabilidad administrativa.

En segundo lugar, las Administraciones y entidades del sector público pueden llegar a obtener tecnologías de Inteligencia Artificial básicamente de dos formas. De un lado, mediante un procedimiento de licitación tras el cual una empresa suministre el *software* de IA. En este supuesto, es evidente que el desarrollo del procedimiento de licitación y especialmente los informes técnicos y el pliego de cláusulas técnicas establecerá las razones por las cuales se exige una de-

²¹ Interesantes resultan a este respecto las reflexiones de MARTÍN DELGADO, I. (2018): “El acceso electrónico a los servicios públicos: hacia un modelo de Administración Digital auténticamente innovador”, en DE LA QUADRA SALCEDO, T. y PIÑAR MAÑAS, J. L. (Dirs.): *Sociedad Digital y Derecho*, Imprenta Nacional del BOE, Madrid, p. 179 y ss., y de VALERO TORRIJOS, J. (2018): “La necesaria reconfiguración de las garantías jurídicas en el contexto de la transformación digital del sector público”, en DE LA QUADRA SALCEDO, T. y PIÑAR MAÑAS, J. L. (Dirs.): *Sociedad Digital y Derecho*, Imprenta Nacional del BOE, Madrid, p. 375 y ss.

²² MARTÍNEZ GUTIÉRREZ, R. (2021): “La plena eficacia de la e-Administración. Comentario y notas fundamentales del Real Decreto 203/2021, por el que se aprueba el Reglamento de actuación y funcionamiento del sector público por medios electrónicos”, en *Derecho Digital e Innovación. Digital Law and Innovation Review*, Núm. 8, enero-marzo, p. 9.

terminada programación del algoritmo, pudiéndose entender que los mismos vendrían a ser una especie de aprobación previa del sistema de IA. De otro lado, el desarrollo propio del *software* por los servicios informáticos de la Administración. En este supuesto, y teniendo en cuenta la regulación actual en la legislación administrativa que mejor encaje tiene con la temática de este estudio que es la de automatización administrativa establecida en la LRJSP, en estos momentos ya no es necesario como sí sucedía con la legislación anterior a 2007 la aprobación previa de los programas, ahora bien, y a nuestro juicio, dicha aprobación o al menos un procedimiento justificativo de las características y de las razones que determinan esa programación sería altamente conveniente. Tanto si se trata de tecnologías de IA obtenidas por compra pública como desarrolladas por medios propios, es evidente que un procedimiento de justificación o si se quiere de aprobación previa del *software* garantiza que ha existido un control de legalidad necesario sobre los elementos que configuran el sistema, no dejando la programación del mismo en manos exclusivas de los técnicos con perfil informático, cuestión que es clave para mitigar los posibles supuestos de responsabilidad administrativa. Precisamente, y en esta línea, la Ley 1/2022, de 13 de abril, de Transparencia y Buen Gobierno de la Comunitat Valenciana, ha establecido una novedosa regulación en cuanto a la publicación de información de tecnologías de IA al establecer en su artículo 16.1.l) que las Administraciones Públicas deben publicar una serie de información de relevancia jurídica entre la que se encuentra “la relación de sistemas algorítmicos o de inteligencia artificial que tengan impacto en los procedimientos administrativos o la prestación de los servicios públicos con la descripción de manera comprensible de su diseño y funcionamiento, el nivel de riesgo que implican y el punto de contacto al que poder dirigirse en cada caso, de acuerdo con los principios de transparencia y explicabilidad”. Este precepto debería incorporarse a la legislación de carácter básico como un elemento claro configurador del parámetro de calidad en el uso de las tecnologías de IA en el sector público.

Y, en tercer lugar, con una cierta conexión con lo que acabamos de señalar pero con un marcado carácter finalista en cuanto a la toma o adopción de decisiones a través de tecnologías de IA, las resoluciones que se adopten mediante estas técnicas deben disponer de una motivación adecuada y suficiente que permitan que las personas destina-

tarias de las mismas puedan disponer de una explicación razonable, es decir, que conozcan las razones en base a las cuales se ha adoptado la decisión. Aunque la decisión administrativa se base en tecnologías de IA es necesario que cuente con la motivación, especialmente en el caso de las decisiones discrecionales a las que se refiere expresamente el artículo 35.1.i) de la LPAC.

5.- SUPUESTOS DE POSIBLE APLICACIÓN DE LA RESPONSABILIDAD ADMINISTRATIVA POR INTELIGENCIA ARTIFICIAL

La aplicación de la responsabilidad en actuaciones derivadas del uso de la inteligencia artificial no es una posibilidad que de partida se perciba como habitual en el momento actual. No obstante, en este apartado vamos a comentar una serie de supuestos en los que cabría la exigencia de responsabilidad, apreciándose esta posibilidad de una manera más o menos clara.

En primer lugar, los tratamientos automatizados vinculados a procedimientos administrativos de toma de decisiones que generan sesgos y con ello perjuicios y daños antijurídicos, como por ejemplo podría ser la adjudicación de plazas funcionariales o el algoritmo utilizado en Inglaterra para la determinación de las notas finales de los estudiantes de bachiller en el curso 2019/2020 ante la suspensión de los exámenes por la COVID y siendo necesarias para el acceso a la Universidad²³. Ya han existido supuestos de cierta relevancia jurídica en los que se han usado algoritmos para la adjudicación de plazas docentes, como en la Sentencia italiana TAR – Lazio – Roma de 22 de marzo de 2017, donde se solicitaba la revelación del código fuente del sistema de inteligencia artificial o algoritmo que finalmente fue admitida por el Tribunal. Si en la revisión del mismo se observan sesgos o cualquier otra cuestión que pudiese suponer una potencial anulación del acto administrativo generado con el mismo, se podría abrir la puerta a la presentación de una reclamación de responsabilidad por los daños que se hubiesen ocasionado a la persona interesada afectada por el mal funcionamiento del sistema.

²³ HUERGO LORA A. (2021), p. 83.

En segundo lugar, se pueden producir filtraciones de datos y con ello una vulneración del derecho a la protección de datos personales y su seguridad en los sistemas de Inteligencia Artificial²⁴, y derivado de ello de forma directa que se generen daños a personas por tales circunstancias. A parte de la propia aplicación de la normativa de procedimiento administrativo y de régimen jurídico en materia de responsabilidad, también se deberá estar a lo previsto en el artículo 82 del Reglamento UE 2016/679 General de Protección de Datos, donde expresamente se señala, entre otras cuestiones, que “toda persona que haya sufrido daños y perjuicios materiales o inmateriales como consecuencia de una infracción del presente Reglamento tendrá derecho a recibir del responsable o el encargado del tratamiento una indemnización por los daños y perjuicios sufridos”, estableciendo en dicho artículo las bases de la reclamación.

En tercer lugar, también pueden plantearse supuestos de responsabilidad por mal funcionamiento de sistemas de inteligencia artificial para asistencia o ayuda por las webs y sedes electrónicas de las Administraciones, es decir, los conocidos como Bots o ChatBots, que guían en la realización de procedimientos administrativos y por defectos o errores en la información puedan generar daños. Se trata de una cuestión en la que claramente se aplica la responsabilidad en cuanto a los supuestos de información incorrecta en webs y sedes electrónicas²⁵, por lo que para el caso que aquí nos ocupa también sería de aplicación el artículo 38.2 de la LRJSP, según el cual “el establecimiento de una sede electrónica conlleva la responsabilidad del titular respecto de la integridad, veracidad y actualización de la información y los servicios a los que pueda accederse a través de la misma”. Dado que los

²⁴ HERNÁNDEZ PEÑA, J. C. (2022): *El marco jurídico de la Inteligencia Artificial. Principios, procedimientos y estructuras de gobernanza*, Thomson Reuters – Aranzadi, Cizur Menor (Navarra), pp. 114-115.

²⁵ Respecto de las cuales hemos tenido ocasión de pronunciarnos en algunos trabajos de investigación previos a los que nos remitimos: MARTÍNEZ GUTIÉRREZ, R. (2012): “Servicio público electrónico y responsabilidad”, en *Revista Española de Derecho Administrativo*, Núm. 155, y, MARTÍNEZ GUTIÉRREZ, R. (2014): “La responsabilidad patrimonial en el acceso, la difusión y la reutilización de la información”, en FERNÁNDEZ SALMERÓN, M. y VALERO TORRIJOS, J: *Régimen jurídico de la transparencia del sector público: del Derecho de acceso a la reutilización de la información*, Thomson Reuters – Aranzadi, Cizur Menor (Navarra).

ChatBots son sistemas de información y ayuda ubicados en las webs y sedes electrónicas de las Administraciones, organismos y entidades del sector público, el régimen de responsabilidad es el mismo que en los supuestos de daños producidos por información, documentación y tramitaciones inadecuadas, cuyo reconocimiento jurisdiccional²⁶ es incluso anterior a la ya derogada Ley 11/2007, de Acceso Electrónico de los Ciudadanos a los Servicios Públicos.

Y, en cuarto y último lugar, se pueden producir lo que se podría denominar daños directos provocados por errores en sistemas de Inteligencia Artificial, como por ejemplo, en el caso de caída de un dron auto pilotado, o en sistemas de reparto de servicios de asistencia en ambulancia cuando el algoritmo decida preferencias, o en medicina, en operaciones asistidas por robots, entre otras muchas cuestiones. En este tipo de supuestos en los que la Administración conoce perfectamente la producción del daño y la consiguiente generación de responsabilidad, debería ser la propia Administración la que de oficio inicie directamente el procedimiento de reclamación de responsabilidad patrimonial tal y como hemos señalado en el presente trabajo y como materialización del principio de buena administración y del principio de legalidad. Es necesario que las Administraciones, organismos y entidades del sector público ajusten su funcionamiento a las previsiones legales, y el inicio del procedimiento de reclamación de responsabilidad de oficio no debe considerarse como una posibilidad opcional, sino más bien como una obligación, ya que si se conoce la producción de daños por sistemas de IA y los posibles afectados no debería existir otra opción diferente al inicio del procedimiento de oficio.

6.- CONCLUSIONES

La responsabilidad patrimonial es un instrumento jurídico esencial para garantizar una adecuada prestación de servicios públicos y de la actividad administrativa, dentro de la cual se deben insertar también los sistemas de Inteligencia Artificial. La actividad administrativa

²⁶ VALERO TORRIJOS, J. (2008): “Responsabilidad de la Administración pública por información publicada en Internet (Comentario a la STSJ de Canarias de 25 de enero de 2007)”, en *Aranzadi Social*, número 17.

producida por IA que desarrollan las distintas Administraciones, organismos y entidades del sector público no puede, ni tampoco debe, considerarse exenta de la aplicación del régimen legal general de la responsabilidad administrativa. La actividad administrativa generada a través de sistemas de Inteligencia Artificial no dispone de una diferenciación propia en relación al resto de actuaciones y actividades administrativas y, en consecuencia, si en su deficiente realización o prestación se presentan los elementos indispensables para exigir responsabilidad patrimonial conforme al régimen general, debe entenderse aplicable el mecanismo de la responsabilidad. Con carácter general, la aplicación del instrumento de la responsabilidad es una garantía para los ciudadanos, que les otorga seguridad jurídica en la realización de actividades y procedimientos administrativos, y en la prestación de servicios públicos.

La aplicación del mecanismo de la responsabilidad en la utilización de sistemas de Inteligencia Artificial para el desarrollo de actuaciones administrativas y procedimientos puede resultar muy favorecedor para la agilidad en la prestación de servicios públicos y para una adecuada resolución de procedimientos en tiempo y forma. La utilización de la IA por las Administraciones, organismos y entidades del sector público puede ser muy beneficiosa para las personas interesadas y para la sociedad en general, pero su buen y correcto funcionamiento depende de diversos factores, siendo uno de ellos la generación de confianza para las personas que vayan a ser destinatarias de resoluciones y actuaciones que les afecten generadas por sistemas de Inteligencia Artificial, y para que ello ocurra así, la posible exigencia por los ciudadanos de responsabilidad en este ámbito debe entenderse como elemento de control de la actividad que las Administraciones desarrollen y también de resarcimiento de los posibles daños que se generen mediante este tipo de actuaciones.

Quizá en el momento actual sea difícil vislumbrar con claridad supuestos de aplicación de la responsabilidad patrimonial en este ámbito, pero tal y como sucedió en el pasado con la responsabilidad en el desarrollo de actividades de e-Administración, es cuestión de tiempo que la generalización del uso de técnicas y sistemas de IA por nuestras Administraciones y entidades del sector público comience a generar problemas, daños y situaciones controvertidas. En estos supuestos que a buen seguro llegarán, la responsabilidad administrativa

debe plantearse como un mecanismo corrector, y si las Administraciones cumplen con los principios de legalidad y buena administración deberían ser conscientes de que cuando detecten con carácter previo a cualquier reclamación a instancia de parte una actuación realizada mediante sistemas de IA que haya provocado daños a terceros deberían instar directamente el inicio de oficio del procedimiento de responsabilidad administrativa. De esta manera, las personas que sean destinatarias de actos y resoluciones de procedimientos dictados mediante sistemas de Inteligencia Artificial podrán tener la garantía y seguridad de que ante cualquier eventual problema que se produzca, nuestras Administraciones sabrán emplear los mecanismos correctores que el ordenamiento jurídico nos brinda, siendo sin duda uno de ellos, la aplicación de la responsabilidad administrativa.

BIBLIOGRAFÍA CITADA

- BOIX PALOP, A. (2020): “Algorithms as Regulations: Considering Algorithms, when Used by the Public Administration for Decision-making, as Legal Norms in order to Guarantee the proper adoption of Administrative Decisions”, en *European Review of Digital Administration & Law – Erdal*, Volume 1, Issue 1-2, june-december.
- CERRILLO i MARTÍNEZ, A. (2019): “El impacto de la inteligencia artificial en el Derecho Administrativo ¿Nuevos conceptos para nuevas realidades técnicas?”, en *Revista General de Derecho Administrativo*, número 50.
- COTINO HUESO, L. (2019): “Derecho y garantías ante el uso público y privado de inteligencia artificial, robótica y big data”, en BAUZÁ REILLY, M. (Dir.): *El Derecho de las TIC en Iberoamérica*, FIADI – La Ley, Montevideo.
- GARCÍA DE ENTERRÍA, E. y FERNÁNDEZ RODRÍGUEZ, T. R. (2002): *Curso de Derecho Administrativo*, Tomo II, 8ª Ed., Civitas, Madrid.
- GAMERO CASADO, E. (2021): “El enfoque europeo de Inteligencia Artificial”, en *Revista de Derecho Administrativo – CDA*, Núm. 20.
- GONZÁLEZ PÉREZ, J. (2006): *Responsabilidad patrimonial de las Administraciones públicas*, 4ª Ed., Thomson-Civitas, Cizur Menor (Navarra).
- HERNÁNDEZ PEÑA, J. C. (2022): *El marco jurídico de la Inteligencia Artificial. Principios, procedimientos y estructuras de gobernanza*, Thomson Reuters – Aranzadi, Cizur Menor (Navarra).
- HUERGO LORA A. (2021): “Administraciones Públicas e Inteligencia Artificial: ¿más o menos discrecionalidad?”, en *El Cronista del Estado Social y Democrático de Derecho*, Núms. 96-97.

- MARTÍN DELGADO, I. (2018): “El acceso electrónico a los servicios públicos: hacia un modelo de Administración Digital auténticamente innovador”, en DE LA QUADRA SALCEDO, T. y PIÑAR MAÑAS, J. L. (Dir.): *Sociedad Digital y Derecho*, Imprenta Nacional del BOE, Madrid.
- MARTÍN MATEO, R. y Díez SÁNCHEZ, J. J. (2012): *Manual de Derecho Administrativo*, 29ª Ed., Thomson Reuters – Aranzadi, Cizur menor (Navarra).
- MARTÍNEZ GUTIÉRREZ, R. (2021): “Inteligencia artificial, algoritmos y automatización en la Justicia. Propuestas para su efectiva implantación”, en *Práctica de Tribunales: Revista de Derecho Procesal Civil y Mercantil*, Núm. 149.
- MARTÍNEZ GUTIÉRREZ, R. (2021): “La plena eficacia de la e-Administración. Comentario y notas fundamentales del Real Decreto 203/2021, por el que se aprueba el Reglamento de actuación y funcionamiento del sector público por medios electrónicos”, en *Derecho Digital e Innovación. Digital Law and Innovation Review*, Núm. 8, enero-marzo.
- MARTÍNEZ GUTIÉRREZ, R. (2014): “La responsabilidad patrimonial en el acceso, la difusión y la reutilización de la información”, en FERNÁNDEZ SALMERÓN, M. y VALERO TORRIJOS, J.: *Régimen jurídico de la transparencia del sector público: del Derecho de acceso a la reutilización de la información*, Thomson Reuters – Aranzadi, Cizur Menor (Navarra).
- MARTÍNEZ GUTIÉRREZ, R. (2012): “Servicio público electrónico y responsabilidad”, en *Revista Española de Derecho Administrativo*, Núm. 155.
- PONCE SOLE, J. (2019): “Inteligencia artificial, Derecho Administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico”, en *Revista General de Derecho Administrativo*, número 50.
- SANTAMARÍA PASTOR, J. A. (2009): *Principios de Derecho Administrativo General*, Tomo II, Madrid, Iustel.
- SARASIBAR IRIARTE, M. (2019): “La Cuarta Revolución Industrial: el Derecho Administrativo ante la inteligencia artificial”, en *Revista Vasca de Administración Pública*, Núm. 115, septiembre-diciembre.
- VALERO TORRIJOS, J. (2020): “The Legal Guarantees of Artificial Intelligence in Administrative Activity: Reflections and Contributions from the Viewpoint of Spanish Administrative Law and Good Administration Requirements”, en *European Review of Digital Administration & Law – Erdal*, Volume 1, Issue 1-2, june-december.
- VALERO TORRIJOS, J. (2019): “Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena Administración”, en *Revista Catalana de Dret Públic*, Núm. 58.
- VALERO TORRIJOS, J. (2018): “La necesaria reconfiguración de las garantías jurídicas en el contexto de la transformación digital del sector públi-

- co”, en DE LA QUADRA SALCEDO, T. y PIÑAR MAÑAS, J. L. (Dirs.): *Sociedad Digital y Derecho*, Imprenta Nacional del BOE, Madrid.
- VALERO TORRIJOS, J. (2008): “Responsabilidad de la Administración pública por información publicada en Internet (Comentario a la STSJ de Canarias de 25 de enero de 2007)”, en *Aranzadi Social*, número 17.

Sección Segunda:
Los derechos y la legislación
administrativa como límites a la IA

Capítulo VII

Discriminación, sesgos e igualdad de la inteligencia artificial en el sector público

LORENZO COTINO HUESO¹

*Catedrático de Derecho Constitucional
Universitat de Valencia*

Resumen: Se definen y se presentan algunos ejemplos de sesgos y discriminación algorítmica y se exponen las causas habituales por los que se generan. Asimismo, se exponen los criterios de cuándo los usos públicos de IA pueden tener impacto y relevancia como discriminatorios. Se proyectan al ámbito de los algoritmos los diferentes conceptos de discriminación, subrayando que por lo general se tratará de discriminación “por error”. Se exponen los casos en los que puede ser útil acudir a la discriminación “indirecta”, se explican las “acciones positivas algorítmicas”, combinadas con las acciones de diferenciación y las modulaciones y compensaciones para minimizar la discriminación. Sobre estas bases, los criterios de cuándo un sistema algorítmico público es discriminatorio son detallados. En primer lugar, cuando NO se trata de discriminación especialmente prohibida, se afirma que el control jurídico se superpone en particular con el de arbitrariedad y las exigencias de motivación. Esto sucede en especial en los usos de sistemas IA para ejercer potestades discrecionales. En segundo lugar, cuando un sistema algorítmico público SÍ tiene en cuenta circunstancias especialmente sospechosas, se sostiene que por lo general las garantías se superponen a las garantías de los tratamientos de datos especialmente protegidos. Así, para que un algoritmo diferencie por circunstancias prohibidas se impone una ley de calidad y con garantías y se da un canon de control “mucho más” estricto, e incluso excepcionalmente estricto. Se analizan supuestos complejos, como el uso por el sistema de IA público de “proxies” o datos afines a los especialmente sospechosos de discriminación, así como los tratos diferentes a grupos que los algoritmos generan dinámicamente. La prueba de la discriminación y sesgo algorítmico es un tema

¹ ORCID: 0000-0003-2661-0010. Senior Research Fellow valgrAI: Valencian Graduate School and Research Network of Artificial Intelligence. El presente estudio es resultado de investigación del proyecto MICINN Retos “Derechos y garantías frente a las decisiones automatizadas...” (RTI2018-097172-B-C21); grupo de investigación de excelencia Generalitat Valenciana “Algorithmic law” (Prometeo/2021/009, 2021-24); estancia (AEST/2021/012) y proyecto “Derecho, Cambio Climático y Big Data”, Grupo de Investigación en Derecho Público y TIC como investigador de la Universidad Católica de Colombia; MININN “Transición digital de las Administraciones públicas e inteligencia artificial” (TED2021-132191B-I00) Y “Algorithmic Decisions and the Law: Opening the Black Box” (TED2021-131472A-I00) 2022-2023.

complejo. Inicialmente quien alega la discriminación debe aportar indicios. A partir de ahí, se sostiene que cuando el sector público usa IA hay una inversión general de la carga de la prueba; a partir de ahí se concreta lo que debe probar el proveedor o el usuario del sistema de IA público para que se entienda que no ha discriminado. También, se exponen las mejores prácticas contra la discriminación algorítmica y se afirma que muchos de estos remedios preventivos, así como los estudios de impacto, son ya obligatorios en el diseño de los usos públicos de la IA. Igualmente se recuerdan las futuras obligaciones del reglamento de datos y de IA de la UE frente a los sesgos. Si la prevención no es suficiente, se detalla qué sucede cuando se detecta una discriminación algorítmica pública: mitigación, corrección e incluso cese de uso del sistema de IA público; nulidad, anulabilidad e inderogabilidad de los sistemas de IA, los efectos de la discriminación algorítmica en el contexto de la contratación administrativa y algunas notas sobre la responsabilidad por el uso de IA público. Finalmente, se detalla quiénes pueden reaccionar ante una discriminación algorítmica y qué autoridades deben actuar.

Abstract: Some examples of biases and algorithmic discrimination, and the usual causes for their generation are presented. It is also set out the criteria for when public uses of AI can have an impact and relevance as discriminatory. The different concepts of discrimination are projected to the field of algorithms, underlining that it will generally be discrimination “by error”. The cases in which so-called “indirect” discrimination can be useful are explained as well as “algorithmic positive actions” are explained, combined with differentiation actions and modulations and compensations to minimize discrimination. On these bases, the criteria for when a public algorithmic system is discriminatory are detailed. First, when it is NOT especially prohibited discrimination, it is asserted that the legal control overlaps especially with the control of arbitrariness and the requirements of motivation. This is particularly the case where AI systems are used for discretionary powers. Secondly, when a public algorithmic system DOES take into account particularly suspicious circumstances, it is supported that the guarantees generally overlap with the guarantees for specially protected data processing. Thus, a law with quality is imposed and a strict scrutiny, even exceptionally strict. Complex cases are analyzed, such as the use by the public AI system of “proxies” or related data, as well as different processing of dynamic algorithmic groups. Proof of algorithmic discrimination and bias is a complex issue. Initially, the party alleging discrimination must provide some clue. From there, there is a general reversal of the burden of proof when the public sector uses AI. It is specified what the administration must prove about the proper functioning of the public IA system. Best practices against algorithmic discrimination are set out; many of these preventive remedies are already mandatory in the design of public uses of AI, so are impact studies. Likewise, future EU data and AI regulation obligations against bias are recalled. If prevention is not enough, it is detailed what happens when public algorithmic discrimination is detected: mitigation, correction and even cessation of use of the public AI system; legal nullity of AI systems, the effects of algorithmic discrimination in the context of administrative procurement and some notes on liability for the use of public AI. Finally, there it is specified who can react to algorithmic discrimination and what authorities must act.

Palabras-clave: inteligencia artificial, algoritmos, administración, gobierno, igualdad, no discriminación, derechos fundamentales, sesgos

Key-words: artificial intelligence, algorithms, administration, government, equality, non-discrimination, fundamental rights, bias

SUMARIO.- 1.- SESGOS Y DISCRIMINACIÓN ALGORÍTMICA ¿DE QUÉ HABLAMOS? Y ALGUNOS EJEMPLOS. 1.1.- Sesgos, errores y discriminación algorítmica y la conveniencia de abordarlos conjuntamente. 1.2.- Algunos ejemplos de discriminación algorítmica pública. 2.- ¿POR QUÉ SE GENERAN LA DISCRIMINACIÓN Y LOS SES-

GOS ALGORÍTMICOS?. 3.- ¿RESPECTO DE QUÉ IA PÚBLICA ES EXIGIBLE LA NO DISCRIMINACIÓN?. 3.1.- La igualdad es exigible respecto de los sistemas de IA públicos, sus usos y sus efectos, aunque resulte difícil su categorización. 3.2.- El grado de legalidad exigible para el diseño, implantación y uso del sistema de IA público. 4.- ¿CUÁNDO EL USO PÚBLICO DE UN SISTEMA DE IA TIENE IMPACTO Y SUFICIENTE RELEVANCIA JURÍDICA Y PUEDE CONSTITUIR UNA DISCRIMINACIÓN ALGORÍTMICA?. 4.1.- Los criterios y variables para determinar el impacto, nivel de riesgo y relevancia jurídica de los sistemas algorítmicos públicos. 4.2.- ¿Cuándo el sesgo y el error algorítmico es relevante para constituir una discriminación? Datos estadísticos en discriminaciones indirectas y la necesidad de “recalibrarlos” a la baja para la IA pública. 5.- DISCRIMINACIONES ALGORÍTMICAS POR ERROR, DIRECTAS, INDIRECTAS, INTERSECCIONALES Y ACCIONES POSITIVAS ALGORÍTMICAS. 5.1.- La discriminación o sesgo algoritmo será en la mayoría de supuestos “discriminación por error”. 5.2.- Discriminaciones algorítmicas directas, interseccionales y por asociación. 5.3.- La complejidad de las discriminaciones algorítmicas indirectas. 5.4.- Acciones positivas algorítmicas, deber de diferenciación y modulaciones y compensaciones para la minimización de la discriminación. 6.- ¿CUÁNDO UN SISTEMA ALGORÍTMICO PÚBLICO QUE NO TIENE EN CUENTA CIRCUNSTANCIAS ESPECIALMENTE SOSPECHOSAS ES DISCRIMINATORIO?. 6.1.- La igualdad “en” la ley o norma que habilite el uso del sistema de IA y el control laxo de la posibilidad de establecer diferencias de tratos, razonables y no arbitrarios. 6.2.- Control de discriminación y arbitrariedad y exigencias de motivación en los casos de aplicación algorítmica de la ley. Decisiones discrecionales y supuestos no especialmente sospechosos. 7.- ¿CUÁNDO UN SISTEMA ALGORÍTMICO PÚBLICO QUE SÍ TIENE EN CUENTA CIRCUNSTANCIAS ESPECIALMENTE SOSPECHOSAS ES DISCRIMINATORIO?. 7.1.- Las garantías frente a las discriminaciones algorítmicas especialmente prohibidas se superponen por lo general a las garantías de los tratamientos de datos especialmente protegidos. 7.2.- En la mayoría de casos será necesaria una ley de calidad y con garantías que habilite a que un sistema de IA pueda diferenciar en razón de datos especialmente sospechosos de discriminación. 7.3.- Un canon de control “mucho más” estricto, e incluso excepcionalmente estricto. 7.4.- La necesidad de regular legalmente el uso de datos especialmente protegidos para poder evaluar y minimizar la posible discriminación. 8.- ZONAS GRISES. GARANTÍAS ANTE EL USO POR EL SISTEMA DE IA PÚBLICO DE “PROXIES” Y GRUPOS ALGORÍTMICOS. 8.1.- Tratamientos diferentes a partir de “proxies” o datos afines o indirectos de grupos especialmente protegidos. 8.2.- Creación dinámica de grupos y colectivos humanos y la privacidad y no discriminación colectiva y de grupo. 9.- ¿CÓMO PROBAR LA DISCRIMINACIÓN ALGORÍTMICA PÚBLICA? INVERSIÓN GENERAL DE LA CARGA DE LA PRUEBA Y LA NECESIDAD DE LA TRANSPARENCIA Y EXPLICABILIDAD. 9.1.- Los indicios que debe aportar la persona afectada por el sistema de IA público. Inversión general de la carga de la prueba y sus excepciones. 9.2.- Lo que debe probar el proveedor o el usuario del sistema de IA público. 10.- MÁS VALE PREVENIR QUE DISCRIMINAR. ¿CUÁLES SON LOS REMEDIOS Y MEJORES PRÁCTICAS PARA EVITAR LA DISCRIMINACIÓN ALGORÍTMICA PÚBLICA?. 11.- LA (CRECIENTE) OBLIGACIÓN DE DISEÑO Y ESTUDIOS DE IMPACTO FRENTE A LA DISCRIMINACIÓN ALGORÍTMICA DE LOS SISTEMAS DE IA PÚBLICOS. 11.1.- Los deberes de responsabilidad proactiva, privacidad en el diseño o estudio de impacto ya son exigibles en general a los usos públicos de la IA. 11.2.- Los usos públicos de IA de alto riesgo habrán de cumplir, además, los deberes del RIA. 11.3.- La extensión de estudios de impacto y la obligatoriedad de medidas para la no discriminación de los sistemas de IA públicos. 11.4.- La necesidad de una visión social y colectiva y de integrar garantías de calidad de datos y de participación. 12.- Y SI LOS REMEDIOS NO VALEN, ¿QUÉ PASA CUANDO SE DETECTA UNA DISCRIMINACIÓN ALGORÍTMICA PÚBLICA?. 12.1.- Obligaciones de mitigación, rectificación, corrección e incluso cese de uso

del sistema de IA público. 12.2.- Nulidad, anulabilidad e inderogabilidad de los sistemas de IA público y sus efectos. 12.3.- Los efectos de la discriminación algorítmica en el contexto de la contratación administrativa. 12.4.- Responsabilidad por el uso de IA público. 13.-¿QUIÉNES PUEDEN REACCIONAR ANTE UNA DISCRIMINACIÓN ALGORÍTMICA Y QUÉ AUTORIDADES DEBEN ACTUAR?. BIBLIOGRAFÍA CITADA.

1.- SESGOS Y DISCRIMINACIÓN ALGORÍTMICA ¿DE QUÉ HABLAMOS? Y ALGUNOS EJEMPLOS

1.1.- *Sesgos, errores y discriminación algorítmica y la conveniencia de abordarlos conjuntamente*

Valga como definición introductoria de la “discriminación algorítmica” como la que “se produce cuando los sistemas automatizados contribuyen a un trato diferente injustificado” o, en su caso, un “impacto que desfavorece a las personas” en función de circunstancias específicas especialmente sospechosas de discriminación (raza, color, sexo, religión, edad, origen nacional, etc.) de modo contrario a Derecho².

El concepto de sesgo algorítmico se ha definido como “una anomalía en la salida de los sistemas de IA, debido a los prejuicios o suposiciones erróneas realizadas durante el proceso de desarrollo del sistema o prejuicios en los datos de entrenamiento, por lo que los resultados del sistema de IA no pueden generalizarse ampliamente.”³ Para ISO es “el grado en que un valor de referencia se desvía de la

² WHITE HOUSE (2022): *The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People*, White House Office of Science and Technology Policy, Washington DC., octubre, pp. 10 y ss. <https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf>

³ JOINT RESEARCH CENTRE; ESTÉVEZ ALMENZAR, M. y otros (2022): *Glossary of human-centric artificial intelligence*, Joint Research Centre, Unión Europea, Luxemburgo, pp. 16-17 <https://publications.jrc.ec.europa.eu/repository/handle/JRC129614> DOI:10.2760/860665

verdad”⁴. EL Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial (HLEG) señala que “Un sesgo es una inclinación que favorece o perjudica a una persona, objeto o posición.”⁵

En una de las mejores aportaciones sobre el tema, el NIST aprecia más de cuarenta variedades o tipologías de sesgos⁶ y los agrupa en tres categorías dominantes: los sesgos sistémicos, los sesgos estadísticos y computacionales y los sesgos humanos, a saber.

- Los sesgos sistémicos son esencialmente sesgos institucionales o históricos. No son voluntarios. Aunque se sigan las reglas o normas existentes. Están en los conjuntos de datos y en las normas, prácticas y procesos institucionales a lo largo del ciclo de vida de la IA y en la cultura y la sociedad en general.
- Los sesgos estadísticos y computacionales se dan cuando la muestra no es representativa de la población. El error puede deberse a datos heterogéneos, representación de datos complejos en representaciones matemáticas más simples, datos erróneos y sesgos algorítmicos como el ajuste excesivo o insuficiente, el

⁴ ISO (2006): *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*, ISO, Tech. Rep. ISO 3534-1:2006, <https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standards/04/01/40145.html>

⁵ HLEG (Comisión Europea), (2019): *Directrices éticas para una IA fiable*, Comisión Europea 2019, p. 48 <https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>

⁶ NIST; SCHWARTZ, R. y otros (2022): *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence*, *Special Publication 1270* (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, pp. 6-9 <https://doi.org/10.6028/NIST.SP.1270>

Así en su amplio glosario (pp. 49-55): de actividad, de comportamiento, cognitivo, de amplificación, de anclaje, de información, de automatización, de disponibilidad, de confirmación, del consumidor, de producción de contenidos, de dragado de datos, de generación de datos, de despliegue, de detección, emergente, de evaluación, de exclusión, de bucle, de financiación, histórico, de información humana, implícito, heredado, institucional, de interpretación, en el procesamiento, de enlace, de pérdida de conciencia de la situación, de medición, de confusión de modo, de selección del modelo, de popularidad, de población, de presentación, de clasificación, de representación, social, de supervivencia, temporal, de incertidumbre, de interacción del usuario.

tratamiento de valores atípicos y la limpieza de datos y los factores de imputación.

- Los sesgos humanos, errores sistemáticos en el pensamiento humano. Son una parte fundamental de la mente humana, así se dan sesgos de anclaje, de disponibilidad, el sesgo de confirmación, entre muchos otros. Asimismo se trata de atajos mentales adaptativos. Obviamente, no son exclusivos de las interacciones humanas con la IA, pero en este ámbito, se insertan en todo el ciclo de vida de la IA y en su uso.

Considero que procede abordar los sesgos y errores juntamente con la discriminación algorítmica. Cualquier uso de sistemas de IA habitualmente genera errores y sesgos, en algunos casos estos pueden tener impactos relevantes. Todo sesgo de cierta relevancia en el uso de la IA potencialmente puede suponer un trato diferente no justificado. Y como se verá, ello puede constituir una discriminación por error, una discriminación especialmente prohibida con relación a circunstancias especialmente sospechosas o una discriminación general.

1.2.- Algunos ejemplos de discriminación algorítmica pública

Los supuestos de discriminación algorítmica son de lo más variado⁷. En contextos más relacionados con el uso público de IA, con los sistemas de identificación biométrica e IA, las posibilidades de

⁷ Del ámbito público y privado, Se pueden seguir diversos ejemplos en CONSEJO DE EUROPA; ZUIDERVEEN BORGESIU F (2018): *Discrimination, artificial intelligence, and algorithmic decision-making*. Directorate General of Democracy, CONSEJO DE EUROPA, pp. 23-31, <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>
 BID- Universidad Adolfo Ibáñez (2022): *Formulación ética de proyectos de ciencia de datos*, Banco Interamericano de Desarrollo, División de Gobierno Digital y Universidad Adolfo Ibáñez, pp. 6-9, <https://goblab.uai.cl/guia-formulacion-etica-de-proyectos-de-ciencia-de-datos/>
 WHITE HOUSE (2022), *op.cit.*, pp. 23-24 y pp. 29-30.
 EXECUTIVE OFFICE OF THE PRESIDENT (2016): *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights*. May 2016. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf
 O'NEIL, C. (2017): *Weapons of Math Destruction*. Penguin Books. 2017.

generar errores, sesgos y discriminaciones son muchas, aunque sean involuntarias. Estos sistemas son probabilísticos y aplicados a miles o millones de personas los errores pueden salir muy caros a los afectados. En Buenos Aires Guillermo Federico Ibarrola fue erróneamente identificado por un sistema inteligente como prófugo y estuvo detenido 6 días; el 12 de abril de 2022 un juez suspendió el sistema que fue anulado por sentencia de 7 de septiembre de 2022. La misma suerte y los mismos días que estuvo detenido Robert Williams en Detroit por otro falso positivo de reconocimiento facial⁸. Escáneres corporales de aeropuertos en EEUU (TSA) han marcado a los viajeros transgénero como más probables sospechosos para efectuar un control particular.⁹ Servicios de empleo también en EEUU han generado la discriminación de discapacitados¹⁰. En el ámbito de la salud, un sistema aplicado aproximadamente 200 millones de estadounidenses cada año preferenciaba para atención sanitaria temprana a quienes habían costado más dinero anteriormente. Como las personas negras recibían

Ya en 2008, en el sector público pueden seguirse supuestos reales en CITRON, D. K. (2008): “Technological Due Process”, 85 *Wash. U. L. Rev.* pp. 1249-1313, pp. 1256, http://openscholarship.wustl.edu/law_lawreview/vol85/iss6/2

⁸ Al respecto pueden seguirse mis estudios recientes sobre el tema con remisión a toda la literatura “Sistemas de inteligencia artificial con reconocimiento facial y datos biométricos. Mejor regular bien que prohibir mal”, en *El Cronista del Estado Social*, IUSTEL, *monográfico Inteligencia artificial*, n° 100, septiembre-octubre 2022, pp. 68-79. Una versión extensa, en la obra que coordino con Francisco Balaguer.

⁹ Al respecto, a partir de los supuestos señalados en WHITE HOUSE (2022), *op.cit.*, cabe seguir, CHRISTIAN, A. (2022): *Four Ways the TSA Is Making Flying Easier for Transgender People*. American Civil Liberties Union. <https://www.aclu.org/news/lgbtq-rights/four-ways-the-tsa-is-making-flying-easier-for-transgender-people>

U.S. TRANSPORTATION SECURITY ADMINISTRATION (2022): *Transgender/ Non Binary / Gender Nonconforming Passengers*. TSA, <https://www.tsa.gov/transgender-passengers>

¹⁰ U.S. EQUAL EMPLOYMENT OPPORTUNITY COMMISSION (2022): *The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees*. EEOC-NVTA-2022-2, <https://www.eeoc.gov/laws/guidance/americans-disabilities-act-and-use-software-algorithms-and-artificial-intelligence>

U.S. DEPARTMENT OF JUSTICE (2022): *Algorithms, Artificial Intelligence, and Disability Discrimination in Hiring*. Mayo, <https://beta.ada.gov/resources/ai-guidance/>

menos asistencia médica y habían hecho menos gasto, eran postergados por el sistema respecto de los blancos con enfermedades y necesidades similares¹¹. Son ya conocidos los supuestos de sesgos raciales en el ámbito policial y de aplicación de la ley¹². Así, la predicción de reincidencia de los individuos bajo custodia federal sobreestimaba o subestimaba a diversos grupos de color y hubo de corregirse¹³. No menos del 70% de estados EEUU utilizan desde 2004 algún instrumento mecánico para la decisión de libertad condicional (ej.: *Level of Services Inventory- Revised*, LSI-R). Los sistemas de predicción de riesgo (*risk assessment instruments*, *Public Safety Assessment*, PSA) después de la condena se emplean en más de veinte jurisdicciones de EEUU en los últimos años. Algunos de estos sistemas se usan por jueces para determinar la condena por el riesgo de reincidencia, siendo especialmente conocido el sistema COMPAS “Correctional Offender Management Profiling for Alternative Sanctions”. En el mismo, los negros tienen casi el doble de probabilidades que los blancos de ser

¹¹ OBERMEYER, Z. y otros (2019), *Dissecting racial bias in an algorithm used to manage the health of populations*. Science. Vol. 366, No. 6464. Oct. 25, 2019 <https://www.science.org/doi/10.1126/science.aax2342>

¹² Sobre el uso penal, policial y de aplicación de la ley, la literatura es muy abundante, me permito destacar: MARTÍNEZ GARAY, L. y MONTES SUAY, F. (2018): “El uso de valoraciones del riesgo de violencia en Derecho Penal: algunas cautelas necesarias”, *Indret: Revista para el Análisis del Derecho*, n.º. 2; BARONA, S. (2021): *Algoritmización del derecho y de la justicia: De la inteligencia artificial a la Smart Justice*, Tirant lo Blanch, Valencia; SIMÓ, P. (2022): *La prisión algorítmica*, Tirant lo Blanch, Valencia; SIMÓ, P. y MAGRO, V. (2021): *Justicia cautelar e inteligencia artificial: la alternativa a los atávicos heurísticos judiciales*, J.M. Bosch Editor, Barcelona, 2021; MIRÓ LLINARES, F. (2022): “Inteligencia artificial, delito y control penal: nuevas reflexiones y algunas predicciones sobre su impacto en el derecho y la justicia penal”, en *El Cronista del Estado Social y Democrático de Derecho*, n.º. 100, 2022, pp. 174-183.

¹³ JOHNSON, C. (2022 a): *Flaws plague a tool meant to help low-risk federal prisoners win early release*. NPR. 26 enero, 2022 <https://www.npr.org/2022/01/26/1075509175/flaws-plague-a-tool-meant-to-help-low-risk-federal-prisoners-win-early-release> y (2022 b): *Justice Department works to curb racial bias in deciding who's released from prison*. NPR. 19 de abril, <https://www.npr.org/2022/04/19/1093538706/justice-department-works-to-curb-racial-bias-in-deciding-whos-released-from-pris>; NATIONAL INSTITUTE OF JUSTICE (2021): *Review and Revalidation of the First Step Act Risk Assessment Tool*. National Institute of Justice NCJ 303859. <https://www.ojp.gov/pdffiles1/nij/303859.pdf>

etiquetados como de mayor riesgo de reincidir. Entre los blancos se produce el error contrario. El uso de este sistema dio lugar al famoso caso *Loomis*¹⁴.

En otros ámbitos del uso público de IA, ya hace bastantes años se mencionaban errores masivos por sistemas automatizados en la asignación de ayudas públicas (como el sistema *CalWIN* en California o en Texas o el CBMS de Colorado (*Colorado Benefits Management System*). Unos 1.500 pasajeros semanalmente no podían subir a un avión al ser incluidos erróneamente en la lista “No Fly”, o excluidos de contratación pública¹⁵. En EEUU, la sentencia *K. W. V. Armstrong* 89 F.3d 962, 976 (9TH Cir. 2015) revisó el sistema del *Department of Health and Welfare* que atribuía beneficios de seguridad social por enfermedad en Idaho y se detectaron diferencias desproporcionadas entre distintas zonas y errores estadísticos en la fórmula utilizada.

En el ámbito educativo también es habitual el uso de sistemas IA. Y con ellos los sesgos y la discriminación. Así, un sistema de predicción de abandono de estudios utilizado en más de 500 universidades de EEUU generaba sesgos que multiplicaban por cuatro la predicción respecto de las personas de color¹⁶. En el ámbito de empleo, los sistemas que predicen quién tiene más probabilidades de hacer clic en un anuncio de trabajo dirigían a las mujeres a trabajos de cajas de supermercado y a los negros a empresas de taxis.¹⁷

¹⁴ Sentencia *State v. Loomis*, 881, N.W.2d 749, 7532 (Wis, 2016). Por todos, DE MIGUEL BERIAIN, I. (2018). “Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the Wisconsin v. Loomis ruling, *Law, Probability and Risk*, Volume 17, Issue 1, 1 March 2018, pp. 45–53, <https://doi.org/10.1093/lpr/mgy001> Ver también, CONSEJO DE EUROPA (2018), *op.cit.*, pp. 23-24. Sobre la herramienta, <http://www.equivant.com>

¹⁵ CITRON, D. K. (2008), *ob.cit.*, p. 1256.

¹⁶ FEATHERS, T. (2021): “Major Universities Are Using Race as a “High Impact Predictor” of Student Success: Students, professors, and education experts worry that that’s pushing Black students in particular out of math and science”, en *The Markup*. Marzo <https://themarkup.org/machine-learning/2021/03/02/major-universities-are-using-race-as-a-high-impact-predictor-of-student-success>

¹⁷ BOGEN, M. (2019): “*All the Ways Hiring Algorithms Can Introduce Bias*”. *Harvard Business Review*. Mayo. <https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias>

Siempre es más fácil ver la paja en el ojo ajeno que en el propio. Y los casos conocidos *a este lado del charco* son muchos. Así, en Francia el sistema de admisión *Parcoursup* distribuye estudiantes en establecimientos de educación superior con base en criterios previstos legalmente. En 2020 el Consejo Constitucional francés reconoció el acceso a las características del sistema por cualquier persona y que cada institución de educación superior debe informar sobre los criterios según los cuales se han examinado las solicitudes de *Parcoursup*.¹⁸ En 2020 en Reino Unido por el Covid no se realizaron los exámenes *A-Level* de acceso a la universidad. El organismo regulador (*Ofqual*) creó un sistema basado en un algoritmo que afectó a casi un millón de personas. Los problemas que se generaron llevaron a la dimisión del Ministro de Educación. La crítica más habitual es que se penalizó a los alumnos de colegios públicos. El diseño del algoritmo dio lugar a “sesgos técnicos que no se identificaron ni se corrigieron de forma proactiva”¹⁹.

Es también famoso el sistema de IA público de empleo austríaco PAMAS. El sistema -correctamente- detectaba que ser mujer y de color eran factores de baja probabilidad de encontrar trabajo. El problema fue humano, dado que se programó el sistema para dedicar menos recursos a estos supuestos de baja probabilidad, de este modo se perpetuaba la exclusión social²⁰. De modo bastante similar, el sis-

¹⁸ Criterios regulados en artículo L. 612.3 del Código de educación. Sobre el tema <https://es.frwiki.wiki/wiki/Parcoursup>. La Decisión del Consejo Constitucional <https://www.conseil-constitutionnel.fr/actualites/communiquede/decision-n-2020-834-qpc-du-3-avril-2020-communique-de-presse>

¹⁹ Analiza el caso exhaustivamente GUTIÉRREZ DAVID, M.E., “Acceso al código fuente y a los algoritmos de las Administraciones inteligentes. Lecciones a partir de experiencias comparadas”, en COTINO HUESO L. y CASTELLANOS CLARAMUNT J. (coords.), *Transparencia y explicabilidad de la inteligencia artificial*, Tirant lo Blanch, Valencia, 2022, pp.135-167. La cita es de WACHTER, S., y otros, “Bias Preservation in Machine Learning: The Legality of Fairness Metrics Under EU Non-Discrimination Law”, *West Virginia Law Review*, Vol. 123, núm. 735, 2021, p. 739. <https://x.gd/hKKKqS>.

²⁰ Sobre el tema, SORIANO ARNANZ, A. (2022 a): “El uso de sistemas automatizados por los poderes públicos. El caso del sistema empleado por el servicio público de empleo austríaco”, en COTINO HUESO, L. y TODOLÍ SIGNES, A. (coord.), *Explotación y regulación del uso del big data e inteligencia artificial para los servicios públicos y la ciudad inteligente*, Tirant lo Blanch, Valencia, pp. 192-216.

tema polaco de empleo introdujo en 2014 un sistema con perfilados (*Publiczne Służby Zatrudnienia*, PSZ). Un algoritmo decidía la categoría y puntuación final que otorgaba al desempleado y a partir de tal puntuación el personal del servicio había de tomar decisiones. Sin embargo, el sesgo de automatización llevó a que habitualmente los funcionarios siguieran estos perfiles y puntuaciones automáticas sin mayor supervisión. El sistema no era transparente de cara al público. La Oficina Superior de Auditoría (*Najwyższa Izba Kontroli*) que lo supervisó detectó la ineficacia del sistema y que las mujeres se puntuaban diferente a los hombres, en general, las personas más vulnerables se atribuían al grupo con más dificultades de recibir asistencia del sistema²¹.

El sistema neerlandés SyRI fue especialmente conocido por generar la decisión judicial de uso público de sistemas algorítmicos más importante en Europa²². El *Systeem Risicoindicatie* (SyRI) es un sistema automatizado para detectar diversas formas de fraude, incluidos beneficios sociales, subsidios y fraude fiscal. El sistema podía llegar a ser de “autoaprendizaje y aprendizaje profundo”, “procesaba una -casi ilimitada- cantidad de datos sustancial” (6.50) de muchas categorías al punto que “es difícil imaginar cualquier tipo de datos personales que no sean elegibles para el procesamiento en SyRI” (6.98). Y para generar las alertas de posibles ilegalidades no había “intervención humana significativa”, luego ya la alerta era investigada para en su caso generar consecuencias administrativas. SyRI contaba con amplia regulación legal y reglamentaria y, todo sea dicho, con muchas garantías, incomparablemente superiores a las que puede haber en España, tanto por los desconocidos sistemas de TGSS y AEAT como los más conocidos como el valenciano SALER²³. Sin embargo, se consideró

²¹ <https://algorithmwatch.org/en/poland-government-to-scrap-controversial-unemployment-scoring-system/>

²² Sentencia de 5 de febrero de 2020 del Tribunal de Distrito de la Haya (C / 09/550982 / HA ZA 18-388). Me remito a mi estudio (2020): “SyRI, ¿a quién sanciono?” Garantías frente al uso de inteligencia artificial y decisiones automatizadas en el sector público y la sentencia holandesa de febrero de 2020, en *La Ley Privacidad*, Wolters Kluwer n° 4, mayo. Acceso en Academia

²³ Con bastante profundidad, mi estudio (2020 a) “Hacia la transparencia 4.0: el uso de la inteligencia artificial y big data para la lucha contra el fraude y la corrupción y las (muchas) exigencias constitucionales”, en Carles Ramió (coord.),

que su falta de transparencia generaba vulneraciones de la igualdad, la privacidad y el debido proceso. En el proceso participó el Relator Especial de la ONU precisamente por sus efectos discriminatorios y consideró que SyRI “tiene un efecto discriminatorio y estigmatizador”, pues al investigar vecindarios más marginados “contribuye a los estereotipos y refuerza una imagen negativa de los ocupantes de dichos vecindarios”. La sentencia afirma que el sistema era discriminatorio por falta de transparencia (6.92), “la transparencia, en aras de la verificabilidad» es necesaria frente a los peligros de «exclusión o discriminación injustificadas» (6.91).

En España, entre otros supuestos, el Gobierno estatal creó la aplicación BOSCO, cedida a las compañías eléctricas para verificar los requisitos de los solicitantes para la atribución del bono social eléctrico. Su funcionamiento era más que sospechoso de errores, que no se han podido probar plenamente por la reiterada falta de transparencia sobre el mismo²⁴. También en nuestro país, el sistema *VeriPol* estima la probabilidad de que una denuncia sea falsa y el sistema *VioGén* señala la peligrosidad de posibles hombres maltratadores y puede acabar determinando las decisiones de prisión preventiva, así como las decisiones de política penitenciaria en Cataluña el sistema *RisCanvi*. La mayor o menor opacidad de estos sistemas no permite valorar ni su incidencia efectiva en las decisiones que se adoptan ni la posible existencia de sesgos o errores que puedan generar discriminación²⁵.

En el ámbito privado son bien conocidos los sistemas de IA que han generado sesgos y discriminación en el ámbito de la contra-

Repensando la administración digital y la innovación pública, Instituto Nacional de Administración Pública (INAP), Madrid, 2021. <https://links.uv.es/FUW2pz6>

²⁴ Me remito a los diversos estudios en COTINO HUESO L. y CASTELLANOS CLARAMUNT J. (coords.) (2022), *ob. cit.*, en particular GUTIÉRREZ DAVID, M.E.), *ob. cit.*, pp. 135-167 y VESTRI, G. (2022): “Denegación vs. Derecho de acceso al código fuente en los sistemas algorítmicos. Una perspectiva jurídico-administrativa”, pp. 111-134. Y (2022): “El acceso a la información algorítmica a partir del caso bono social vs. Fundación Ciudadana Civio”, *Revista General de Derecho Administrativo*, n° 61, 2022.

²⁵ Al respecto, me remito a la investigación coordinada por el equipo liderado por Lucía Martínez Garay Universidad de Valencia-Amnistía Internacional sobre las insuficiencias de estos sistemas.

tación, los buscadores, asociación de textos, generación de contenidos, etiquetado de imágenes, personalización de precios, de publicidad, análisis biométrico de sentimientos y lecturas de emociones, concesión de préstamos, fijación de primas de seguro y un largo etcétera. Baste ahora señalar que en muy buena medida estos sesgos y discriminaciones son claramente exportables a los usos públicos de IA.

2.- ¿POR QUÉ SE GENERAN LA DISCRIMINACIÓN Y LOS SESGOS ALGORÍTMICOS?

Los sesgos y errores se generan antes de concebirse un sistema de IA, en cualquiera de sus fases o bien en el uso posterior de estos sistemas por los usuarios o de sus resultados de salida²⁶. Los motivos por los que un sistema de IA genera sesgos y discriminaciones pueden ser de lo más variado²⁷. No obstante, como señalé hace años,

²⁶ NIST; SCHWARTZ, R. y otros (2022), *op.cit.*, pp. 6-7.

²⁷ Sobre las causas, además de la referida guía del NIST, resultan de especial interés los clásicos de O'NEIL, C. (2017), *ob. cit.*, pp. 15-31; BAROCAS, S. y SELBST, A. D. (2016): "Big Data's disparate impact". *104 Calif Law Rev* 671, <http://dx.doi.org/10.2139/ssrn.2477899>. Una buena descripción, TISCHBIREK, A. (2020): "Artificial Intelligence and Discrimination: Discriminating Against Discriminatory Systems", en WISCHMEYER, T y RADEMACHER, T. (ed.), (2020): *Regulating Artificial Intelligence*, Springer, pp. 103-121, ver pp. 104-109. <https://doi.org/10.1007/978-3-030-32361-5> ; HLEG (2019), *op.cit.*, p. 48, CONSEJO DE EUROPA (2018), *op.cit.*, p. 15, nota. 28; VALLE ESCOLANO, R., "Transparencia en la inteligencia artificial y en el uso de algoritmos: una visión de género", COTINO HUESO L. y CASTELLANOS CLARAMUNT J. (coords.) (2022), *ob. cit.*, pp. 85-110, ver pp. 96 y ss., 98-99; ZLOTNIK, A. (2019): "Inteligencia Artificial en las Administraciones Públicas: definiciones, evaluación de viabilidad de proyectos y áreas de aplicación", *Boletic*, núm. 84, pp. 27-28, <http://apiscam.blogspot.com/2019/11/boletic-revista-de-astic-monografico-de.html> así como los diversos trabajos de SORIANO ARNANZ, A. (2021 b): "Decisiones automatizadas: problemas y soluciones jurídicas. Más allá de la protección de datos", *Revista de Derecho Público: Teoría y Método*, Vol. 3, pp. 85-127, DOI:10.37417/RPD/vol_1_2021_535 o SORIANO ARNANZ, A. (2022 b): "Discriminación algorítmica: garantías y protección jurídica", en COTINO HUESO, L. (coord.), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Aranzadi, Cizur Menor, 2022, págs. 139-169.

los grandes problemas de la inteligencia artificial por lo general se deben a la *estupidez humana*²⁸. Son los humanos quienes deciden crear el sistema, para qué usos, a partir de qué datos alimentarlo, de dónde obtenerlos, qué valor otorgarles a los datos y a los resultados que arroja el sistema.

Un problema importante es la “tendencia a confiar automáticamente o en exceso en la información de salida generada por un sistema de IA de alto riesgo (“sesgo de automatización”, art. 13.4º b) RIA. El ser humano está sesgado de fábrica y ello se refleja en la forma en que el sistema fue programado o los datos seleccionados o procesados. Y especialmente se pueden dar los problemas y sesgos respecto de diversos colectivos si tres de cada cuatro profesionales, normalmente del ámbito STEM, son hombres blancos heterosexuales acomodados católicos-cristianos (77%). Sólo el 15 % de los científicos de datos son mujeres y menos del 10% entre los de científicos de datos²⁹. Por ejemplo, los contenidos de Wikipedia (y de tantos otros sitios de los que luego *beben las máquinas*) según ella misma proceden de hombre blanco, técnico, con educación formal, de habla inglesa, de 15 a 49 años, de un país cristiano y del hemisferio norte. Son muy intensas las infrarrepresentaciones.³⁰

Una IA diseñada por colectivos de este tipo involuntariamente ha de tener falta de empatía hacia los colectivos tradicionalmente discriminados, así como “puntos ciegos”, como los que tenemos con el retrovisor del vehículo.³¹

También, la causa de sesgos y discriminaciones puede ser un *erróneo diseño y elección de algoritmos, del peso que atribuyen a unos u otros factores* o a los errores en el desarrollo de los sistemas de apren-

²⁸ COTINO HUESO, L. (2017): “Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales”, en *Dilemata. Revista Internacional de Éticas Aplicadas*, nº 24, 2017. pp. 131-150, <https://goo.gl/iERVha>

²⁹ https://www.burtchworks.com/wp-content/uploads/2019/06/Burtch-Works-Study_DS-PAP-2019.pdf

³⁰ https://en.wikipedia.org/wiki/Wikipedia:Systemic_bias#The_%22average_Wikipedian%22

³¹ Sobre este punto ciego, un clásico, PRONIN, E. y otros, (2002).: “The Bias Blind Spot: Perceptions of Bias in Self Versus Others”, *Volume 28, Issue 3*, SAGE, <https://doi.org/10.1177/0146167202286008>

dizaje automático. Pese a sistemas y personas del todo adecuadas, es extremadamente difícil operacionalizar todos los aspectos relevantes de una realidad social. De ahí que los programadores pueden obviar algunas variables, introducir sesgos en el sistema o escoger bases de datos que ofrecen una representación parcial de la realidad. Frente a ello, los seres humanos tienen la capacidad de apreciar determinados elementos intangibles que no son captados por los sistemas automatizados.³² Un sistema de IA correcto utilizado en un contexto diferente o para una población diferente para los cuales fue originalmente desarrollado puede generar problemas³³.

Cabe recordar que la existencia de correlaciones que fácilmente encuentran los sistemas de *machine learning* no implica causalidad, por lo que deben revisarse las reglas que se generen. Por ejemplo, se suele mencionar con ánimo jocoso que hay una correlación casi perfecta entre el consumo de *mozzarella* en Estados Unidos y el número de doctorados en Ingeniería Civil entre 2000 y 2009, o la edad de Miss América y los asesinatos con objetos de calor o vapor³⁴.

Con especial significación, *los sesgos y discriminaciones se generan por causa de los datos de entrenamiento*. Se afirma así “Garbage in, garbage out” (GIGO), “basura entra, basura sale”³⁵: datos de baja calidad llevan a resultados baja calidad. El Parlamento UE en su resolución sobre macrodatos ha insistido en que “los datos de capacitación a menudo son de una calidad cuestionable y no son neutrales” (Letra B), la “baja calidad” de los datos o los procedimientos “podrían dar lugar a algoritmos sesgados, correlaciones falsas, errores, una subestimación de las repercusiones éticas, sociales y legales” (Considerando

³² En esta dirección, SORIANO ARNANZ, A. (2021 b), *ob.cit.*, p. 90.

³³ VALLE ESCOLANO, R. (2022), *ob. cit.*, p. 99.

³⁴ <http://tylervigen.com/spurious-correlations>

³⁵ https://es.wikipedia.org/wiki/Garbage_in,_garbage_out. Ver FRA, (2018): UE. (2018). #BigData: *Discrimination in data-supported decision making*, European Union Agency for Fundamental Rights, Viena, <https://fra.europa.eu/en/publication/2018/big-data-discrimination> ; FRA- Agencia de la Unión Europea para los Derechos Fundamentales, (2019): *Data quality and artificial intelligence- mitigating bias and error to protect fundamental rights*, Luxembourg, Publications Office, junio, <https://fra.europa.eu/en/publication/2019/artificial-intelligence-data-quality>
NIST; SCHWARTZ, R. y otros (2022), *op.cit.*, pp. 15-16 y 17-20.

m)³⁶. Muchos conjuntos de datos no son generalizables o representativos, sin embargo se utilizan para entrenar los sistemas. Los grupos desfavorecidos están sistemáticamente subrepresentados, así como los datos utilizados para el procesamiento del lenguaje natural difieren significativamente de sus aplicaciones del mundo real. También se alerta de que se reutilizan conjuntos de datos, a partir de sistemas ya diseñados con deficiencias o los generados por usuarios finales (por ejemplo, de redes sociales), por lo que el sistema es vulnerable a sesgos humanos o estadísticos aditivos. Los daños pueden perpetuarse y amplificarse. Las mismas series históricas de datos que alimentan al sistema de IA incorporan, reproducen y perpetúan estructuras de discriminación. Lo peor, además, es que se pueden generar espirales de sesgo, error y discriminación³⁷ pues los sistemas muy posiblemente acentuarán sus decisiones al nutrirse de nuevos datos cada vez más negativos para los sectores perjudicados.

Y, por supuesto, los sesgos y discriminaciones pueden provenir porque el proveedor del sistema de IA público sea privado y puede perder de objetivos de interés general como la protección de los segmentos más vulnerables.

3.- ¿RESPECTO DE QUÉ IA PÚBLICA ES EXIGIBLE LA NO DISCRIMINACIÓN?

3.1.- *La igualdad es exigible respecto de los sistemas de IA públicos, sus usos y sus efectos, aunque resulte difícil su categorización*

El derecho a la igualdad tiene efecto directo respecto de la actuación de los poderes públicos (art. 53.1 CE) y como se verá prohíbe la

³⁶ PARLAMENTO EUROPEO (2017): *Resolución de 14 de marzo de 2017, sobre las implicaciones de los macrodatos en los derechos fundamentales: privacidad, protección de datos, no discriminación, seguridad y aplicación de la ley* (2016/2225(INI)). Acceso en <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0076+0+DOC+XML+V0//ES>

³⁷ En particular BAROCAS, S. y SELBST, A. D. (2016), *ob. cit.* y los diversos trabajos de Soriano.

discriminación tanto en general cuanto la relativa a colectivos sospechosos de ser discriminados. La Ley 15/2022, de 12 de julio integral para la igualdad de trato y la no discriminación en principio sólo sería aplicable para los supuestos de discriminación y sesgos algorítmicos vinculados con las categorías especialmente sospechosas de discriminación pero no para los supuestos generales. En todo caso, seguir esta ley facilita la comprensión de la atracción para el ámbito jurídico del uso de sistemas públicos de IA. Sus obligaciones son de “aplicación al sector público” (art. 2.4º Ley 15/2022). Cualquier uso material de IA por el sector público quedaría expresamente en su ámbito de aplicación y obligaciones (art. 3), pues alcanzan expresamente a la “inteligencia artificial y gestión masiva de datos, así como otras esferas de análoga significación”, además de “(n) internet, redes sociales y aplicaciones móviles”. De igual modo, yendo más allá del estricto procedimiento administrativo, la no discriminación también se exige en las actividades relativas al empleo público (b), educación (d), sanidad (e), seguridad ciudadana (h), administración de justicia (i), protección social, las prestaciones y los servicios (j) sociales, acceso, oferta y suministro de bienes y servicios a disposición del público (k), transporte (f) o cultura (g).

Más allá de la estricta actuación administrativa y procedimental, un “incidente discriminatorio” (art. 25. 3º Ley 15/2022) puede darse por “toda disposición, conducta, acto, criterio o práctica” (art. 6 Ley 15/2022), así como “la inacción, dejación de funciones, o incumplimiento de deberes” (art. 4.1º Ley 15/2022).

Como más adelante se señala, la existencia de un sistema de IA público podrá y en ocasiones deberá estar regulado por leyes o reglamentos y disposiciones generales. En algunos casos estas normas podrán ser objeto de posible impugnación directa o indirecta, especialmente si las normas regulan la posibilidad de establecer diferencias de trato. En la mayoría de los casos la discriminación o sesgos algorítmicos se deberán a decisiones adoptadas o no en el desarrollo, implantación y uso del sistema probablemente no formalizadas y de más que de dudosa naturaleza jurídica. O en su caso, por el funcionamiento autónomo del sistema de IA. En cualquier caso, serán “criterios o prácticas” que pueden ser discriminatorios. Aunque la existencia del sistema de IA o las decisiones sobre el su funcionamiento o efectos no adquieran forma

jurídica o formalmente no formen parte propiamente del procedimiento, ello no es obstáculo para generar una discriminación.

No es necesario que una discriminación o sesgo algorítmico se vincule necesariamente a actuaciones administrativas “en el marco de un procedimiento administrativo” como reza el artículo 41 LRJSP. Ahora bien, si se está en el terreno del artículo 41 LRJSP se habrá determinado el órgano competente del sistema “que debe ser considerado responsable a efectos de impugnación” (art. 41.2º LRJSP). También en el ámbito de este precepto, en el caso de la AGE puede focalizarse la atención en la “resolución” de autorización del sistema (art. 13. 2º Real Decreto 203/2021, RAFESP), igual que en el caso de Cataluña (art. 54.2º Decreto 76/2020, de 4 de agosto, de Administración digital de Cataluña). En el ámbito de la contratación de sistemas de IA, hay que acudir a las vías de impugnación que pudieran darse en este contexto, así como a la posibilidad de exigencia de responsabilidades a los contratistas. Como señala Boix, no está muy claro si se puede impugnar el algoritmo o programa empleado en sí mismo, o solo es posible atacar sus actos de aplicación³⁸. El uso del sistema de IA en su caso puede ubicarse en las diligencias preliminares, en muchas ocasiones puede ser considerado un acto de trámite como un eslabón del procedimiento. En todo caso, de cara a su impugnabilidad habría que determinar la fase de procedimiento específica, su naturaleza como acto de iniciación, ordenación, o de instrucción y distinguir si tiene mayor o menor sustantividad propia. Habría que determinar el tipo de acto de trámite, si es acto de terminación o firme. En cualquier caso, ante dificultades de esta categorización siempre se podrá acudir a los conceptos habituales y menos formales en el ámbito de la no discriminación.

Igualmente, la discriminación se puede dar por “inacción” (art. 4.1º Ley 15/2022) o el “incumplimiento de los deberes”, como puedan ser las obligaciones de responsabilidad proactiva en el diseño, control o evaluación del sistema de IA que exige la normativa de protección de datos u otras regulaciones que luego se analizan. La

³⁸ BOIX PALOP, A. (2020): “Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la administración para la adopción de decisiones”, en *Revista de Derecho Público: Teoría y Método*, Vol. 1, 2020, p. 261 https://doi.org/10.37417/RPD/vol_1_2020_33

inacción puede ser atacable por las vías de acción frente a la inactividad de la Administración (art. 29 LJCA). En tanto en cuanto está implicado el artículo 14 CE en principio será una posibilidad acudir directamente a la jurisdicción contenciosa a través del procedimiento para la protección de los derechos fundamentales (arts. 114 y ss.). Y no genera particular problema acudir a la vía de reclamación de responsabilidad patrimonial por el funcionamiento normal o anormal de los servicios públicos en razón del uso de sistemas de IA (arts. 32 y ss. LRJSP). En este punto también es de interés recordar que “acreditada la discriminación se presumirá la existencia de daño moral” (art. 27 Ley 15/2022, de 12 de julio) y cabe acudir al régimen de la responsabilidad patrimonial de las Administraciones Públicas.

En muchos supuestos, una vía de *aterrizar* jurídicamente el uso de un sistema de IA público será el tratamiento de datos personales que ello implica. Esencialmente: cuando la salida o decisión del sistema de IA público se individualice y aplique a una persona física; así como cuando el sistema de IA público suponga la elaboración de perfiles o decisiones automatizadas a partir de datos personales facilitados, o también los inferidos por el sistema. Igualmente habrá tratamiento de datos en la mayoría de los supuestos de personalización y servicios 360° de la Administración y un largo etcétera. En estos supuestos y como consecuencia, ello supondrá la aplicación en bloque del régimen de protección de datos al sistema de IA público. Y como se verá, una discriminación algorítmica muy posiblemente implicará el incumplimiento de los principios básicos como el de calidad de los datos o las muchas obligaciones de responsabilidad proactiva.

Además de un tratamiento de datos, el uso público de IA podrá suponer también una decisión automatizada (art. 22 RGPD) con sus especiales garantías. Y como se dirá, en muchos casos el régimen y garantías de las discriminaciones especialmente prohibidas es fácil que converja y se superponga con la regulación del tratamiento de datos especialmente protegidos (art. 9 RGPD).

Igualmente y por último, no hay que descartar que las actuaciones del sistema de IA público discriminatorias, se enmarquen en contextos de infracción sancionadora o penal en su caso, particularmente sensibles y ajenos por lo general a la necesidad de que el hecho se enmarque en el procedimiento Administrativo.

3.2.- *El grado de legalidad exigible para diseño, implantación y uso del sistema de IA público*

Hay que contextualizar normativamente la existencia, usos y efectos del sistema de IA público y su relación con el principio de legalidad y con el ejercicio de potestades administrativas³⁹. Como es sabido, el principio de legalidad es muy variable en sus intensidades y así habrá de serlo respecto del uso público de IA y todas las acciones tendentes a ello. Habrá de exigirse una mayor regulación de la posibilidad de uso de estos sistemas según su impacto y cuando pueda implicar actuaciones restrictivas de derechos, penales, sancionadoras o determinar la posibilidad de ejercer derechos. La exigencia de legalidad se debilita en los casos de usos de IA en actuaciones favorables o ampliatorias de derechos como en el contexto de prestación de servicios públicos, sanidad, educación, etc. En la misma dirección, habrá que intentar delimitar si el sistema de IA se utiliza para el ejercicio de potestades administrativas, pues la decisión de empleo, diseño y desarrollo del sistema de IA habrán de contar con variable cobertura legal y normativa. No obstante, no siempre la ley y las normas determinan expresamente la potestad administrativa en sí. La potestad puede ser implícita o inherente al cumplimiento de una competencia o atribución. Lamentablemente, es muy posible que las acciones más importantes para el diseño, configuración, implantación y evaluación del sistema de IA queden al margen del reconocimiento normativo y se consideren potestades implícitas o inherentes. Sin embargo, en la mayoría de los casos estas decisiones son los elementos determinantes para generar el sesgo o discriminación algorítmica. Así las cosas, el uso de la IA pública puede contar con una muy variable cobertura y delimitación normativa. Aquí se apuesta firmemente por el uso de los sistemas IA para el ejercicio de potestades discrecionales, precisamente como garantía frente a los *humanos*. Y, como luego se insiste, dado que estos usos de IA suponen más riesgos, debe exigirse una mayor habilitación normativa que incluya garantías. En esta línea hay que tener en cuenta el XVIII. 6 d) de la Carta de De-

³⁹ Se proyecta para este ámbito GAMERO CASADO, E. y FERNÁNDEZ RAMOS, S. (2016): *Manual básico de Derecho Administrativo*, (13ª ed.), Tecnos, 2016, págs. 77-78.

rechos Digitales que impone normativa con garantías adecuadas para las decisiones discrecionales automatizadas.

La exigencia de densidad legal normativa serán mayores según mayor sea el peligro e impactos en derechos e intereses, el mayor grado de automatización y participación humana, el peligro por la tecnología empleada. A lo anterior hay que añadir específicamente las variables propias al ámbito de la igualdad. Mayor cobertura normativa será necesaria especialmente si el sistema de IA establece diferencias a colectivos y especialmente a los sospechosos de discriminación. Aunque sin excesiva claridad, en el artículo 2.2º Ley 15/2022 exige una ley o disposición general). Asimismo, si el sistema de IA público trata datos personales, cosa que no debe ser nada extraña, la base de legitimación legal es la exigencia general (Informe 175/2018 AEPD) y sólo excepcionalmente el consentimiento puede servir como base de legitimación. Y como se expondrá *infra* esta exigencia de legalidad se hará muy intensa en los supuestos de que el sistema de IA trate datos especialmente protegidos, que es precisamente lo más habitual en los casos de supuestos especialmente sospechosos de discriminación.

4.- ¿CUÁNDO EL USO PÚBLICO DE UN SISTEMA DE IA TIENE IMPACTO Y SUFICIENTE RELEVANCIA JURÍDICA Y PUEDE CONSTITUIR UNA DISCRIMINACIÓN ALGORÍTMICA?

4.1.- Los criterios y variables para determinar el impacto, nivel de riesgo y relevancia jurídica de los sistemas algorítmicos públicos

Cabe partir de un concepto inclusivo de sistemas públicos de IA que no se ciña estrictamente al concepto de “IA”, sino que gire sobre el uso público de algorítmicos, esto es, sistemas informáticos que integren fórmulas más o menos complejas y las apliquen a los datos. Todos los sistemas algorítmicos potencialmente pueden generar sesgos y discriminaciones. Con acierto la Ley 1/2022, de 13 de abril valenciana es inclusiva cuando impone la publicidad activa “de sistemas algorítmicos o de inteligencia artificial”, no sólo de estos últimos. Ahora

bien, los riesgos serán mayores conforme el sistema sea más complejo y especialmente debido a la autonomía del sistema propiamente de IA. Y en particular los riesgos se dan con los sistemas predictivos así como los sistemas de autoaprendizaje y aprendizaje profundo, por su falta de explicabilidad, que hará muy difícil justificar los tratos diferentes o demostrar que no hay sesgos relevantes.

Por cuanto a la atracción para el ámbito del Derecho, cabe seguir algunos criterios básicos, algunos de ellos ya están establecidos normativamente y lo que propongo es proyectarlos también para los sistemas IA público y en nuestro caso en relación con los sesgos y la discriminación.

Así, hay que tener en cuenta que el sistema de IA público “produzca efectos jurídicos en él o le afecte significativamente de modo similar” al individuo (art. 22 RGPD). Y acierta también la Ley 1/2022, de 13 de abril valenciana cuando un elemento esencial (para que deba o no hacerse publicidad activa de un sistema) es “que tengan impacto en los procedimientos administrativos o la prestación de los servicios públicos” (art. 16.1 l). Este concepto de “impacto”, tan habitual en las evaluaciones de riesgos y en el contexto de la protección de datos, pasa a ser clave.

Para *valorar el impacto* hay que tener especialmente en cuenta, de un lado, la finalidad del uso o propósito del sistema de IA. Es un buen punto de partida seguir los usos de IA públicos de “alto riesgo” del futuro RIA, a partir de su artículo 6 combinado con los anexos II y III. En esta dirección y para el sector público considero que fundamentalmente serán sistemas de IA de alto riesgo los de seguridad para infraestructuras críticas públicas, el uso de dispositivos de salud, la gestión y priorización de emergencias, el uso de sistemas de identificación biométrica en el contexto de la seguridad pública y fronteras, los usos relacionados con la aplicación policial y judicial de la ley: evaluación de riesgos individuales de delinquir y reincidir, evaluación de perfiles, pruebas y polígrafos, detección estado emocional, análisis de documentación. Asimismo, en el Anexo III del RIA se consideran de alto riesgo usos de IA pública para determinar la admisión y programas en el ámbito de la educación, evaluación y seguimiento de aprendizaje. También los sistemas relativos a los servicios de empleo y evaluación de rendimiento y, más próximo a la actividad administra-

tiva, los sistemas de evaluación, concesión, revocación de prestaciones y servicios esenciales de asistencia pública. Se dice que un tercio de los sistemas de IA públicos serán de alto riesgo⁴⁰. Aunque son de alto riesgo de sesgos y discriminación todos los que están, ni mucho menos están todos los que son. Dejando ahora al margen los usos vinculados a defensa y seguridad nacional, me permito escalar de mayor a menor impacto y riesgo de discriminación, así como de atracción jurídica diversas categorizaciones de usos de IA públicos, a saber:

- Sistemas que identifican o priorizan objetivos para la aplicación de la ley o realizan inspecciones en el ámbito de infracciones penales, administrativas y persecución de ilicitudes y fraude, cada vez más habituales en ámbitos de seguridad, mercado y competencia, trabajo, salud, cuidado del medio ambiente, etc.
- Sistemas para la adjudicación de contratos, subvenciones, privilegios del gobierno.
- Sistemas para personalizar, priorizar o apoyar la prestación de servicios a los ciudadanos en salud, educación, empleo, servicios sociales, etc.
- Sistemas de extracción de información, investigación, recopilación, la supervisión y el análisis de datos, para la elaboración de políticas, la toma de decisiones, monitoreo general y análisis de riesgos.
- Sistemas para la gestión de la organización interna, recursos humanos y las adquisiciones y la gestión de los recursos tecnológicos.

⁴⁰ JOINT RESEARCH CENTRE, TANGI, L. y otros (2022): *AI Watch European landscape on the use of Artificial Intelligence by the Public Sector*, JRC Science For Policy Report, Publications Office of the European Union, Luxembourg, 2022, p. 58 doi:10.2760/39336, <https://publications.jrc.ec.europa.eu/repository/handle/JRC129301>

También se afirma que en general, sólo un 10% de sistemas de IA serán de alto riesgo, COMISIÓN EUROPEA, RENDA A. (2021): *Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe. Final Report (D5)*, abril 2021. <https://op.europa.eu/es/publication-detail/-/publication/55538b70-a638-11eb-9585-01aa75ed71a1>

- Sistemas para interactuar y comunicarse con el público sobre sus derechos y obligaciones y su participación⁴¹.

Señala Huergo que los sistemas de selección de sujetos para ser inspeccionados “quedarían en un segundo plano y con una “irrelevancia jurídica casi total”⁴². Por el contrario aquí se consideran del mayor impacto. Y es que pese a que las normas y jurisprudencia actuales en España parecen darle la razón, es momento de evitar esta huida del Derecho y aplicar las necesarias garantías a estos sistemas automatizados que encienden las luces rojas determinando o guiando la actuación humana. Los falsos positivos y negativos pueden generar tremendas discriminaciones. El caso judicializado más importante hasta la fecha de Países Bajos de SyRI, con mucho acierto determinó la nulidad del sistema que se utilizaba sólo en la fase de selección de alertas.

Para determinar el mayor o menor riesgo del sistema de IA público y la posibilidad de que haya generado o no el sesgo o la discriminación son importantes los *niveles de mayor o menor automatización e intervención o supervisión humana*. Siguiendo los estándares 2016 para el

⁴¹ Elaboración propia, combinando diversas descripciones de usos de IA públicos, en JOINT RESEARCH CENTRE, TANGI, L. y otros (2022), *ob. cit.* antes en JOINT RESEARCH CENTRE, MISURACA, G. (2020), *AI Watch. Artificial Intelligence in public services. Overview of the use and impact of AI in public services in the EU*, Joint Research Centre, Unión Europea. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC120399/jrc120399_misuraca-ai-watch_public-services_30062020_def.pdf De especial interés, en EEUU destaca FREEMAN ENGSTROM, D. y otros (2020): “Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies. Report submitted to the Administrative conference of the United States”, *NYU School of Law, Public Law Research Paper No. 20-54*, 122 págs., en concreto, pp. 21-69, <http://dx.doi.org/10.2139/ssrn.3551505>; HERMSTRÜWER, Y. (2020): “Artificial Intelligence and Administrative Decisions Under Uncertainty”, en WISCHMEYER, T y RADEMACHER, T. (ed.), (2020), *ob. cit.* pp. 199-223. También, VELASCO RICO, C. I. (2022): “Tecnologías disruptivas en la Administración Pública. Inteligencia artificial y Blockchain”, en CERRILLO I MARTÍNEZ, A. (coord.) *La Administración Digital*, Dykinson, Madrid, pp. 228-256, pp. 236-240.

⁴² HUERGO LORA, A. J. (2020): “Una aproximación a los algoritmos desde el derecho administrativo”, en HUERGO LORA, A. J. (dir.), DÍAZ GONZÁLEZ, G. M. (coord.) *La regulación de los algoritmos*, Aranzadi Thomson Reuters, Cizur, 2020, p. 77.

vehículo autónomo⁴³, recientemente Roehl⁴⁴ ha afirmado seis niveles de automatización de la actuación administrativa: “automatización mínima”; “recuperación y tratamiento de datos”; cuando el sistema sugiere los “pasos procedimentales a seguir”; “decisiones asistidas”; “decisiones automatizadas”, hasta el grado “decisiones autónomas” con sistemas dinámicos de machine learning no supervisado.

En este punto cabe recordar que las especiales garantías del artículo 22 RGPD sólo se dan respecto de las decisiones “únicamente” automatizadas del sector público. Igualmente, las escasas garantías del artículo 41. 1º LRJSP se reservan a “cualquier acto o actuación realizada íntegramente a través de medios electrónicos” y que “no haya intervenido de forma directa un empleado público”. Hoy por hoy estos requisitos sirven para una huida del Derecho, esto es, en tanto en cuanto se considera que hay intervención humana en la decisión administrativa que se adopte, no se aplican las especiales garantías. No obstante, el Grupo del artículo 29 ha puesto límites, consciente del seguimiento rutinario por los humanos del sistema automatizado. Para que no rijan estas especiales garantías, la intervención humana ha de ser “significativa, en vez de ser únicamente un gesto simbólico” y llevada a cabo por “persona autorizada y competente”. El estudio de impacto que debe hacerse ha de registrar el grado de intervención humana⁴⁵. La Carta de Derechos Digitales exige el estudio de impacto también en las decisiones “semiautomatizadas” públicas (apartado XVIII. 4º). El futuro RIA para considerar si el sistema es de alto riesgo ha introducido en el artículo 6.3º

⁴³ <https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic>
⁴⁴ ROEHL, U. (2022): “Understanding Automated Decision-Making in the Public Sector: A Classification of Automated, Administrative Decision-Making”, en Juell-Skielse, G. y otros (eds.), *Service Automation in the Public Sector. Progress in IS*. Springer, Cham, pp. 35-63, DOI: 10.1007/978-3-030-92644-1_3, sigo por MORAL SORIANO, L. (2022): “Inteligencia artificial y filosofía del derecho decisiones automatizadas, derecho administrativo y argumentación jurídica”, en LLANO ALONSO F. H. (dir.), *Inteligencia artificial y filosofía del derecho*, Laborun, Madrid, 2022, pp. 475-498, https://idus.us.es/bitstream/handle/11441/137250/Inteligencia%20artificial_Llano%20Alonso.pdf?sequence=1&isAllowed=y

⁴⁵ GRUPO DEL ARTÍCULO 29 (2018): *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*. 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, p. 23.

de la versión de la República Checa que “la salida del sistema no es puramente accesoria con respecto a la acción o decisión pertinente que debe tomarse”. Así pues, hay que valorar si la salida del sistema de IA público, por ejemplo, una propuesta de resolución, puede pasar a ser elemento sustancial para la decisión final que adopte el usuario de su sistema.

Acierta Huergo cuando señala que los mayores riesgos se dan respecto de los sistemas predictivos por cuanto inciden en la acción administrativa añadiendo un “contenido de elaboración propia”⁴⁶. Por el contrario, otros sistemas tienen “un valor meramente auxiliar en la aplicación de la norma”⁴⁷ y cabe ubicarlos en el ámbito de la actuación reglada. Valero en 2007 y Martín Delgado en 2009 hablaban de los sistemas automatizados de “baja intensidad” frente a los de “discrecionalidad política o discrecionalidad técnica no parametrizable”⁴⁸, que afirmaban que no eran posibles en la legislación española.

Sobre estas bases, considero que el peligro del uso de sistemas de IA públicos de mayor a menor pasaría por el uso meramente instrumental (procesadores, traductores, etc.); también meramente sería instrumental el uso para actos reglados, siempre que el sistema se limite a seguir la decisión normativa. Más riesgos se dan en el uso de IA pública en ámbitos de discrecionalidad técnica y más aún en el uso para ámbitos de discrecionalidad política o discrecionalidad técnica no parametrizable. También podrían añadirse a esta escala las decisiones IA administrativas para la aplicación de conceptos jurídicos indeterminados.

Ahora bien, pese a la utilidad de estas escalas de riesgo, hay que insistir que desde el nivel más básico de sistemas automatizados y algorítmicos es más que posible que el sistema genere sesgos, errores

⁴⁶ HUERGO LORA, A. J. (2020), *ob. cit.* ver pp. 65 y ss. en particular 66 y 75 y ss.

⁴⁷ *Ibídem*, p. 65)

⁴⁸ VALERO TORRIJOS, J. (2007): *El régimen jurídico de la e-Administración. El uso de los medios informáticos y telemáticos en el procedimiento administrativo común*, 2.^a ed., Comares, Granada, pp. 74 y 75. Y MARTÍN DELGADO, I. (2009): “Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada”, *Revista de administración pública*, n° 180, pp. 353-386. Ver pp. 368 y ss.

y discriminaciones. Basta una mala selección o calidad de datos o un mero error en un simple algoritmo o fórmula para que quede viciado e impacte en la decisión final que se adopte, pese a que sea formalmente humana.

4.2.- *¿Cuándo el sesgo y el error algorítmico es relevante para constituir una discriminación? Datos estadísticos en discriminaciones indirectas y la necesidad de “recalibrarlos” a la baja para la IA pública*

Debe valorarse en cada supuesto específico si el sesgo o trato diferente que genera el sistema de IA tiene una relevancia o efectos suficientes para considerar que puede constituir una discriminación. Los errores, sesgos o posibles discriminaciones deben ser evaluados y determinarse hasta qué punto son asumibles por la organización⁴⁹. Como recuerda Zlotnik hay que utilizar métricas de impacto y no sólo métricas de error. El impacto no es lo mismo que el error, dado que un 0,1% de errores de que una persona muera, se prorrogue su estancia en la cárcel, sea detenida o que se produzca un incidente en una central nuclear, nada tiene que ver con la posibilidad de que se le conceda erróneamente una ayuda escolar, se le priorice para una cita médica general o se le concedan entradas gratuitas para a una actividad cultural.

No es fácil establecer apriorísticamente los umbrales asumibles o no, pero sí es posible exigir que se hayan analizado, determinado, evaluado y adoptado medidas al respecto. Si estamos en el ámbito de diferencias de trato de las especialmente sospechosas o concurren otros derechos fundamentales en juego, serán menos tolerables los sesgos y errores.

Tiene interés señalar ahora que el juicio de si se da efectivamente una discriminación indirecta -que luego se expone-, se centra precisamente en la determinación de los “efectos sustancialmente más perjudiciales para un grupo protegido”⁵⁰. En nuestro caso, se trata de valorar datos

⁴⁹ ZLOTNIK, A. (2019), *ob. cit.* pp. 27-28.

⁵⁰ STC 145/1991, FJ 2º; STC 128/1987, FJ 6º, ver voto particular a la STC 240/1999.

estadísticos, de la realidad, de las circunstancias concretas del caso en el que el sistema de IA -aparentemente neutro- afecta de manera negativa especialmente a determinados colectivos. El TJUE ha indicado que la diferencia de trato debe ser considerable y que el resultado perjudicial debe tener una incidencia de en torno al 80 o 90% para las personas pertenecientes al grupo protegido. Lo habitual no es centrarse en magnitudes numéricas, sino que “para presumirse discriminatoria, la medida debe afectar a un número *mucho más elevado* de mujeres que de hombres [Rinner-Kühn], o a un porcentaje *considerablemente inferior* de hombres que de mujeres [Nimz, Kowalska] o a un porcentaje *muy superior* de mujeres que de hombres [De Weerd]”⁵¹.

Estos parámetros serían en principio aplicables a los sistemas de IA. Sin embargo, considero que para sistemas de IA públicos que se aplican de forma masiva hay que “recalibrar” estos umbrales y ser mucho menos tolerantes.⁵² Para establecer el umbral de afectación del sesgo o diferencia algorítmica en un caso concreto, no sólo hay que tener en cuenta la afectación al concreto afectado por un sistema de IA, sino que habrá de ponderarse el peligro que supone que ese error o sesgo masivo se replique en miles o millones de decisiones. Además, ha de tenerse en cuenta que si el error no se controla, analiza y en su caso se corrige, las decisiones erróneas pasarán a ser *big data* que alimentará a los futuros algoritmos, haciendo que el sesgo se cronifique. En consecuencia, en vez de considerar como relevantes diferencias de trato “muy elevadas”, “considerablemente inferiores” o “muy superiores”, habrán de bastar tasas de afectación mucho menores. Se seguiría así para el ámbito algorítmico la doctrina Sentencia TJUE 9 de febrero de 1997.⁵³

⁵¹ Sobre el papel de las estadísticas y otros datos FRA- CONSEJO DE EUROPA (2018), *ob. cit.* pp. 275-282. Ahí las referencias a sentencias TJUE mencionadas.

⁵² Se aprovecha la idea del recalibrado de garantías de CITRON, D. K. (2008), *ob.cit.*, pp. 1254 y ss. p. 1286.

⁵³ C-167/97, Regina contra Secretary of State for Employment, ex parte Nicole Seymour-Smith y Laura Perez, párrafo 61. Sigo por SORIANO ARNANZ, A. (2021 a).

5.- DISCRIMINACIONES ALGORÍTMICAS POR ERROR, DIRECTAS, INDIRECTAS, INTERSECCIONALES Y ACCIONES POSITIVAS ALGORÍTMICAS

No procede extenderse ahora en la protección internacional, constitucional y legal de la igualdad y la no discriminación.⁵⁴ En España, la igualdad se ha conformado jurídicamente tanto como un valor superior del ordenamiento jurídico (art. 1. 1º CE), así como un principio y derecho fundamental; también su constitucionalización ha incluido las dos vertientes de la igualdad formal (art. 14 CE) y la igualdad material (art. 9. 2º CE), entre otros preceptos⁵⁵. Baste simplemente adelantar que las leyes o reglamentos en su contenido, así como quienes los aplican, tienen en general prohibido establecer diferencias de trato que no tengan una justificación objetiva y razonable. Y esta prohibición se hace más intensa si las diferencias se basan en circunstancias especialmente sospechosas o prohibidas de discriminación. En estos casos especiales, las diferencias de las normas o de su aplicación se someten a un escrutinio de justificación objetiva y proporcionalidad más riguroso y, de hecho, se presume que sí que son discriminatorias. Puede adelantarse también que en el ámbito del sector público

⁵⁴ Entre otros, destaco FRA- CONSEJO DE EUROPA (2018): *Manual de legislación europea contra la discriminación*, revisión 2018 340 pp., Oficina de Publicaciones de la Unión Europea, https://www.echr.coe.int/documents/handbook_non_discrim_law_spa.pdf; GONZÁLEZ-TREVIJANO, P. (2020): *Los principios de igualdad y no discriminación, una perspectiva de Derecho Comparado. España*, EPRS Servicio de Estudios del Parlamento Europeo, PE 659.297, octubre, [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/659297/EPRS_STU\(2020\)659297_ES.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/659297/EPRS_STU(2020)659297_ES.pdf); RUBIO LLORENTE, F. (1993): "Igualdad en la jurisprudencia del Tribunal Constitucional. Introducción", en *La Forma del Poder (Estudios sobre la Constitución)*, CEC, Madrid, pp. 637-684; GIMÉNEZ GLUCK, D. (2004): *Juicio de igualdad y Tribunal Constitucional*, J.M. Bosch Editor, Barcelona; REY MARTÍNEZ, F. (2017): *Igualdad y prohibición de discriminación: de 1978 a 2018*, *Revista de derecho político*, Nº 100, 2017, pp. 125-171 y REY MARTÍNEZ, F. (2020): *Derecho antidiscriminatorio*, Aranzadi Thomson Reuters, Cizur.

⁵⁵ No agotan estos preceptos la recepción expresa de la igualdad en la Constitución, así, el artículo 23. 2º CE supone una concreta manifestación de la igualdad ante la ley y de la igualdad política. De otra parte, los artículos 139 y 149. 1.º. 1ª CE vienen a reconocer el carácter estatal de la igualdad en el marco de la descentralización autonómica.

el control de la igualdad queda muy ligado y se superpone con el del cumplimiento de otros mandatos constitucionales: la actividad administrativa no puede ser arbitraria, ha de estar motivada con suficiencia y racionalidad jurídica (arts. 9. 3 y 24 CE), ha de someterse a la legalidad y al mandato de objetividad (art. 103 CE) y el control incluirá que no se dé un desvío de poder (art. 106 CE). Como se verá, esta superposición se da más en los supuestos generales de igualdad en la aplicación de la ley, no en los especialmente sospechosos de discriminación.

A continuación se proyectan los variados conceptos y categorías básicas de la igualdad respecto del uso público de algoritmos. Y para ello es conveniente apoyarse en las recientes ente definiciones del artículo 6 Ley 15/2022.

5.1.- La discriminación o sesgo algoritmo será en la mayoría de supuestos “discriminación por error”

Aunque no sea un concepto asentado en el Derecho antidiscriminatorio, resulta especialmente útil acudir a la noción de “discriminación por error”, esto es, “aquella que se funda en una apreciación incorrecta acerca de las características de la persona o personas discriminadas” (art. 6. 2º b Ley 15/2022). Este concepto permite subsumir prácticamente todos los supuestos en los que el sistema de IA público genera diferencias de trato por error, sin que se haya diseñado para establecer estas diferencias. Cabe adelantar que en estos casos no habrá de aplicarse el test de igualdad, sino que el análisis jurídico se centrará en probar la existencia del sesgo o discriminación, determinar si es relevante y qué consecuencias tiene y si se han cumplido las medidas preventivas para evitarlo.

5.2.- Discriminaciones algorítmicas directas, interseccionales y por asociación

Las discriminaciones directas (art. 6. 1º Ley Orgánica 3/2007, de 22 de marzo y art. 6. 1º Ley 15/2022)⁵⁶ se dan cuando una persona

⁵⁶ FRA- CONSEJO DE EUROPA (2018), *ob. cit.* pp. 47-58.

recibe un trato menos favorable por alguna de las circunstancias especialmente sospechosas de discriminación. En el caso de un sistema de IA tendría lugar si se introducen en el sistema de IA datos sobre la pertenencia a un colectivo especialmente discriminado y a dicha pertenencia se asocia un factor negativo, o si los algoritmos y variables se diseñan para perjudicar a estos colectivos.

Ya se trate de unas u otras discriminaciones, los sistemas algorítmicos públicos también pueden generar discriminación múltiple o interseccional (art. 6. 3º) en tanto se acumulen errores o tratamientos diferenciados basados en dos o más circunstancias especialmente prohibidas (por ejemplo, mujeres musulmanas). Ello no será muy extraño dada la ingente cantidad de datos y perfiles que se manejan.

La categoría de la “discriminación por asociación” (art. 6. 2º a) Ley 15/2022) puede ser útil para abordar las diferenciaciones algorítmicas a partir de datos de aproximación, características sociales o *proxies*⁵⁷. Como señala Soriano esta categoría “abre la puerta a la posibilidad de reconocer como discriminación algorítmica (directa o indirecta)”⁵⁸ estos supuestos en los que se tienen en cuenta “afinidades” étnicas, multiculturales.

5.3.- *La complejidad de las discriminaciones algorítmicas indirectas*

Es bastante habitual acudir a la categoría de las discriminaciones indirectas en el caso de algoritmos, aunque la cuestión en modo alguno resulta clara. Cabe recordar que se da una discriminación indirecta “cuando una disposición, criterio o práctica aparentemente neutros ocasiona o puede ocasionar a una o varias personas una desventaja particular con respecto a otras por razón de las causas” de las especialmente sospechosas. (art. 6. 2º Ley orgánica 3/2007, de 22 de

⁵⁷ Cabe recordar que se define “discriminación por asociación cuando una persona o grupo en que se integra, debido a su relación con otra sobre la que concurra alguna de las causas” especialmente sospechosas, art. 6. 2 a) Ley 15/2022. Al respecto de este tipo de discriminación FRA- CONSEJO DE EUROPA (2018), *ob. cit.* pp. 56-58.

⁵⁸ SORIANO ARNANZ, A. (2021 a), *ob. cit.*

marzo y art. 6. 1º b) Ley 15/2022)⁵⁹. En nuestro caso, se acude a esta categoría por cuanto lo “aparentemente neutral” sería el sistema de IA en su conjunto, dado que está diseñado sobre la base de datos, criterios o variables como mínimo aparentemente objetivas⁶⁰. Lo normal es que el sistema de IA no se haya diseñado para diferenciar, pero se hayan generado involuntariamente sesgos que dan lugar a una discriminación por error. Es por ello que el análisis jurídico en las discriminaciones indirectas acaba siendo bastante similar al caso de la discriminación por error, centrado esencialmente en determinar los efectos del sistema y el umbral de admisión de los sesgos, errores y diferencias que genera. No tiene sentido aplicar el test general de discriminación, puesto que no se trata de justificar si la diferenciación que produce el sistema está justificada y es razonable y proporcional⁶¹.

Aunque cae a medio caballo entre las discriminaciones directas e indirectas, puede ser útil acudir a las discriminaciones indirectas respecto de problemas del uso de *proxies*, como se verá *infra*, datos o variables que son “aparentemente neutros”, pero de los que el sistema de IA acaba derivando un tratamiento perjudicial para colectivos especialmente discriminados. En estos supuestos también se puede acudir al concepto de discriminación por asociación algorítmica y para Soriano estos supuestos como abordables como discriminación directa⁶². A mi juicio, lo importante es apreciar que estos datos *proxie* que se manejan son afines a datos de colectivos especialmente discriminados, o bien, que es el sistema de IA al tratar con *proxies* el que produce los efectos diferentes para los colectivos especialmente

⁵⁹ FRA- CONSEJO DE EUROPA (2018), *ob. cit.* pp. 58-77.

⁶⁰ Afirma SORIANO ARNANZ, A. (2021 a), *ob. cit.* que una discriminación algorítmica indirecta se da por “la disposición, práctica o criterio aparentemente neutro puede ser la variable específica introducida en el sistema que, por el valor que le otorga este, genera un resultado discriminatorio para el grupo desaventajado, o bien el algoritmo generalmente considerado”. Por cuanto a la variable específica será necesario acceder al sistema y la complejidad del mismo y de los datos que procesa dificultará especialmente el conocimiento.

⁶¹ SORIANO *ibídem*, acude erróneamente a mi juicio a este test en las discriminaciones algorítmicas indirectas, si bien, acaba reconduciendo su análisis a la necesidad de analizar la precisión del sistema.

⁶² SORIANO ARNANZ, A. (2021 b) y (2022 a).

protegidos. Si es así, se han de aplicar en lo posible las garantías de las discriminaciones especialmente prohibidas que luego se siguen.

También, en principio, se estaría propiamente en el ámbito de la discriminación algorítmica indirecta cuando el sistema de IA público no diferencia en sus efectos y resultados a los colectivos especialmente discriminados, es decir, cuando arroja resultados iguales para categorías diferentes⁶³. Este tratamiento algorítmico aparentemente neutral puede ser inaceptable. Este supuesto sería similar a los que se abordan en el siguiente apartado y deben ser analizados bajo el test de la igualdad y no tanto en el marco de las discriminaciones indirectas.

5.4.- Acciones positivas algorítmicas, deber de diferenciación y modulaciones y compensaciones para la minimización de la discriminación

El mandato constitucional de hacer efectiva la igualdad (art. 9.2º CE) ha de proyectarse también en la acción y la transformación digital del sector público. Entre otros, queda recogido en la Carta de Derechos Digitales, especialmente en su apartado VIII, así como los IX y XII respecto las luchas contra las brechas digitales, que incluye también “la promoción de “políticas públicas específicas dirigidas a abordar las brechas de acceso atendiendo a posibles sesgos discriminatorios” (2º)⁶⁴. También como proyección del artículo 9.2º CE, para asegurar la igualdad de todas las personas en el disfrute de los derechos, las administraciones, los empleadores y los prestadores de servicios deben adoptar medidas especiales o específicas para adaptar sus disposiciones y prácticas a las personas con características diferentes, denominadas generalmente como acciones positivas⁶⁵.

⁶³ En sentido contrario a lo señalado por SORIANO *ibidem*.

⁶⁴ Al respecto, ABA CATOIRA, A. (2022), “Derechos de igualdad, personas con discapacidad y mayores en el entorno digital (VIII, XI Y XII)”, COTINO HUESO, Lorenzo (editor), *La Carta de Derechos Digitales*, Tirant Lo Blanch, Valencia, pp. 123-154.

⁶⁵ Sobre el tema, entre otros, FRA- CONSEJO DE EUROPA (2018), *ob. cit.* pp. 78-91. La bibliografía es muy abundante, además de obras generales mencionadas, BALAGUER CALLEJÓN, M. L. (2018): “Acciones positivas para la igualdad efectiva de mujeres y hombres”, VENTURA FRANCH, A. y GARCÍA

Asimismo, aunque es un tema conflictivo, el TEDH afirma desde 2000 que se vulnera de derecho de igualdad “cuando [...] los Estados no aplican un trato diferente a personas cuyas situaciones son sensiblemente diferentes”⁶⁶ y puede ser obligatoria una regulación o trato diferente en algún caso concreto para salvar el derecho de igualdad⁶⁷.

En consecuencia, el sistema de IA público en ocasiones habrá de diseñarse para instrumentar una medida especial o acción positiva, esto es, con la intencionada finalidad de beneficiar a determinados grupos sociales que están en situación desfavorable respecto de la media de la población. Se trata de medidas algorítmicas “orientadas a prevenir, eliminar y, en su caso, compensar cualquier forma de discriminación o desventaja en su dimensión colectiva o social” (art. 6.7º Ley 15/2022, también, art. 11 Ley Orgánica 3/2007, de 22 de marzo), al amparo del artículo 9.2 CE o el artículo 23 de la Carta de derechos fundamentales de la UE. Si el sistema y su funcionamiento está regulado, habrá que analizar la concreta disposición y la justificación objetiva razonable de la medida de acción positiva y su concreta proporcionalidad. No obstante y en especial, habrá que tener en cuenta las medidas concretas efectivamente adoptadas en el sistema de IA. En estos supuestos se modulará el funcionamiento del sistema y se introducirán los criterios y variables para preferenciar a los colectivos determinados, o perjudicar al resto, según se mire. Habrán de reco-

CAMPÁ, S. (coord.) *El derecho a la igualdad efectiva de mujeres y hombres: una evaluación del primer decenio de la Ley Orgánica 3/2007*, pp. 285-312; GIMÉNEZ GLUCK, D. *Una manifestación polémica del principio de igualdad: acciones positivas moderadas y medidas de discriminación inversa*, Tirant lo Blanch, Valencia, 1999; GONZÁLEZ-TREVIJANO, P. (2020), *ob. cit.* pp. 75-77.

⁶⁶ Especialmente importante fue la sentencia TEDH, Thlimmenos contra Grecia [GS], n.o 34369/97, 6 de abril de 2000, apartado 44. De manera análoga, TEDH, Pretty contra Reino Unido, n.o 2346/02, 29 de abril de 2002, apartado 88.

⁶⁷ TEDH, Biao contra Dinamarca [GS], n.o 38590/10, 24 de mayo de 2016, apartado 103; TEDH, D. H. y otros contra República Checa [GS], n.o 57325/00, 13 de noviembre de 2007, apartado 184. Ciertamente, nuestro TC siempre ha afirmado lo contrario, esto es, que no hay un derecho a la diferenciación (por todas, sentencia 181/2000 (FJ 14º) o STC 128/2004, de 21 de julio). Sobre el tema, de escasa atención, COBREROS MENDAZONA, E. (2997): “Discriminación por indiferenciación: estudio y propuesta”, *Revista española de derecho constitucional*, Año nº 27, Nº 81, 2007, págs. 71-114.

gerse las evidencias relativas a los datos de entrenamiento del sistema y las variables utilizadas, así como los mecanismos de evaluación del funcionamiento del sistema con los criterios correctores introducidos. Sobre estas evidencias, será posible analizar las medidas respecto del sistema de IA público y valorar si se han adoptado conforme a la normativa -si es que la hay- y ya en particular, la justificación objetiva⁶⁸. Ciertamente se trata de evaluar la legitimidad de los objetivos seguidos y la proporcionalidad de los medios utilizados por el sistema de IA para alcanzarlos.

Cabe señalar que es un análisis similar al que hay que realizar respecto de las correcciones, medidas de compensación, modulación o mitigación que se han de adoptar precisamente en el sistema de IA cuando se detecta la discriminación y el sesgo algorítmicos y hasta dónde hacerlo.

6.- ¿CUÁNDO UN SISTEMA ALGORÍTMICO PÚBLICO QUE NO TIENE EN CUENTA CIRCUNSTANCIAS ESPECIALMENTE SOSPECHOSAS ES DISCRIMINATORIO?

6.1.- *La igualdad “en” la ley o norma que habilite el uso del sistema de IA y el control laxo de la posibilidad de establecer diferencias de tratos razonables y no arbitrarios*

La igualdad formal pasó a incluir también la igualdad *en* la ley, esto es, la necesidad de que las normas respeten también en su contenido el principio de igualdad⁶⁹. La acción normativa legal o reglamentaria implica otorgar -o no- el mismo tratamiento a grupos o colectivos definidos por la norma. Bajo esta perspectiva, la

⁶⁸ Sobre la justificación FRA- CONSEJO DE EUROPA (2018), *ob. cit.* p. 103-125.

⁶⁹ La STC 144/1988, de 12 de junio. En su FJ 1º se afirma que “frente al legislador o frente al poder reglamentario, impidiendo que uno u otro puedan configurar los supuestos de hecho de la norma de modo tal que se dé trato distinto a personas”. Diversas referencias para el TC de la igualdad en la ley, GONZÁLEZ-TREVIJANO, P. (2020), *ob. cit.* p. 55-58.

función de las normas es atribuir efectos jurídicos diferentes a las situaciones jurídicas que crean. El diseño, implantación y uso del sistema de IA público queda sometido al principio de legalidad de la Administración, de modo más o menos intenso según diversas variables que se han expuesto *supra*. Sobre esta base, es fundamental tener en cuenta la cobertura legal o reglamentaria del sistema de IA y de la posibilidad de que dicha normativa habilite a que el sistema de IA genere diferencias de trato. Así, por ejemplo, la norma puede determinar requisitos para una contratación, una prestación, una subvención y, en consecuencia, el diseño, configuración y datos del sistema de IA quedarán más o menos recogidos en la norma. Si es así, habrá que llevar a cabo un análisis de igualdad *en* dicha norma. En este sentido, la igualdad opera “frente al legislador o frente al poder reglamentario, impidiendo que uno u otro puedan configurar los supuestos de hecho de la norma de modo tal que se dé trato distinto a personas que, desde todos los puntos de vista legítimamente adoptables, se encuentran en la misma situación”. Así pues, no hay que dar “relevancia jurídica a circunstancias que, o bien no pueden ser jamás tomadas en consideración por prohibirlo así expresamente la propia Constitución, o bien no guardan relación alguna con el sentido de la regulación que, al incluirlas, incurre en arbitrariedad y es por eso discriminatoria” (STC 144/1988, de 12 de junio FJ 1º)⁷⁰. La igualdad -y también la igualdad en la ley- no sólo se predica respecto de las discriminaciones especialmente sospechosas. Ahora bien, lo cierto es que salvo que se trate de discriminaciones prohibidas o vinculadas con otro derecho fundamental, el análisis de admisibilidad se hace muy laxo en general, limitado a la razonabilidad, no arbitrariedad y existencia de justificación objetiva. En muchos casos basta con que el establecimiento de la diferencia tenga una finalidad y ésta sea razonable, básicamente que no sea contraria al sistema de valores

⁷⁰ Y, “para que las diferenciaciones normativas puedan considerarse no discriminatorias resulta indispensable que exista una justificación objetiva y razonable, de acuerdo con criterios y juicios de valor generalmente aceptados, cuya exigencia deba aplicarse en relación con la finalidad y efectos de la medida considerada, debiendo estar presente, por ello, una razonable relación de proporcionalidad entre los medios empleados y la finalidad perseguida” STC 46/1999, de 22 de marzo FJ 2º. También clara STC 75/1983, de 18 de febrero”.

constitucionales. En principio es necesario que el trato sea congruente y proporcionado⁷¹, si bien desaparece -o casi- el análisis de la proporcionalidad. Los términos del análisis de la posible discriminación quedan abiertos al decisionismo judicial y a los criterios políticos⁷². También es cierto que en el ámbito de la potestad reglamentaria y frente a la arbitrariedad se van incrementando las necesidades de justificar la finalidad y contenido de la norma reglamentaria, esencialmente en el procedimiento de elaboración de la misma⁷³. Por el contrario, el análisis sí será riguroso respecto de las normas que habiliten a que un sistema de IA pueda establecer diferencias de trato por motivos especialmente sospechosos de discriminación.

⁷¹ “La diferencia de trato a personas en situaciones sustancialmente similares [...] es discriminatoria si no tiene una justificación objetiva y razonable; en otras palabras, si no persigue un fin legítimo o no hay una relación de proporcionalidad razonable entre los medios empleados y el fin perseguido” (TEDH, Burden contra Reino Unido, 29 de abril de 2008, ap. 60). FRA- CONSEJO DE EUROPA (2018), ob. cit. pp. 103-108 y ss. Así desde la STEDH de 23 de julio de 1968 (caso régimen lingüístico belga), asumida por el TC en el FJ 3º de la pronta STC 22/1981, de 2 de julio.

⁷² GIMÉNEZ GLUCK, D. (1999), ob. cit. p. 34 ha llegado a afirmar en esta dirección que “En materia de constitucionalidad de la diferencia en el trato, las decisiones son principalmente políticas, aunque se revistan de un entramado jurídico justificativo de la opción política que se adopta. No es menos cierto que esto es lo que ocurre con la mayoría de las decisiones denominadas jurídicas, pero en el terreno de la igualdad esta dinámica general se hace mucho más evidente.” (Cfr. *Una manifestación polémica del principio de igualdad...* cit. pág. 34). “La discrecionalidad de las decisiones del Tribunal, que en general ya puede ser considerada excesiva, alcanza en la determinación de la constitucionalidad del trato desigual un grado ilimitado. No obstante, la construcción jurisprudencial sobre la justificación objetiva y razonable otorga, al menos, un marco jurídico a estas decisiones políticas” (pp. 38 -39).

⁷³ Así, STS de 3 de noviembre de 2010, recurso 480/2009 y la más reciente STS de 4 de abril de 2022 (rec.300/2020), <https://delajusticia.com/wp-content/uploads/2022/04/STSregla.pdf>

6.2.- *Control de discriminación y arbitrariedad y exigencias de motivación en los casos de aplicación algorítmica de la ley. Decisiones automatizadas discrecionales y supuestos no especialmente sospechosos*

La ley o reglamento es muy posible que no establezcan ninguna diferencia de trato, pero que sea el sistema de IA el que la produce. En estos supuestos, ya no se trata de una cuestión de igualdad *en* la ley, sino de igualdad *ante* la ley o en la aplicación de la misma. La Administración y los jueces deben aplicar la ley por igual a quienes estén en la misma situación. La igualdad no exige un estricto trato igual, sino que exige que el diferente trato no sea discriminatorio. Pues bien, ello mismo debe predicarse de los sistemas IA públicos que tengan efectos jurídicos sobre las personas. Si el sistema de IA y la diferencia de trato cuenta con cobertura normativa, ya habrá tenido que realizarse en abstracto el análisis de la posible discriminación *en* la ley.

A partir de ahí, habrá que apreciar si las diferencias de trato que genera el sistema son relevantes en los términos ya vistos. Si es así, procede distinguir *si el trato diferenciado se genera por sesgos y errores del sistema*. Si es así, por lo que *se trataría de una discriminación por error*. En estos casos, el debate jurídico se centrará en una cuestión de prueba de la existencia del sesgo o discriminación, de umbrales de precisión admisibles, del cumplimiento de las medidas preventivas en el diseño, implantación y evaluación del sistema y de las consecuencias del sesgo o discriminación. No habrá de aplicarse el juicio de igualdad: si el sesgo, error o trato diferenciado es relevante y no cuenta con cobertura normativa o el sistema se ha desviado de la norma, no podrá considerarse justificado, razonable y no arbitrario.

Otra situación es que *las normas no hayan previsto la diferencia de trato, pero en el diseño e implantación del sistema se han adoptado conscientemente decisiones para que el sistema de IA genere tratos diferentes*. También puede darse el caso de que *el sistema autónomo, especialmente de aprendizaje automático, funciona generando diferencias de trato no previstas*. En estos casos, habrá que hacer un análisis del concreto trato diferente que depara el sistema, su posible finalidad objetiva y los efectos reales del sistema de IA. Así, no sólo habría que estar a las finalidades perseguidas o declaradas por los desarrolladores, sino a los resultados efectivos del sistema. Y a partir de ahí, para

que se considere que el sistema no es discriminatorio es preciso que la finalidad del sistema y sus resultados puedan considerarse razonables y no contrarios al sistema de valores, a la Constitución, así como al marco legal de la actuación administrativa, así como congruentes y proporcionados. Como se adelantó, el juicio de igualdad es más laxo cuando no se trata de una discriminación especialmente prohibida (por ejemplo, que el sistema preferencia quienes estudiaron en colegios privados). Ahora bien, en estos supuestos *el análisis ha de concurrir llegando a fusionarse con el de no arbitrariedad, de motivación con suficiencia y racionalidad jurídica* (arts. 9. 3 y 24 CE) exigibles a la actuación administrativa.

En los supuestos en los que los sistemas automatizados más simples se emplean en decisiones regladas, el análisis será más sencillo y simplemente se limitaría a apreciar si el sistema y los algoritmos se ajustan completamente a la norma que lo regula.

Las *curvas* se dan especialmente si se usa un sistema de IA para *decisiones y ejercicio de potestades discrecionales*⁷⁴. Es en estos ámbitos donde el uso de IA público más puede aportar para lograr mayor neutralidad y objetividad, reducir la discriminación y la discrecionalidad y los ineludibles sesgos y opacidad humanos, además de lograr mayor justicia con la personalización de individualización de los casos⁷⁵. De hecho, a mi juicio, no tiene sentido una prohibición general del uso de IA respecto de potestades discrecionales, como se interpreta habitualmente la actual normativa alemana. Eso sí, el uso en estos contextos discrecionales tiene un peligro potencial mayor y, en consecuencia, mayores deben ser las garantías que deben darse. Es así como debe interpretarse el derecho de la Carta (XVIII. 6º d): “Que la adopción de decisiones discrecionales quede reservada a personas, salvo que normativamente se prevea la adopción de decisiones automatizadas

⁷⁴ Se sigue a continuación una proyección de los conceptos y garantías exigibles al respecto en GAMERO, E. y FERNÁNDEZ, S. (2016), *ob. cit.* pp. 81-84.

⁷⁵ En esta línea, MINISTERIE VAN BINNENLANDSE ZAKEN (2021), *ob. cit.* p. 11. TISCHBIREK, A. (2020), *ob. cit.* pp. 103 y ss.; HERMSTRÜWER (2020), *ob. cit.* pp. 199 y ss. En cierto modo, HUERGO LORA, A. J. (2021), *ob. cit.* Y especialmente habría que acudir a la excelente reflexión global de BEN- SHAHAR, O. y PORAT, A. (2021): *Personalized law : different rules for different people*, Oxford University Press.

con garantías adecuadas.” Y precisamente la labor normativa debe ser más intensa en estos casos. En las decisiones discrecionales la Administración y el sistema de IA pueden optar entre diversas soluciones hacia la más favorable al interés general. En estos casos se ha de poder comprobar que el sistema de IA no es arbitrario y el trato diferente tiene justificación objetiva y no es irrazonable. El sistema de IA habrá de garantizar que sus decisiones están claramente respaldadas por los hechos determinantes sobre los que incide la decisión administrativa. Y para ello la importancia de la motivación administrativa es esencial, y para un sistema de IA este examen pasa por la suficiente transparencia y explicabilidad del sistema, especialmente relacionada con la motivación⁷⁶. Como he analizado⁷⁷, acierta la Carta de Derechos digitales al reconocer el derecho a “Obtener una motivación comprensible en lenguaje natural de las decisiones que se adopten en el entorno digital, con justificación de las normas jurídicas relevantes, tecnología empleada, así como de los criterios de aplicación de las mismas al caso” (XVIII. 6º c). Hay que eludir, por ello, una presunción jurídica de acierto de las decisiones automatizadas o de sistemas inteligentes y ofrecer posibilidades y garantías eficaces —no sólo en la vía judicial— de probar los posibles errores de diseño o funciona-

⁷⁶ Al respecto, una mayor atención en HERMSTRÜWER, Y. (2020), *ob. cit.* ; GAMERO CASADO, E. (2021): “Necesidad de motivación e invalidez de los actos administrativos sustentados en inteligencia artificial o en algoritmos”, *Almacén de Derecho*, 4 febrero, <https://almacendederecho.org/necesidad-de-motivacion-e-invalidez-de-los-actos-administrativos-sustentados-en-inteligencia-artificial-o-en-algoritmos> Desde la perspectiva de la discrecionalidad, HUERGO LORA, A. J. (2021): “Administraciones Públicas e inteligencia artificial: ¿más o menos discrecionalidad?”, *El Cronista del Estado Social y Democrático de Derecho*, nº 96-97, 2021, pp. 78-95, sigo por <https://laadministracionaldia.inap.es/noticia.asp?id=1512494> ; CERRILLO MARTÍNEZ, A. (2020): “¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?”, *European Review of Digital Administration & Law* Volume 1 – Issue 1-2 – 2020, DOI: 10.4399/97888255389603, págs. 18-36, ver 23-26, y 29 y ss.. Con relación a la revisión judicial ulterior, en especial, DE LA SIERRA MORÓN, S. (2020): “Inteligencia artificial y justicia administrativa: una aproximación desde la teoría del control de la Administración Pública”, *Revista General de Derecho Administrativo*, nº. 53, 2020.

⁷⁷ COTINO HUESO, L. (2022): “Derechos ante la administración digital y la inteligencia artificial (XVIII Y XV)”, en COTINO HUESO, Lorenzo (editor) (2022) *ob. cit.* pp. 251-284.

miento. En el caso de la discrecionalidad técnica, será necesario que se conozcan las fuentes de información sobre las que el sistema de IA realiza el juicio técnico, los criterios de valoración cualitativa que el sistema ha empleado, y que el resultado de salida del sistema permita motivar la aplicación de tales criterios al resultado individualizado que ha preferenciado una decisión o solución. El sistema de IA debe poder hacerlo, así como generar evidencias de cómo lo hace al tiempo de motivar su decisión.

Estas obligaciones de motivación están muy ligadas a la justificación objetiva atenuada, que es lo que esencialmente se exige en los supuestos de discriminaciones no especialmente prohibidas⁷⁸. En principio, habría de haber cierta tolerancia de errores en estos supuestos de uso de IA en ámbitos de discrecionalidad técnica.

Propuse en la fase inicial de la Carta⁷⁹ -y afortunadamente se ha incluido- la garantía de que “El interesado tendrá derecho a que se motive o se explique la decisión administrativa cuando esta se separe del criterio propuesto por un sistema automatizado o inteligente” (XVIII. 6º c). El interesado debe poder conocer las razones humanas que imponen a no seguir las propuestas o decisiones automatizadas, se trata de una extensión necesaria precisamente del control o supervisión humana.

Las exigencias al sistema de IA serán bastante similares cuando se utilicen para la interpretación de conceptos jurídicos indeterminados en supuestos concretos por la Administración. En estos casos, para actuar conforme a Derecho se da un control de racionalidad y razonabilidad. Es por ello que el sistema de IA que se aplique debe permitir que se controle si la Administración ha respetado la realidad de los hechos, si ha considerado todos los factores relevantes y si en sus criterios y preferencias ha seguido el valor que el ordenamiento confiere a estos factores. Respecto de la transparencia y explicabilidad me remito al trabajo en esta obra, así como a recientes aportaciones.⁸⁰

⁷⁸ FRA- CONSEJO DE EUROPA (2018), ob. cit. pp. 103-108 y ss.

⁷⁹ Véase mi aportación en esta misma dirección en RED DAIA (2020 a): *Aportación a consulta de derechos digitales de julio, 2020* <https://acortar.link/q1wl3i>

⁸⁰ (2023): “Qué concreta transparencia e información de algoritmos e inteligencia artificial es la debida”, Revista Española de Transparencia y “Transparencia y explicabilidad de la inteligencia artificial y “compañía” (comunicación, interpre-

7.- ¿CUÁNDO UN SISTEMA ALGORÍTMICO PÚBLICO QUE SÍ TIENE EN CUENTA CIRCUNSTANCIAS ESPECIALMENTE SOSPECHOSAS ES DISCRIMINATORIO?

7.1.- Las garantías frente a las discriminaciones algorítmicas especialmente prohibidas se superponen por lo general a las garantías de los tratamientos de datos especialmente protegidos

El derecho a la igualdad excluye toda clase de discriminación, si bien hay una serie de rasgos diferenciadores que por la experiencia de la Historia y el presente son especialmente sospechosos de motivar una discriminación. El artículo 14 CE así como su interpretación conforme a los tratados y jurisprudencia internacionales parecen a garantizar la igualdad en general y a incluir un mandato antidiscriminatorio específico, una prohibición específica de discriminar “por razón de nacimiento, origen racial o étnico, sexo, religión, convicción u opinión, edad, discapacidad, orientación o identidad sexual, expresión de género, enfermedad o condición de salud, estado serológico y/o predisposición genética a sufrir patologías y trastornos, lengua, situación socioeconómica, o cualquier otra condición o circunstancia personal o social”. De igual modo, cabe tener en cuenta las diferencias de trato en razón de la “nacionalidad, de si son menores o mayores de edad o de si disfrutan o no de residencia legal” (art. 2.1º Ley 15/2022), así como el régimen más específico de los tratos diferentes en razón de enfermedad (art. 2. 3º). En la normativa europea, las características protegidas se limitan expresamente al sexo, el origen racial o étnico, la edad, la discapacidad, la religión o las creencias, y

tabilidad, inteligibilidad, auditabilidad, testabilidad, comprobabilidad, simulabilidad...). Para qué, para quién y cuánta”, en COTINO HUESO L. Y CASTELLANOS J. (coords.), *Transparencia y explicabilidad de la inteligencia artificial*, Tirant lo Blanch, Valencia, 2022. Una clara visión de conjunto de interés en JIMÉNEZ LÓPEZ, Jesús (2022): “Oscuridad algorítmica en el sector público”, en VESTRI, G. (2022): *Disrupción tecnológica en la administración pública retos*, Aranzadi, Cizur.

la orientación sexual⁸¹. Pues bien, cuando se dan estas circunstancias, se incrementan y mucho las garantías frente a la no discriminación. Y lo que es más, en muchos casos a estas garantías hay que añadir las -muchas- de protección de datos en general y de los datos especialmente protegidos en particular. La superposición de los dos regímenes lleva a una solución bastante más garantista que la que sólo otorga el artículo 14 CE, especialmente para el sector público.

7.2.- En la mayoría de casos será necesaria una ley de calidad y con garantías que habilite a que un sistema de IA pueda diferenciar en razón de datos especialmente sospechosos de discriminación. Y no la hay

En primer lugar, hay que analizar si la norma que habilita a la actuación administrativa y al uso del sistema de IA recoge expresamente la posibilidad de tener en cuenta circunstancias especialmente sospechosas y en qué medida (datos de recogida, variables a tener en cuenta por el sistema, etc.). Como se vio, las exigencias del principio de legalidad administrativa en principio son más intensas debido al impacto en derechos. Por ello, el punto de partida es que la norma sí que debería recoger expresamente esta posibilidad. De hecho, el artículo 2.2º Ley 15/2022 parece exigir “norma con rango de ley, o [...] disposiciones normativas o decisiones generales de las administraciones públicas.” Pero es más, en razón de la aplicación de la normativa de protección de datos en muchos supuestos sí que es obligatoria una norma con rango de ley -y no reglamentaria-. Además, no se tratará de la general base legal que se exige para que sector público trate datos personales (y no el consentimiento, Informe 175/2018 AEPD). Y es que el tratamiento de datos relativos a categorías especialmente sospechosas de discriminación coincide mucho con el tratamiento de datos especialmente protegidos del artículo 9 RGPD⁸². Además, en diversos casos confluirá

⁸¹ FRA- CONSEJO DE EUROPA (2018), *ob. cit.* p. 182, ver pp. 177-260.

⁸² “origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, y el tratamiento de datos genéticos, datos biométricos dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o datos relativos a la vida sexual o las orientación sexuales”.

con el artículo 22 RGPD para decisiones automatizadas, que también impone ley⁸³. A la vista del antedicho art. 2.1º Ley 15/2022, únicamente *no* estaríamos en supuestos de tratamientos de datos especialmente protegidos cuando el sistema de IA trate datos de “nacimiento”, género, “edad”, “lengua” o “situación socioeconómica” y en algunos de estos casos, como la edad o el género no sería difícil que se exijan especiales garantías de protección de datos. Cuando sí estemos bajo el régimen especial del artículo 9 RGPD no basta cualquier ley, sino que la misma ha de contar con una especial calidad según STC 76/2019, de 22 de mayo, en especial FJ 8º y, por supuesto, con las exigencias del artículo 23 RGPD. En el contenido de estas leyes claramente se tiene que habilitar a la Administración y legitimar el tratamiento concreto del sistema de IA y además justificar el mismo. Esta justificación habrá de conectarse por lo general, en aras de un “interés público esencial” (art. 9. 2º g) RGPD). Asimismo, la jurisprudencia constitucional ha sido muy rigurosa respecto de la necesidad de que la ley integre en su contenido no sólo el detalle de la restricción y sus presupuestos, sino que también se han de regular las garantías concretas *compensatorias* para estos tratamientos de datos especialmente protegidos (y sospechosos de discriminación).

Resulta ciertamente difícil pensar que algún sistema público de IA en España actualmente cumpla con estos requisitos. Lo más cercano sería la Ley 22/2018, de 6 de noviembre, de la Generalitat de la Comunidad Valenciana que reguló bastante concienzudamente un sistema de alertas tempranas que en principio no manejaba datos sensibles y lo cierto es que el desconocido informe de la Informe AEPD 385661/2017 de marzo sobre el mismo no consideró suficiente la regulación⁸⁴. Como excepción, es bastante positivo el artículo 10.5º (y Considerando 44) RIA, que habilita al uso de datos especialmente

⁸³ La conexión de estos regímenes se aborda en cierto modo Consejo de Europa (2018), *op.cit.*, p. 38, así como SORIANO ARNANZ, A. (2021 a).

⁸⁴ Al respecto dediqué mi estudio COTINO HUESO, L. (2021). Ver AEPD (2017), *Informe AEPD 385661/2017 de marzo relativo al anteproyecto de ley de sistema de alertas tempranas de la Comunidad Valenciana*. No publicado. He dispuesto un enlace en <https://acortar.link/AJ6mYz>

Asimismo, el 29 de julio de 2020 ejercí un derecho de acceso a la información pública con 17 preguntas sobre el sistema, resuelto favorablemente el 20

protegidos, precisamente, para evitar los sesgos y la discriminación. Algunas normas en España tienen en cuenta la necesidad de base de legitimación para la gestión de datos en sistemas de personalización del sector público, fenómeno ampliamente estudiado por Cerrillo. Es positivo el artículo 19, Ley 18/2020, de 28 de diciembre, de facilitación de la actividad económica sobre proactividad de la Administración, de Cataluña. Si bien, ni siquiera sería suficiente para datos especialmente protegidos. En el ámbito de la proactividad, las meras referencias a la personalización de servicios en modo alguno valdrían como base de legitimación legal⁸⁵. Los artículos 5 a), 25. 1º y 67. 3º Ley 4/2019, de 17 de julio de Administración digital de Galicia. También el artículo 28.6º Ley 8/2019, de 5 de abril, para una Administración más ágil, de Extremadura, es bastante general. Con ser positiva, tampoco sería suficiente al menos para datos especialmente protegidos la referencia del artículo 50.3º Ley 5/2021, de 29 de junio, de organización y régimen jurídico del sector público autonómico de Aragón. El Decreto 76/2020, de 4 de agosto, de Administración digital de Cataluña, (que obviamente no es ley) tiene acertadas referencias a la proactividad y la personalización (at. 4), que por sí no serían suficientes como habilitación, ni tampoco la mención a las garantías (art. 31). Cuestión diferente es que parece tratarse de usos automatizados de bajo impacto, una personalización limitada a la “finalidad de informar a las personas, de manera predictiva y anticipada, sobre los servicios públicos a los cuales pueden acceder” (art. 31.1º). Y por lo general el consentimiento al que acude el artículo 25.3º Ley 4/2019, de 17 de julio de Administración digital de Galicia solo valdría para legitimar servicios personalizados puramente informativos y relativamente intrascendentes o que el ciudadano tenga la alternativa a los mismos.

de agosto (GVRTE/2020/1163199). He dispuesto un acceso a la resolución que incluye las preguntas formuladas en <https://bit.ly/37f2apR>

⁸⁵ El tema lo aborda convenientemente aunque no como aquí desde el punto de vista de los datos especialmente protegidos o las discriminaciones especialmente prohibidas. CERRILLO I MARTÍNEZ, A. (2022): “La personalización de servicios digitales”, en CERRILLO I MARTÍNEZ, A. (coord.) *La Administración Digital*, Dykinson, Madrid, pp. 311-370, pp. 333 y ss.

7.3.- *Un canon de control “mucho más” estricto, e incluso excepcionalmente estricto*

La norma legal o en su caso reglamentaria que habilite a que el sistema de IA genere tratos diferenciados especialmente prohibidos, habrá de ser sometida a un juicio de igualdad estricto. El artículo 2.2º Ley 15/2022 permite tratos diferentes en estos supuestos específicos si son razonables y objetivos, con propósitos legítimos, también finalidades de proteger a personas o grupos de población. Asimismo, en razón del artículo 4.2º “No se considera discriminación” el trato diferente “que pueda justificarse objetivamente por una finalidad legítima y como medio adecuado, necesario y proporcionado para alcanzarla.” Hay que advertir que la superación de estos estándares habrá de *mirarse con lupa*: “el canon de control, al enjuiciar la legitimidad de la diferencia y las exigencias de proporcionalidad *resulta mucho más estricto*, así como más rigurosa la carga de acreditar el carácter justificado de la diferenciación” (entre otras muchas, sentencia 75/1983, de 3 de agosto, FFJJ 6 y 7). Y ello se agravará en la mayoría de los casos en los que se traten datos especialmente protegidos en los que aún rige con mayor rigor la excepcionalidad y la prohibición. Así pues, en principio serán pocos los supuestos en los que no será necesaria una ley de calidad y muchas garantías para que un sistema de IA público trate datos especialmente protegidos y lo haga para establecer diferencias de trato.

En consecuencia, estos casos en los que hay ley pasarán a ser en principio una cuestión de igualdad en la aplicación de esa ley. Y en todo caso habrán de controlarse y bajo estos estándares estrictos todas las medidas de gestión de datos, diseño del sistema y técnicas y organizativas que se hayan adoptado para que el sistema de IA sí que establezca diferencias de trato.

En EEUU la clave jurídica para admitir que los algoritmos y el big data incluyan información racial reside en que el algoritmo no haya elegido el factor raza con finalidad discriminatoria (“because of” “discriminatory purpose”), que es lo que prohíbe la doctrina *McCleskey v. Kemp* 481 U.S. 279, 291-92 (1987). Por el contrario, quizá se podría admitir que se tome en cuenta el factor raza por ser un elemento que aumenta la precisión “a pesar de” (“in spite of”) sus efectos sobre un grupo identificable. Habrá que analizar rigurosamen-

te cada caso concreto el funcionamiento y resultados de la IA, eso sí, bajo la presunción de discriminación.

7.4.- *La necesidad de regular legalmente el uso de datos especialmente protegidos para poder evaluar y minimizar la posible discriminación*

Como se recuerda en Países Bajos⁸⁶, precisamente para contrarrestar los efectos involuntarios e indeseables de los algoritmos, puede ser necesario tratar datos personales especialmente protegidos como variables de control para el desarrollo de modelos y métodos. De este modo se pueden eliminar sesgos y corregir las deficiencias ocultas de los conjuntos de datos o los modelos. En la misma dirección, señala Soriano que “es más sencillo enseñar a los sistemas automatizados a no discriminar con base en las categorías sospechosas si dichos datos se introducen y consideran expresamente”.⁸⁷ Si en el sector privado el consentimiento en muchos casos podría permitir el uso de estos datos para esta finalidad, no es así como regla general para el uso de IA en el sector público, por lo que se hace necesaria una especial regulación habilitante y con garantías. En esta línea, el RIA en su Considerando 44 afirma que “los proveedores deben poder tratar también categorías especiales de datos personales, como cuestión de interés público sustancial [...] con el fin de garantizar el seguimiento, la detección y la corrección de los sesgos en relación con los sistemas de IA de alto riesgo” y el artículo 10. 5º RIA contiene la habilitación expresa.

⁸⁶ De particular interés los criterios generales afirmados en el documento MINISTERIE VAN JUSTITIE EN VEILIGHEID (Ministerio de Justicia y Seguridad), Rijksoverheid (Gobierno central) (2021), *Richtlijnen voor het toepassen van algoritmen door overheden en publieksvoorlichting over data-analyses (Pautas para la aplicación de algoritmos por parte de los gobiernos y educación pública sobre análisis de datos)*, Directiva (Richtlijn), de 08-03-2021, <https://acortar.link/6C226N>, pág. 23 y ss. <https://acortar.link/6C226N> Se remite especialmente a LAMMERANT, H. y otros (2019): *Big data decision-making processes and shortcuts of discrimination*, NTM/NJCM-bull. 2018/1, p. 10-11 y ejemplos como van der SANGEN, M. (2019), *Research on fair algorithms for policy*, 6-8-2019, <https://www.cbs.nl/nl-nl/corporate/2019/32/research-on-fair-algorithms-for-policy>

⁸⁷ En la misma línea SORIANO ARNANZ, A. (2021 a), *ob. cit.*

8.- ZONAS GRISES. GARANTÍAS ANTE EL USO POR EL SISTEMA DE IA PÚBLICO DE “PROXIES” Y GRUPOS ALGORÍTMICOS

8.1.- *Tratamientos diferentes a partir de “proxies” o datos afines o indirectos de grupos especialmente protegidos*

Hay que tener en cuenta el fenómeno de los datos llamados *proxies* o indirectos, esto es, datos en principio no vinculados a circunstancias especialmente sospechosas de discriminación, ni datos especialmente protegidos (por ejemplo, datos geográficos, de pautas y productos de consumo, formas de vida, etc.). Recuerda Watcher cómo los dueños de perros, los adolescentes tristes, los jugadores de videojuegos, los padres solteros o los pobres se utilizan habitualmente para asignar recursos y tomar decisiones como qué publicidad mostrar, precio ofrecer o servicio público financiar⁸⁸. En principio, no aplicarían las garantías más intensas. Sin embargo, fácilmente el sistema de IA puede inferir de estos datos factores especialmente prohibidos o datos especialmente protegidos. Es por ello que hay que evitar que se eludan intencionalmente las garantías que corresponderían y analizar posibles “enmascaramientos”⁸⁹. Apunta Van der Sloot en esta línea que los datos no personales pueden convertirse en datos sensibles en una fracción de segundo, que basar la protección jurídica en la naturaleza de los datos “no es la mejor manera de avanzar”.⁹⁰ Como he

⁸⁸ WACHTER, S. (2022): “The Theory of Artificial Immutability: Protecting Algorithmic Groups under Anti-Discrimination Law”, 15 de febrero, *Tulane Law Review*, pp. 5-6, <http://dx.doi.org/10.2139/ssrn.4099100>

⁸⁹ Entre otros, TERRY NICOLAS, “Big Data Proxies and Health Privacy Exceptionalism”, en *Health Matrix* n° 24, pp. 65-108 (2014). Sobre el enmascaramiento de datos especialmente prohibidos, BAROCAS, S. y SELBST, A. D. (2016), *ob. cit.* pp. 692-693. NICOLAS, T. (2014) “Big Data Proxies and Health Privacy Exceptionalism”, en *Health Matrix* n° 24, pp. 65-108. Para el ámbito de la justicia, SIMMONS, R. (2016): “Quantifying Criminal Procedure: How to Unlock the Potential of Big Data in Our Criminal Justice System”, *Mich. St. L. Rev.* 947 (Ohio State Public Law Working Paper No. 362. pp. 977 y ss. <http://dx.doi.org/10.2139/ssrn.2816006>

⁹⁰ VAN DER SLOOT, B. (2021): “Regulating non-personal data in the age of Big Data”, en TZANOU, M. (Ed.) (2021): *Health data privacy under the GDPR* :

sostenido al respecto en el ámbito de salud, el tratamiento jurídico ha de venir claramente ligado con la finalidad de su uso y el riesgo que se genera⁹¹. De este modo, hay que poner especialmente el foco en la afinidad de los datos con datos sensibles o de categorías sospechosas de discriminación. O, directamente, analizar si los efectos que genera el tratamiento impactan en estos colectivos sospechosos de discriminación, aunque los datos no sean especialmente sensibles. Si es así, no puede considerarse que el sistema sea aparentemente neutro y habrá que aplicar las especiales garantías de las discriminaciones prohibidas. Como he adelantado, en estos supuestos puede resultar práctico acudir a las categorías de discriminación por asociación y especialmente la discriminación indirecta. Desde EEUU, se afirma que algunos *proxies* como la información demográfica si se identifican “deben evitarse” por su “alto riesgo de conducir a la discriminación algorítmica”⁹².

8.2.- Creación dinámica de grupos y colectivos humanos y la privacidad y no discriminación colectiva y de grupo

Una de las novedades que implica la IA es la creación dinámica de grupos y colectivos humanos que no comparten en principio rasgos sospechosos. Apunta Watcher que la IA también crea grupos aparentemente incomprensibles, definidos por parámetros que desafían la comprensión humana, como píxeles en una imagen, comportamiento de clic, señales electrónicas o tráfico web y a resultados de ellos se deciden préstamos o solicitudes de empleo, que tienen un impacto significativo en la vida de las personas⁹³. Ello además sin que los interesados percibamos ni la inclusión en estos grupos, ni las implicaciones que

Big Data challenges and regulatory responses, Routledge pp. 85-105; p. 16 de la versión SSRN.

⁹¹ COTINO HUESO, L. (2020 b): “El alcance e interacción del régimen jurídico de los datos personales y big data relacionado con salud y la investigación biomédica”, *Revista de derecho y genoma humano: genética, biotecnología y medicina avanzada*, n° 52, pp. 57-96, n°52 enero-junio 2020. acceso

⁹² WHITE HOUSE (2022), *op.cit.*, p. 23 y 26.

⁹³ WACHTER, S. (2022), *ob. cit.* pp. 6, 10, 24.

puede tener. Para Jaume,⁹⁴ en estos casos la identidad de los individuos no es relevante y se trata de la creación de grupos fuera de las categorías sociales convencionales de una cultura, e incluso al margen de los grupos tradicionalmente discriminados. Floridi y el grupo de Tilburg, que abordan especialmente la “privacidad de grupo en la era de los datos”⁹⁵; exponen -y así lo recuerda también la AEPD⁹⁶-, que la mayoría de las personas no están perfiladas individualmente, sino como miembros de un grupo específico. Los datos empleados para generar el perfil de estos colectivos o grupos podrían no tener la consideración de datos personales. Y afirma la AEPD que estos grupos “no están establecidos como tal y no están sujetos a ninguna forma jurídica”, por ello concluye que el perfilado y las inferencias sobre estos grupos “carecerían de cualquier protección legal.”⁹⁷

Pues bien, como ya he sostenido⁹⁸ ciertamente no lo considero así en tanto en cuanto estas clasificaciones, aunque no supongan tratamiento de datos personales, pasan a serlo en el momento en el que se aplican a una persona concreta. En muchos casos serán supuestos de tratamiento de *datos inferidos*⁹⁹, a los que Palma ha prestado especial

⁹⁴ JAUME PALASÍ, L. (2020): “Cómo la inteligencia artificial está impactando en las sociedades”, CERRILLO I MARTÍNEZ, A. Y PEGUERA POCH, M. (coords.), (2020): *Retos jurídicos de la inteligencia artificial*, Aranzadi, Cizur, pp. 27-39, p. 28.

⁹⁵ En particular, KAMMOURIEH, L y otros, “Group privacy in the age of big data”, en TAYLOR, L.; VAN DER SLOOT, B.; FLORIDI, L. (eds.) (2017): *Group Privacy: New Challenges of Data Technologies*, Springer, Capítulo 3, pp. 48-83, ver p. 54. <https://www.stiftung-nv.de/sites/default/files/group-privacy-2017-authors-draft-manuscript.pdf>

⁹⁶ AEPD (2020 b): *Privacidad de grupo*, 19 de octubre, <https://www.aepd.es/es/prensa-y-comunicacion/blog/privacidad-de-grupo>

⁹⁷ *Ibidem*.

⁹⁸ (2022): “Nuevo paradigma en la garantías de los derechos fundamentales y una nueva protección de datos frente al impacto social y colectivo de la inteligencia artificial”, en COTINO HUESO, Lorenzo (editor), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Thompson-Reuters Aranzadi, FIADI (Federación Iberoamericana de Asociaciones de Derecho e Informática), Cizur.

⁹⁹ Especialmente cabe seguir, VAN DER SLOOT, B. (2021), *ob. cit.* y en España a POLO ROCA, A. (2021):, “Datos, datos, datos: el dato personal, el dato no personal, el dato personal compuesto, la anonimización, la pertenencia del dato y otras cuestiones sobre datos”. *Estudios de Deusto: revista de la Universidad de Deusto*, Vol. 69, n°. 1, pp. 165-194, DOI: 10.18543/ed-69(1)-2021

atención¹⁰⁰, esto es, datos de salida del sistema de IA a partir del tratamiento de datos personales aportados por el interesado o vinculados al interesado. Aunque el TJUE no es tan claro, sin duda alguna para el Grupo del Artículo 29 estos datos son datos personales, aunque no hayan sido facilitados el interesado¹⁰¹. En consecuencia, si el sistema de IA público confiere tratos diferentes en razón de la pertenencia a estos colectivos, como mínimo se aplicarán las garantías ordinarias de la discriminación, así como de la protección de datos. Y si puede considerarse que hay parámetros afines a los datos especialmente protegidos o a las categorías especialmente sospechosas, nos situaríamos ante un supuesto de *proxies* antes expuesto. Las soluciones propuestas por Watcher y su “teoría de la inmutabilidad artificial” y la proyección de su el “derecho a inferencias razonables”¹⁰², pese al interés de su reflexión, en cierto modo son reconducibles a la solución aquí expuesta.

9.-¿CÓMO PROBAR LA DISCRIMINACIÓN ALGORÍTMICA PÚBLICA? INVERSIÓN GENERAL DE LA CARGA DE LA PRUEBA Y LA NECESIDAD DE LA TRANSPARENCIA Y EXPLICABILIDAD

En general la discriminación se enfrenta a graves problemas de prueba y más en los casos de la discriminación algorítmica¹⁰³.

Para empezar, lamentablemente hoy por hoy, ni siquiera se conoce la existencia misma de buena parte de los sistemas IA públicos, por lo

¹⁰⁰ Sobre el tema, PALMA ORTIGOSA, A. (2021): *Régimen jurídico de la toma de decisiones automatizadas y el uso de sistemas de inteligencia artificial en el marco del derecho a la protección de datos personales*, Tesis doctoral Universidad de Valencia, 2021, que tuve el gusto de dirigir. Sobre la noción, WACHTER, S. y MITTELSTADT, B. D. (2019): “A right to reasonable inferences: re-thinking data protection law in the age of big data and AI”, *Columbia Business Law Review*, vol. 2019, No. 2, pp. 494-620, p. 2 versión <https://ssrn.com/abstract=3248829>

¹⁰¹ GRUPO DEL ARTÍCULO 29 (2018), *ob. cit.* p. 10.

¹⁰² WACHTER, S. y MITTELSTADT, B. D. (2019), *ob. cit.* entre otras pp. 49-50.

¹⁰³ En general y por todos, respecto de la prueba de la discriminación, FRA- CON- SEJO DE EUROPA (2018), *ob. cit.* pp. 261-282. Respecto de algoritmos TISCH- BIREK, A. (2020), *ob. cit.*

que se hace más difícil saber si discriminan. Si se conocen, por lo general no se tratará de discriminaciones algorítmicas directas y evidentes en las que la norma o los responsables y proveedores del sistema de IA expresamente indiquen que en su diseño y medidas adoptadas han manejado el uso de datos de colectivos especialmente discriminados o que las variables del sistema preferencian o perjudican a unos colectivos frente a otros y en qué modo concreto. Si así es, obviamente será más sencillo probar la posible discriminación. De igual modo, especialmente por cuanto al funcionamiento autónomo del sistema de IA o de *machine* o *deep learning*, no será fácil que el mismo sistema muestre cómo efectúa tratos diferentes no previstos en el diseño. Y obviamente, si la causa de la discriminación son errores, éstos difícilmente serán plenamente visibles y parametrizados para que se puedan conocer y probar.

La tendencia del Derecho antidiscriminatorio en general ha sido la reducción de la necesidad o la inversión de la carga de la prueba por parte de la persona afectada. La regulación europea y española de no discriminación es bastante incisiva en terreno de la prueba. A continuación se señalan los elementos clave proyectados para el ámbito específico del uso de sistemas públicos de IA.

9.1.- Los indicios que debe aportar la persona afectada por el sistema de IA público. Inversión general de la carga de la prueba y sus excepciones

La persona afectada por un sistema de IA público (o la persona legitimada para su impugnación) debe aportar “indicios fundados” sobre la existencia de la discriminación (art. 30. 1º Ley 15/2022, de 12 de julio). Así pues, se requieren datos, documentación, estadísticas u otras pruebas sobre el sistema de IA que permitan hacer pensar que genera errores o resultados diferentes respecto de grupos o colectivos vinculados a categorías especialmente prohibidas, o diferentes niveles de falsos negativos o positivos. O que el sistema tiene en cuenta datos o variables relacionadas con estos colectivos. En algunos supuestos de discriminación indirecta, se tratará de mostrar indicios de que se trata de modo idéntico a quienes habrían de ser tratados de modo diferente por el sistema.

Señala Soriano que en principio no ha de ser problemático aportar términos de comparación del funcionamiento del sistema, al apreciar los resultados que arroja el sistema en otros muchos supuestos¹⁰⁴. Asimismo, en muchos supuestos se podrá contar con información sobre el sistema.

En los casos de discriminaciones especialmente prohibidas, si se logra aportar suficientes indicios fundados, pasará a presumirse que el sistema de IA sí que es discriminatorio. Se invierte la carga de la prueba: “la carga de demostrar el carácter justificado de la diferenciación recae sobre quien asume la defensa de la misma” (sentencia 81/1982, de 21 de diciembre, FJ 2)¹⁰⁵. Es al responsable del sistema de IA público a quien compete demostrar que no es discriminatorio. Ello sucedería en los supuestos de discriminaciones algorítmica directa, indirectas y encubiertas, así como discriminaciones por error. Esta inversión de la carga de la prueba también habrá de darse cuando el uso del sistema de IA quede sea vinculable al ejercicio de un derecho fundamental (salud, por integridad y vida, uso con, profesores, centros educativos, alumnos o sus padres educación, así como con sindicatos partidos, asociaciones ideológicas o religiosas, etc.).

En principio, en los *supuestos en los que el sistema de IA público no parezca comprometer o afectar a colectivos tradicionalmente discriminados*, ni quede vinculado con otros derechos fundamentales, no habría inversión de la carga de la prueba y el afectado habría de probar que el sistema contiene errores o sesgos relevantes o establece diferencias de trato que son claramente irrazonables. Sin embargo, considero que una vez probado el indicio de discriminación del sistema de IA público por la parte afectada, *también habrá de darse en general la inversión de la carga de la prueba* y que sea el responsable del sistema de IA público quien pruebe que no hay discriminación, sesgos y errores relevantes. A mi juicio debe reconocerse esta

¹⁰⁴ Así SORIANO ARNANZ, A. (2022 a y b), se apoya en XENIDIS, R. y SENDEN, L. (2020): “EU Non-discrimination law in the era of artificial intelligence: mapping the challenges of algorithmic discrimination”, en BERNITZ, U. y otros, (eds.), *General Principles of EU Law and the EU Digital Order*, Alphen aan den Rijn, Kluwer Law International, 2020, p. 171..

¹⁰⁵ Al respecto, véanse las STC 81/1982, de 21 de diciembre, 34/1984, de 9 de marzo, 20/1985, de 14 de febrero, entre otras.

inversión general de la carga de la prueba en razón del llamado *data due process* desarrollado en buena medida en EEUU desde hace una década¹⁰⁶. Y es que desde el punto de vista de los afectados por un sistema público de IA su posición es de una fuerte asimetría informativa frente al sector público y sus proveedores de los sistemas de IA. Ya sólo reunir “indicios” puede ser una labor extremadamente compleja.

Sin embargo, no será necesaria esta inversión de la prueba si no se da esta asimetría. No será necesaria si la diferencia de trato la establece la propia normativa que regula el sistema de IA o se trata de decisiones expresas en el diseño e implantación del sistema y sobre las mismas hay publicidad activa con suficiente información y explicabilidad del sistema de IA para la prueba de la discriminación.

Para que el afectado pueda hacerse con estos indicios o en su caso pruebas, pasa a ser esencial la existencia y cumplimiento de obligaciones de registros públicos de algoritmos (para algunos sistemas de alto riesgo del futuro RIA), de transparencia activa obligatoria de los sistemas IA utilizados (p. e. art. 16.1 l) Ley 1/2022, de 13 de abril), o las autorizaciones del uso de estos sistemas y su difusión en sede electrónica (p.e. arts. 11 y 13 RAFESP)¹⁰⁷, la publicidad de los estudios de impacto que se efectúen, o del registro de tratamientos de datos (art. 31 Ley Orgánica 3/2018). Asimismo, puede ser clave el ejercicio del derecho de acceso a la información pública (arts. 20 CE y 105 CE, Ley 19/2013) o las de acceso de protección de datos y transparencia (arts. 15 y 12 y ss. RGPD), reforzado en el caso de decisiones automatizadas (art. 22 y 13 RGPD). Al respecto cabe remitir a este estudio, así como a diversos trabajos sobre las obligaciones de transparen-

¹⁰⁶ Especialmente CRAWFORD, K. Y SCHULTZ, J. (2014): “Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms”. *Boston College Law Review*, Vol. 55, No. 93 <https://ssrn.com/abstract=2325784>.; CITRON, D. K. (2008), *ob.cit.* No expresan esta inversión general de la carga de la prueba, pero sí asentaron las obligaciones de generar registros, poder auditar los sistemas, obligaciones de información mínima, etc.

¹⁰⁷ Sobre este reglamento y su regulación al respecto, puede seguirse mi estudio (2021 b): “El nuevo reglamento de Administración electrónica, que no innova en tiempos de transformación digital”, *Revista Catalana de Derecho Público* nº 63,, pp. 118-136, DOI: <http://dx.doi.org/10.2436/rcdp.i63.2021.3672>

cia¹⁰⁸. El seguimiento de políticas de transparencia no obligatorias por proveedores o usuarios de los sistemas de IA públicos puede ser también esencial para facilitar que los afectados conozcan la existencia del sistema, del impacto e importancia de este, su diseño, datos de entrada y funcionamiento.

Tampoco se daría esta inversión general de la carga en el uso de IA contextos en los que se da intensamente la presunción de inocencia: “procesos penales ni a los procedimientos administrativos sancionadores, ni a las medidas adoptadas y los procedimientos tramitados al amparo de las normas de organización, convivencia y disciplina de los centros docentes” (art. 13. 2º Ley Orgánica 3/2007, art. 30. 3º Ley 15/2022, de 12 de julio).

Hay que llamar la atención que en septiembre de 2022 se ha presentado una Propuesta de Directiva sobre responsabilidad en materia de IA¹⁰⁹. La misma es especialmente destacable en materia de carga de la prueba y presunción de responsabilidad, en particular su artículo 4. Salvo en el ámbito penal, no hay que excluir su proyección respecto de la responsabilidad por uso público de la IA, así como en el caso de la colaboración público-privada.

9.2.- *Lo que debe probar el proveedor o el usuario del sistema de IA público*

En general, a partir de la aportación de los indicios suficientes por el afectado, es el responsable del sistema de IA público quien debe probar que no discrimina. Así, según los casos, el responsable del sistema habría de demostrar:

¹⁰⁸ La literatura ya es abundante, me remito a los trabajos de la monografía CO-TINO, L. Y CASTELLANOS, J. (coords.) (2022), el ya referido de JIMÉNEZ LÓPEZ, Jesús (2022) *ob.cit.*, en la misma obra, BENÍTEZ PALMA, E. “Mucho ruido y pocas nueces: algunas consideraciones sobre los registros públicos de algoritmos”, pp. 83-96.

¹⁰⁹ <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52022PC0496> Sobre el tema en la UE https://ec.europa.eu/info/business-economy-euro/doing-business-eu/contract-rules/digital-contracts/liability-rules-artificial-intelligence_en

- Que el sistema no es erróneo ni genera sesgos o que estos son irrelevantes.
- Que la diferencia de trato se basa en un factor objetivo no relacionado con un motivo especialmente prohibido.
- Que el tratamiento diferente que genera el sistema cuenta con “justificación objetiva y razonable, suficientemente probada, de las medidas adoptadas y de su proporcionalidad” (art. 30. 1º Ley 15/2022, de 12 de julio)¹¹⁰. Así, el proveedor y en su caso la Administración como usuaria del sistema de IA habrían de expresar y detallar la finalidad o finalidades pretendidas y cómo se ha configurado el sistema, la selección de datos de entrada y entrenamiento y otros elementos, y que estas decisiones técnicas y organizativas arrojan resultados proporcionales respecto de tales finalidades y con relación a las diferencias que generan.
- Que, en su caso, si el sistema trata de modo más favorable a determinados colectivos y lo hace de forma razonable o proporcional, en supuestos de acciones positivas, obligaciones de diferenciación o medidas de minimización de sesgos, por ejemplo.

Puede ser muy importante para el responsable del sistema de IA -presunto discriminador- contar con las evidencias sobre el cumplimiento de todas las exigencias proactivas y preventivas, tanto las que se exijan debido a la protección de datos (como los estudios de impacto, análisis de riesgos, etc.) o de otra normativa aplicable, como en su caso, el futuro reglamento de IA. El cumplimiento de estos deberes genera toda una serie de información y documentación. Asimismo, la demostración del cumplimiento de estos deberes prueba la diligencia debida, así como la falta de cualquier ánimo discriminatorio o finalidad contraria a la ley. En su caso, las evidencias que puedan aportarse serán reactivas, esto es, a partir del conflicto y cuestionamiento del funcionamiento del sistema. La falta de esta información podrá ser la prueba de la discriminación por “inacción” (art. 4.1º Ley 15/2022) u omisión.

¹¹⁰ Igualmente cabe tener en cuenta, en los mismos términos que esta última ley, el artículo 77. 3º bis LPAC, de 1 de octubre, del Procedimiento y artículo 60. 7º Ley 29/1998, de 13 de julio, reguladora de la Jurisdicción Contencioso-Administrativa.

La regulación de igualdad también indica que el órgano judicial, a instancia de parte, podrá recabar, si lo estimase útil y pertinente, informe o dictamen de los organismos públicos competentes (art. 13. 1º Ley Orgánica 3/2007, art. 30. 2º Ley 15/2022, de 12 de julio). A lo anterior cabe aplicar todo lo procedente sobre las facultades administrativas o judiciales de solicitud de prueba, proyectadas al ámbito tecnológico y a la especialidad de los sistemas de IA en los que pueden concurrir importantes cuestiones de propiedad o secreto industrial, secreto y materia clasificadas u otras restricciones.

10.- MÁS VALE PREVENIR QUE DISCRIMINAR. ¿CUÁLES SON LOS REMEDIOS Y MEJORES PRÁCTICAS PARA EVITAR LA DISCRIMINACIÓN ALGORÍTMICA PÚBLICA?

Desde hace años instituciones y expertos vienen señalando los mejores remedios y buenas prácticas para evitar los sesgos, errores y discriminación. A continuación se destilan los que se consideran más relevantes. Hay que adelantar que buena parte de ellos no sólo son buenas prácticas o recomendaciones, sino que pasarían a integrarse en las medidas técnicas y organizativas que las autoridades de protección de datos ya exigen, que por ejemplo se han de adoptar en los estudios de impacto según los riesgos. Muchas medidas también se integrarían entre las muchas obligaciones preventivas que impone el futuro RIA respecto de un sistema de alto riesgo.

Desde 2018 el Grupo del Artículo 29 G29-UE¹¹¹ ha detallado una serie de medidas respecto decisiones automatizadas esencialmente “para garantizar que las personas reciben un trato justo y no discriminatorio”, sin “resultados discriminatorios, erróneos o injustificados”, entre otras insistía en los controles periódicos de calidad, auditorías algorítmicas, en su caso independientes de terceros, garantías contractuales respecto de los algoritmos de terceros que garanticen comprobaciones y cumplimiento normativo. Asimismo, mecanismos de certificación para operaciones de tratamiento, códigos de conducta

¹¹¹ GRUPO DEL ARTÍCULO 29 (2018), *ob. cit.* pp. 37-38.

para procesos de auditoría, comités de ética para evaluar los daños y beneficios.

También desde 2018 y adoptadas finalmente en 2019, las Directrices para la ética en el diseño en la UE del HLEG¹¹² incluyen una lista de evaluación, un *checklist* con 150 cuestiones y existe también una adecuación para el ámbito de salud (www.future-ai.eu). Muchas de ellas son de interés para lograr la exactitud y fiabilidad (bloque 2)¹¹³, las integridad y calidad de los datos (bloque 3) y, específicamente, un bloque sobre “Diversidad, no discriminación y equidad” (5) con 19 cuestiones para evitar el sesgo y la discriminación¹¹⁴. También se incluyen cuestiones sobre accesibilidad y diseño universal y los bloques sobre bienestar social y ambiental y de responsabilidad también incluyen cuestiones de interés¹¹⁵. Se trata de una herramienta básica para la evaluación y autoevaluación de sistemas IA que perfectamente puede ser proyectada para el sector público, al igual que sucede con la Guía de auditoría de IA de la AEPD¹¹⁶.

¹¹² HLEG (2019), *op.cit.*, pp. 26-31. De interés el portal ALTAI <https://futurium.ec.europa.eu/en/european-ai-alliance/pages/welcome-altai-portal>

¹¹³ ¿Ha implementado medidas para garantizar que los datos utilizados sean completos y estén actualizados?; ¿se necesitan datos adicionales?, ¿qué daño se causaría si el sistema de AI hace predicciones inexactas?, ¿el sistema está haciendo una cantidad inaceptable de predicciones inexactas?; ¿pasos para aumentar la precisión del sistema?, etc.

¹¹⁴ Entre otras, ¿Estableció una estrategia o un conjunto de procedimientos para evitar crear o reforzar un sesgo injusto en el sistema AI, tanto en el uso de los datos de entrada como en el diseño del algoritmo?, diversidad y representatividad en los datos; inclusión poblaciones específicas, uso de pruebas en las fases de desarrollo, incorporación de mecanismos específicos para que otros puedan identificar problemas de sesgo o discriminación; mecanismos para consultar estos problemas; consideración de afectados indirectos. Evaluación si variación de resultados en mismas condiciones, cómo medirlo. ¿Cómo se ha medido la imparcialidad?

¹¹⁵ El bloque sobre bienestar social y ambiental (6) cuenta con alguna cuestión sobre si, por ejemplo, se tuvo en cuenta el impacto social general, como pérdidas de empleo. Finalmente, en el bloque de responsabilidad (7) también incluye cuestiones sobre auditoría independiente, evaluación de riesgo o impacto, capacitación y educación, introducción de mecanismos de “junta de revisión ética de AI” o similares.

¹¹⁶ AEPD (2021): *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021.

En cualquier caso, *supra* se han expuesto las causas de los sesgos y discriminaciones algorítmicas, agrupadas respecto de los datos, la prueba evaluación y validación de los sistemas y métricas y la consecuente adopción de medidas de mitigación y los factores humanos. Sobre estos ejes cabe recordar los principales remedios.

En el *primer grupo de remedios, relativos a los datos*, el NIST o la Casa Blanca¹¹⁷ señalan que los datos definen en gran medida la tarea de aprendizaje automático en sí misma, por lo que hay que hay que “ir donde están los datos”, una cultura centrada en qué conjunto de datos podría ser el más adecuado y no tanto en cuáles son accesibles. Emplear una representación estadística equilibrada en los conjuntos de datos con técnicas como medidas de desequilibrio de clase.

Tras la relevante sentencia de nulidad del sistema de IA SYRI, Países Bajos han destacado en Europa -con el Reino Unido- en la adopción de medidas preventivas y de transparencia algorítmica como ningún otro país¹¹⁸. Además de su estudio de impacto algorítmico IAMA¹¹⁹, destaca una guía específica para la no discriminación¹²⁰. Se insiste en que hay que investigar la calidad de las fuentes de datos y si son suficientes para el propósito para el que se utilizan; identificar si existen limitaciones en el uso de una fuente de datos, algoritmo o método de análisis. Se recuerdan los requisitos legales de calidad, cantidad y actualización de

¹¹⁷ Destaca especialmente NIST; SCHWARTZ, R. y otros (2022), *op.cit.*, pp. 14-19 o WHITE HOUSE (2022), *op.cit.*, pp. 23-29, en especial pp. 26 y ss.

¹¹⁸ Destacan sus directrices MINISTERIE VAN JUSTITIE (2021), *ob. cit.*, para no discriminación y sesgos en especial pp. 23-24. Asimismo el Excel desarrollado por el ALGEMENE REKENKAMER (Tribunal de Cuentas), (2020): *Digitale Toetsingskader Algoritmes (Marco de evaluación digital de Algoritmos)*, Tribunal de Cuentas, Excel, <https://www.rekenkamer.nl/binaries/rekenkamer/documenten/publicaties/2021/01/28/download-het-toetsingskader/Toetsingskader+algoritmes+v1.0.xlsx>

¹¹⁹ MINISTRY OF THE INTERIOR AND KINGDOM (2022): *Impact Assessment. Fundamental rights and algorithms (IAMA)*, Países Bajos, marzo. <https://www.government.nl/binaries/government/documenten/reports/2022/03/31/impact-assessment-fundamental-rights-and-algorithms/Fundamental+Rights+and+Algorithms+Impact+Assessment.pdf>

¹²⁰ MINISTERIE VAN BINNENLANDSE ZAKEN; VAN DER SLOOT, B. y otros (2021): *Non-discriminatie by design*, 69 pp., encargo para (Ministerio del Interior), Tweede Kamer (Cámara de representantes), <https://www.tweedekamer.nl/downloads/document?id=2021D2272>

los datos, siendo los necesarios para los fines del análisis. Hay que describir la finalidad del sistema, cómo se manejarán los datos y los errores que se detecten¹²¹. Por cuanto a los datos de entrenamiento, hay que justificar la elección y poder parametrizar los posibles factores discriminatorios y análisis sobre las consecuencias de una elección diferente de la parametrización¹²².

Desde EEUU en 2022 se insiste en que los datos sean “representativos y sólidos”. Y el uso de los *proxies*, como la información demográfica, “debe evitarse” por su “alto riesgo de conducir a la discriminación algorítmica”. Deben realizarse pruebas proactivas para identificar los *proxies* y si se identifican eliminarse o buscar alternativas¹²³. Respecto la prueba, evaluación, validación y verificación de los sistemas y los algoritmos y las métricas empleadas, el NIST aporta toda una guía frente a los sesgos, necesidades de modelación del sistema, correlaciones espurias, la complejidad algorítmica, validación del rendimiento, también el aseguramiento en el despliegue y el sesgo de implementación, la necesidad de garantizar actualizaciones periódicas del modelo, y probar y recalibrar los parámetros del modelo en conjuntos de datos representativos actualizados, la actualización del modelo para evitar premisas que hayan quedado obsoletas¹²⁴.

Por cuanto al *segundo grupo, la elección de las técnicas de análisis de datos*, se insiste en valorar si es necesario usar métodos de inteligencia artificial, pues otros métodos de análisis suelen ser también ade-

¹²¹ ¿Son los datos relevantes para el objetivo?; ¿Es la agregación de datos en un nivel suficientemente informativo para el propósito?; ¿Se mantienen los datos de forma coherente a lo largo del tiempo?; ¿Se registran los datos de la misma manera?; ¿Hay suficiente variación en los datos? ¿Están determinadas subpoblaciones (minorías) sobrerrepresentadas en los datos, lo que podría provocar un sesgo en el modelo de predicción?; ¿El conjunto de datos utilizado es correcto, oportuno y completo?; ¿Pueden vincularse las fuentes de datos (si se trabaja con varias tablas)?; ¿Se pueden rastrear las mutaciones de los datos (fuente)?; ¿Ha habido un cambio en la base de datos? MINISTERIE VAN JUSTITIE (2021)... *ob. cit.* Parte II.

¹²² Elección de datos de entrenamiento, posibles factores discriminatorios y documentarlos, registrar qué datos de entrenamiento y qué parámetros se han utilizado, cómo afectaría a los resultados una elección diferente de los parámetros, cómo afectaría a los resultados un conjunto de datos de entrenamiento diferente. *Id.*

¹²³ WHITE HOUSE (2022), *op.cit.*, pp. 23-29, en especial 26 y ss.

¹²⁴ NIST; SCHWARTZ, R. y otros (2022), *op.cit.*, pp. 20-27.

cuados. En cualquier caso, hay que justificar la elección de la técnica y las hipótesis preconcebidas¹²⁵. Respecto de la metodología, hay que describir los pasos del desarrollo del modelo respecto de los datos, el tratamiento de datos (preparación y limpieza de datos, composición y estructura de las pruebas modelo, marco de datos básico, creación de variables, corrección de errores de datos). El desarrollo del modelo y la evaluación del modelo. También, por cuanto al uso del algoritmo en la práctica, hay que tener en cuenta el riesgo de que los algoritmos de autoaprendizaje hagan uso de correlaciones y relaciones estadísticas sobre las que es cuestionable que sean adecuadas para la toma de decisiones. Respecto del aprendizaje automático (supervisado), se apunta que hay que hacer explícitos (cuantitativamente) los requisitos relativos a la aplicabilidad, la previsibilidad, la explicabilidad, la discriminación indeseable, etc., de modo que se reduzca aún más el espacio de soluciones. En concreto, se señala que hay que indicar a qué población y las circunstancias en las que debe aplicarse el modelo. Hay que fijar por expertos cómo se evaluará el poder predictivo del modelo, con qué métricas y valores de umbral; o cuándo se determinará si el modelo es suficiente, si falsos positivos son menos malos que los falsos negativos. Así, cabe fijar la discriminación o sesgo indeseable, que sea medible y objetivo. Asimismo, se debe cuantificar el sesgo deseado si lo hay.

En el ámbito de la evaluación y minimización del sesgo y la discriminación ya en 2017 la resolución del Parlamento UE sobre macrodatos “insta” a “minimizar la discriminación y el sesgo algorítmicos” (nº 20) y afirma también la necesaria “mitigación algorítmica” (nº 21, ver también 32). Subraya la posibilidad de corrección de datos y de recurrir decisiones algorítmicas. Los propios algoritmos pueden ser el antídoto contra el sesgo algorítmico y se pueden programar para ignorar o minimizar la importancia que los factores prohibidos en sus decisiones.

AI Watch respecto del sector público señala que los riesgos deben evaluarse sistemáticamente con un procedimiento estructurado y bien definido, se deben identificar las medidas de mitigación adecuadas y ello “debe convertirse en una rutina para las organizaciones

¹²⁵ MINISTERIE VAN JUSTITIE (2021), *ob. cit.* pp. 23-24.

públicas.”¹²⁶. Hay que probar los sistemas para evaluar si se producen sesgos y habrá que documentar los y en su caso ver si es aceptable. Hay que tener en cuenta los cambios en el sistema después del despliegue o de cambios en el contexto de uso o en los datos asociados. La evaluación debe realizarse periódicamente y más los sistemas de mayor riesgo, y siempre que se produzca un patrón de resultados inusuales¹²⁷.

Cuando se detecta la disparidad o sesgo, hay que mitigarlo o eliminarlo. Si no se puede mitigar, debe reconsiderarse el uso del sistema automatizado. Sin embargo, hay que garantizar que el sistema inteligente no pierda su eficacia, su propia naturaleza o lleve a generar resultados absurdos por la eliminación de patrones o datos. No siempre la eliminación es la mejor solución. Por otra parte, resulta prácticamente imposible eliminar todos aquellos datos que revelen la pertenencia a un grupo especialmente protegido sin hacer que el algoritmo sea inservible¹²⁸.

Deben realizarse evaluaciones independientes. Y especialmente se señala que en el caso de los usos del sector público las evaluaciones deberían hacerse públicas salvo restricciones legales o por seguridad nacional. Asimismo, deben presentarse informes incluyendo quién evalúa el sistema y las medidas correctivas que se derivan. Estos informes han de hacer referencia a las evaluaciones de equidad en la fase de diseño, los diseños y pruebas de accesibilidad, las pruebas de disparidad, documentar cualquier disparidad restante y detallar cualquier aplicación y evaluación de mitigación¹²⁹.

El *tercer gran bloque es relativo a los factores humanos*, incluidos los sesgos sociales e históricos, respecto de los que desde NIST, Unesco y otros ámbitos se establecen orientaciones de actuación bá-

¹²⁶ JOINT RESEARCH CENTRE(2022), *ob. cit.* p. 6.

¹²⁷ WHITE HOUSE (2022), *op.cit.*, p. 24.

¹²⁸ BID- Universidad Adolfo Ibáñez (2022), *op.cit.*, pp. 6-9. También, GOODMAN, B. W. (2016): “A step towards accountable algorithms?: Algorithmic discrimination and the European Union general data protection”, *29th Conference on Neural Information Processing Systems*, <http://www.mlandthelaw.org/> sigo en este caso por SORIANO ARNANZ, A. (2021 b), *ob. cit.*

¹²⁹ WHITE HOUSE (2022), *op.cit.*, p. 28.

sicas y se contienen numerosas propuestas.¹³⁰ Se destacan como remedios la necesidad de evaluaciones de impacto, la integración en el diseño, desarrollo, implementación y evaluación de las múltiples partes interesadas, priorizar la diversidad, la equidad y la inclusión en los equipos involucrados en la capacitación y el despliegue de sistemas de IA. Además, el NIST analiza las cuestiones de la necesaria gobernanza y supervisión a través de canales de reclamación, políticas y procedimientos, obligaciones de documentación, fijación de responsabilidades y otras recomendaciones¹³¹.

Especialmente por cuanto a la transversalidad y la integración de los equipos humanos desde Unesco y el Consejo de Europa se insiste en la necesidad de concienciación especial en las profesiones STEM, así como la necesidad de insertar en equipos el enfoque de género¹³². Como recuerda Valle Escolano, se necesitan equipos de trabajo diversos y multidisciplinares, se ha de insistir en una cultura de la ética y la responsabilidad en las con controles y equilibrios¹³³. Si es necesario deben crearse nuevos conjuntos de datos, así como adaptar los existentes. Hay que concienciar y sensibilizar de los problemas que se generan por la falta de representatividad y sesgo, para que así se incremente la recopilación y el uso de datos interseccionales, especialmente de grupos subrepresentados. Se afirma también que hay que comprometerse a usar conjuntos de datos que hayan sido desarrollados con un prisma de igualdad de género (pese a que ello pudiera suponer un aumento de costes).

¹³⁰ NIST; SCHWARTZ, R. y otros (2022), *op.cit.*, pp. 35-42; VALLE ESCOLANO, R. (2022), *ob. cit.*; UNESCO, *Artificial Intelligence and Gender Equality*, cit.

¹³¹ NIST; SCHWARTZ, R. y otros (2022 a), *op.cit.*, pp. 42-47.

¹³² Sobre este tema y necesidad de concienciación especial en profesión, CONSEJO DE EUROPA (2018), *op.cit.*, pp. 55 y ss.; UNESCO (2020): *Artificial Intelligence and Gender Equality. Key findings of UNESCO's Global Dialogue*, pp. 19 y ss., 22-23, así como la posibilidad de paneles éticos, p. 24, <https://unesdoc.unesco.org/ark:/48223/pf0000374174>. Para el ámbito laboral, EUROPEAN INSTITUTE OF GENDER EQUALITY (EIGE), *Artificial intelligence, platform work and gender equality*, 2021, <https://eige.europa.eu/publications/artificial-intelligence-platform-work-and-gender-equality-report>

¹³³ VALLE ESCOLANO, R. (2022), *ob. cit.*, p. 99-100.

11.- LA (CRECIENTE) OBLIGACIÓN DE DISEÑO Y ESTUDIOS DE IMPACTO FRENTE A LA DISCRIMINACIÓN ALGORÍTMICA DE LOS SISTEMAS DE IA PÚBLICOS

11.1.- *Los deberes de responsabilidad proactiva, privacidad en el diseño o estudio de impacto ya son exigibles en general a los usos públicos de la IA*

Como he tenido ocasión de exponer (Cotino, 2022) ante la IA hay que asumir técnicas de responsabilidad “proactiva, “demostrada” o “accountability” en el diseño, bajo el modelo jurídico del “más vale prevenir que curar”¹³⁴. A modo de las leyes de la robótica de Asimov, supone de la inserción del cumplimiento ético y legal en el mismo “Código”, en los conocidos términos de Lessig. Se trata de la filosofía del *compliance* o cumplimiento normativo, de corte anglosajón y en buena medida ajena al Derecho continental europeo. No obstante, es una técnica ya conocida el sector público en ámbitos como el medio ambiente o los ya asentados estudios de impacto de género para disposiciones generales¹³⁵, que sin duda habrán de incorporar la discriminación algorítmica pública. En todo caso, especialmente este sistema de cumplimiento ha sido desarrollado en el contexto de la

¹³⁴ Sobre la aplicación del principio de responsabilidad proactiva y el diseño para el ámbito de la IA y el big data mi estudio (2022 a). Destaca, HERNÁNDEZ PEÑÁ, J. C. (2022): *El marco jurídico de la inteligencia artificial. Principios, procedimientos y estructuras de gobernanza*, Aranzadi, Cizur, pp. 119-172 y también con especial atención a la gobernanza (pp. 173-206). Cabe seguir: ALBERTO GONZÁLEZ, P. (2017): “Responsabilidad proactiva en los tratamientos masivos de datos”, *Dilemata*, nº 24, pp. 115-129; NAVAS NAVARRO, S. (2017): “Derecho e inteligencia artificial desde el diseño. Aproximaciones”, NAVAS NAVARRO, S. (coord.). *Inteligencia artificial: tecnología, derecho*, Tirant lo Blanch, Valencia, pp. 23-72; MARTÍNEZ MARTÍNEZ, R. (2019): “Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo”, *Revista catalana de dret públic*, nº 58, pp. 64-81.

¹³⁵ Impuestos a nivel estatal por artículos 22.2 para leyes y 24.1.b para reglamentos, Ley 50/1997, de 27 de noviembre, del Gobierno; artículo 19 Ley Orgánica 3/2007 para disposiciones de carácter general y los planes de especial relevancia económica y artículos 26. 3 f) Ley 40/2015. Existe abundante normativa autonómica.

protección de datos y, de ahí, se proyecta para los sistemas IA en aplicación de dicha normativa o en la futura regulación del RIA. Precisamente el “*Ethics & Rule of law by design* X-by design” es la enseña de la “IA confiable” “Made in Europe”¹³⁶ que está en la base de las políticas y normativas de IA en la UE. Pues bien, su adopción ha de ser también la viga maestra de la garantía de la evitación de los sesgos, la no discriminación y, en general, el cumplimiento normativo del uso de la IA en el sector público.

En el ámbito de la protección de datos el principio de la responsabilidad proactiva (art. 5, Considerando 78) se traduce en muchas obligaciones¹³⁷, como el registro interno de las actividades del tratamiento, identificando el análisis del riesgo en cada tratamiento (art. 30, art. 31 Ley Orgánica 3/2018, de 5 de diciembre, en conexión con el inventario de actividades en el sector público, art. 77.1º). En todo caso, hay que destacar especialmente la obligación de la protección de datos desde el diseño y por defecto (art. 25) antes del inicio del tratamiento de datos y, singularmente, las evaluaciones de impacto (art. 35 RGPD), esto es, el análisis y descripción de todas las operaciones, su necesidad y la proporcionalidad y la evaluación de los riesgos¹³⁸.

El estudio de impacto será obligatorio en casi todo supuesto de uso público de IA que suponga un tratamiento de derecho de protección de datos siguiendo los criterios de la AEPD de 2019¹³⁹.

¹³⁶ HLEG (2019), *op.cit.*, en versión 2018, pp. 19, 21, 29.

¹³⁷ d como la diligente elección de un encargado, las garantías suficientes para aplicar medidas técnicas y organizativas apropiadas (art. 28) y la necesidad de decidir cuáles son las adecuadas según los riesgos (art. 32 RGPD), la designación de un delegado de protección de datos (art. 36), la notificación de una violación de la seguridad a las autoridades (art. 33) y su comunicación a los interesados (art. 34).

¹³⁸ GRUPO DEL ARTÍCULO 29 (2018), *ob. cit.* pp. 3 y ss.

¹³⁹ Afirma que se han de dar al menos dos de estos supuestos: toma de decisiones automatizadas; perfiles de comportamiento; uso de datos a gran escala; asociación, combinación o enlace de registros de bases de datos para varias finalidades o utilización de nuevas tecnologías o un uso innovador de tecnologías consolidadas, incluyendo la utilización de tecnologías a una nueva escala, con un nuevo objetivo o combinadas con otras, AEPD (2019): *Listas de tipos de tratamientos de datos que requieren EIPD* (art 35.4), <https://www.aepd.es/media/criterios/listas-dpia-es-35-4.pdf>

Sobre el cumplimiento de estas medidas cabe remitir a los instrumentos desarrollados por las autoridades europeas¹⁴⁰ y las autoridades de datos españolas, también para el ámbito concreto del sector público¹⁴¹. De hecho, en buena medida hay que remitir al conjunto de remedios que se ha hecho referencia que en muy buena medida pueden introducirse en los análisis de riesgos y estudios de impacto como medidas a adoptar para prevenir los riesgos.

A todo este arsenal de medidas preventivas, tecnológicas y de garantías habrá que añadir las especiales garantías de transparencia, explicabilidad y recurribilidad especialmente reforzadas en el caso de las decisiones automatizadas (arts. 22 y 13 RGPD). Además, hay que subrayar que con carácter general todas estas exigencias se harán más intensas en los supuestos de tratamiento de datos especialmente protegidos del artículo 9 RGPD. Como se ha dicho, ello se traducirá en un claro fortalecimiento de garantías en muchos de los supuestos de discriminaciones especialmente protegidas.

Hay que llamar la atención de que, a partir de la actual normativa y principios de la protección de datos, las autoridades de protección de datos ya han interpretado y *extraído* de forma expansiva muchas obligaciones sobre gobernanza de los datos, calidad y robustez de los sistemas de IA, evitación de sesgos y discriminación, obligación de auditorías, así como importantes obligaciones de transparencia. Y todo ello sin esperar a la aprobación de la normativa de IA. Baste seguir las

¹⁴⁰ Así, GRUPO DEL ARTÍCULO 29 (2017): *Directrices sobre la evaluación del impacto de la protección de datos (DPIA) y determinar si el procesamiento es “probable que genere un alto riesgo” a los efectos del Reglamento 2016/679, WP248 rev.01*, Revisadas por última vez y adoptadas el 4 de octubre de 2017, <https://www.aepd.es/sites/default/files/2019-09/wp248rev01-es.pdf>

¹⁴¹ AEPD (2021 b): *Gestión del riesgo y evaluación de impacto en tratamientos de datos personales*, <https://www.aepd.es/es/node/46578>; AEPD (2019): *Modelo de informe de Evaluación de Impacto en la Protección de Datos (EIPD) para Administraciones Públicas*, <https://www.aepd.es/es/node/816> ; AUTORIDAD CATALANA DE PROTECCIÓN DE DATOS (2017): *Guía práctica sobre la evaluación de impacto relativa a la protección de datos*, https://apdcat.gencat.cat/es/documentacio/guies_basiques/Guies-apdcat/Guia-sobre-la-evaluacion-de-impacto-relativa-a-la-proteccion-de-datos-en-el-RGPD/ ; AEPD (2018): *Guía práctica para las evaluaciones de impacto en la protección de datos personales*, <https://www.lopdencastellon.com/wp-content/uploads/2018/02/Gu%C3%ADa-Evaluaci%C3%B3n-de-impacto-protecci%C3%B3n-de-datos.pdf>

Guías sobre IA de la AEPD (2020 y 2021)¹⁴² o del Grupo del artículo 29 (2018) y ahora del CEPD.

Pues bien, las medidas proactivas de protección de datos y sus proyecciones para la IA serán obligatorias y exigibles en tanto en cuanto el uso público de un sistema de IA suponga un tratamiento de datos personales, que como se expuso, no será nada extraño.

11.2.- Los usos públicos de IA de alto riesgo habrán de cumplir, además, los deberes del RIA

El enfoque preventivo del cumplimiento de obligaciones según el riesgo es la base del régimen de la propuesta de Reglamento de IA de 2021¹⁴³. Y los sistemas de alto riesgo públicos que se han identificado *supra* tienen que cumplir todo el grueso de la regulación preventiva y de gobernanza¹⁴⁴. Tiene especial relación con la evitación de sesgos y discriminación las obligaciones de evaluación y mitigación de riesgos (art. 9 RIA), así como en especial las obligaciones de utilizar conjuntos de datos de alta calidad con buena gobernanza (art. 10 y 16 RIA). Se parte de que los datos “deben estar tan libres de errores y ser tan completos como sea posible” (Considerando 44). Se regula la concreta obligación de hacer un “examen a la vista de posibles sesgos que puedan afectar a la salud y la seguridad de las personas físicas o dar lugar a discriminaciones prohibidas” (art. 10. 2º f). También en el RIA hay conjuntos de obligaciones de los desarrolladores de generar documentación técnica para los clientes o usuarios del sistema (art. 11), el diseño de la IA para que genere registros y logs de funcionamiento (art. 12). Estas obligaciones, además de demostrar el cumplimiento de los deberes, pueden ser decisivas para la comprobación *ex post* de los

¹⁴² AEPD (2020 a): *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*. <https://www.aepd.es/es/media/guias/adequacion-rgpd-ia.pdf>, en especial, pp. 30-48 y, con más claridad en AEPD (2021 a) *ob. cit.*

¹⁴³ Un primer análisis de este en España puede seguirse en COTINO HUESO, L. y otros (2021): “Un análisis crítico constructivo de la Propuesta de Reglamento de la Unión Europea por el que se establecen normas armonizadas sobre la Inteligencia Artificial (Artificial Intelligence Act)”, *Diario La Ley*, 2 de julio, Wolters Kluwer. <https://links.uv.es/2FK3xc4>

¹⁴⁴ En especial, HERNÁNDEZ PEÑA, J. C. (2022), *ob. cit.*

mismos. Las medidas proactivas frente a la discriminación y el sesgo pasan también en general por las obligaciones de supervisión humana (art. 14) y evitación del “sesgo de automatización”. Y especialmente los sistemas de IA públicos evitarán los sesgos y la discriminación si están diseñados con un nivel adecuado de precisión, solidez y ciberseguridad (art. 15). La “solidez” técnica implica obligaciones para prevenir o minimizar “errores, fallos, incoherencias, situaciones inesperadas” (Considerando 50) y la eliminación o reducción de resultados sesgados una vez comercializados (art. 15. 3º). Los sistemas habrán de seguir procedimientos de evaluación de la conformidad (art. 19) y una vez en funcionamiento se someten al sistema de vigilancia del mercado y de supervisión. Todas estas exigencias serán desarrolladas además por todo un armazón de normas técnicas y de armonización, que serán acompañadas de la normalización técnica de la IA que está en fase de desarrollo.

Previsiblemente dos tercios de los sistemas de IA públicos no serán de alto riesgo¹⁴⁵, por lo no estarán obligados a adoptar estas medidas preventivas. Ahora bien, estas medidas serán los referentes de las buenas prácticas a seguir voluntariamente en códigos, sellos o certificaciones a las que se sumen (Considerando 81, art. 69) los proveedores públicos o privados de esos sistemas públicos de IA. De igual modo, estas medidas podrán integrarse en las que hay que adoptar por protección de datos o en razón de las medidas que se exijan por no discriminación, por ejemplo, en los estudios de impacto de género.

Pues bien, el cumplimiento de todas estas obligaciones de responsabilidad proactiva acaba secuenciándose en toda una serie de medidas a adoptar en las diversas etapas de interés: al momento de concebir y diseñar el sistema de IA público, o al momento de preparar los pliegos y mecanismos de contratación de este. También durante el desarrollo o la adquisición del sistema y antes de su despliegue, momento en el que el entrenamiento del sistema, la comprobación de errores y sesgos o la seguridad de este es especialmente relevante. Tras el despliegue del sistema cobra especial protagonismo garantizar la intervención humana y la supervisión de los resultados y la reevaluación de su fun-

¹⁴⁵ Ver supra, JOINT RESEARCH CENTRE (2022), *ob. cit.* p. 58 Cabe recordar que también en principio, sólo un 10% de sistemas de IA serán de alto riesgo.

cionamiento, así como la garantía de los derechos de los interesados afectados de acceso de información y frente a posibles vulneraciones de derechos¹⁴⁶.

11.3.- *La extensión de estudios de impacto y la obligatoriedad de medidas para la no discriminación de los sistemas de IA públicos*

La FRA fue de las primeras instituciones en prestar especial atención al fenómeno de la discriminación algorítmica. En su estudio de 2021, su dictamen n° 2 invita a que el legislador de la Unión haga obligatorias las evaluaciones de impacto que cubran toda la variedad de derechos fundamentales¹⁴⁷, en un sentido similar, el European Law Institute¹⁴⁸. Como en otros lugares he señalado¹⁴⁹, siguiendo especialmente a Mantelero¹⁵⁰ el estudio de impacto no ha de ser exclusivo de la protección de datos, sino a todos los derechos digitales y, por supuesto, para el caso de la igualdad, no discriminación y evitación del sesgo. La necesidad de implantar instrumentos preventivos y del estudio de impacto para la igualdad se percibe con mucha claridad en

¹⁴⁶ Un buen ejemplo a seguir lo brinda en especial CEPD (2022): *Guidelines 05/2022 on the use of facial recognition technology in the area of law enforcement*, Version 1.0, 12 mayo, https://edpb.europa.eu/our-work-tools/documents/public-consultations/2022/guidelines-052022-use-facial-recognition_en Ver, Anexo II, pp. 29-35.

¹⁴⁷ Ver todas las referencias a los diversos trabajos. En particular ahora FRA(2021): *Getting the future right artificial intelligence and fundamental rights*, FRA, <https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights>

¹⁴⁸ EUROPEAN LAW INSTITUTE (2022): *Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration*, European Law Institute-University of Viena, https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Model_Rules_on_Impact_Assessment_of_ADMSs_Used_by_Public_Administration.pdf

¹⁴⁹ COTINO HUESO, L. (2022 a), *ob. cit.*

¹⁵⁰ Son muchas las obras del máximo nivel del autor desde hace años, puede seguirse en su reciente monografía MANTELERO, A. (2022): *Beyond Data. Human Rights, Ethical and Social Impact Assessment*, Springer, Information Technology and Law Series IT&LAW 36, 2022, <https://link.springer.com/book/10.1007/978-94-6265-531-7>

Recomendación Unesco de 2021¹⁵¹, especialmente en el capítulo dedicado a la “Equidad y no discriminación” (nº 28 y ss.)¹⁵²; en el ámbito de actuación 1º de “evaluación del impacto ético” (nº 50-53)¹⁵³ y el ámbito de actuación 6 de “género” (nº 87-93) el nº87¹⁵⁴.

De modo más concreto, la evitación de sesgos y de discriminación se integran en los marcos de gestión de riesgos que para la IA se están estandarizando, por ejemplo, desde el NIST en 2022¹⁵⁵. Para el sector público cabe seguir grandes avances como en Países Bajos, ya referenciados y especialmente incisivos respecto de la no discriminación. Hay que estimular que en España se sigan estas buenas prácticas.

La misma AEPD respecto de los estudios de impacto de protección de datos (que serán por lo general obligatorios en los usos de IA pública) recuerda la necesidad de atender a otros derechos, así como específicamente a la no discriminación. Pues bien, como se expuso, los sistemas de IA públicos que hagan diferencias de trato por circunstancias especialmente sospechosas, por lo general supondrán tratamientos de datos especialmente protegidos. Y en estos supuestos los estudios de impacto han de ser especialmente incisivos en la detección de riesgos y las medidas compensatorias frente a los mismos,

¹⁵¹ Así, todos los derechos que se han de tener en cuenta “a lo largo del ciclo de vida de los sistemas de IA” (nº 16) y en particular “la protección y la promoción de la diversidad y la inclusión” con especial “participación activa de todas las personas o grupos” (nº 19).

¹⁵² Los actores de la IA deberían hacer todo lo razonablemente posible por reducir al mínimo y evitar reforzar o perpetuar aplicaciones y resultados discriminatorios o sesgados a lo largo del ciclo de vida de los sistemas de IA, a fin de garantizar la equidad de dichos sistemas.” (nº29), también “abordar las brechas digital y de conocimientos” (nº 30).

¹⁵³ “Los gobiernos deberían adoptar un marco regulador que establezca un procedimiento para que las autoridades públicas, en particular, lleven a cabo evaluaciones del impacto ético de los sistemas de IA [...] Se debería exigir a las autoridades públicas que supervisen los sistemas de IA que hayan implantado o desplegado, mediante la introducción de mecanismos e instrumentos adecuados.” (nº 53).

¹⁵⁴ El nº87 insiste en que la evaluación del impacto ético debería incluir una perspectiva transversal de género. Y que los estereotipos de género y los sesgos discriminatorios no se trasladen a los sistemas de IA, sino que se detecten y corrijan de manera proactiva.” (nº 90).

¹⁵⁵ NIST (2022): *AI Risk Management Framework: Initial Draft*, 17 de marzo, <https://www.nist.gov/document/ai-risk-management-framework-initial-draft>

habiendo de integrar en la medida de lo posible las buenas prácticas y recomendaciones habitualmente señaladas. Es más, en estos supuestos, se hace precisa una regulación legal y con garantías para que la Administración efectúe estos tratamientos de datos. Por ello, los instrumentos legales habrán de integrar y concretar las garantías para evitar el sesgo, error y discriminación en los sistemas públicos de IA. De este modo, no se tratará de la adopción voluntaria de buenas prácticas y recomendaciones.

Y ahí es donde han de cobrar especial importancia los estudios de impacto de género obligatorios para disposiciones generales, conectados con los muchos supuestos en los que las diferencias de trato y los sistemas de IA deben ser regulados por ley o reglamento. Hoy por hoy, las herramientas y guías para la realización de estos estudios de impacto normativo¹⁵⁶ no incluyen las cuestiones que aquí interesan, pero sin duda que se van a ir generalizando e incluyendo estos problemas debido a la evolución y los cambios normativos recientes.

En España, además de las actuales exigencias por la protección de datos, es del todo acertada la Carta de Derechos Digitales que afirma en general “el principio de cumplimiento normativo desde el diseño” para todos los ámbitos en su apartado I. 4 ¹⁵⁷. Ya para la IA hace expresa mención a “garantizar el derecho a la no discriminación

¹⁵⁶ Entre otros, cabe seguir, DELGADO GOBOY y otras (2003): *Guía de aplicación práctica, para la elaboración de informes de impacto de género de las disposiciones normativas que elabore el gobierno*, de acuerdo a la ley 30/2003, 2003, Instituto De La Mujer (Ministerio de Trabajo Y Asuntos Sociales), <https://www.inmujeres.gob.es/publicacioneselectronicas/documentacion/Documentos/DE0259.pdf>

LIKADI (2018): *Guía para la evaluación de impacto de género de la normativa guía para la evaluación de impacto de género de la normativa*, Diputación Foral de Álava, 2018, <https://web.araba.eus/documents/105044/1393927/Gu%C3%ADa+para+la+evaluaci%C3%B3n+de+impacto+de+g%C3%A9nero+%C3%81lava+castellano+navegable.pdf/74905ae0-fc50-7b15-a5c3-2ab4b2f16fa3?t=1557493191803>

Pueden seguirse otras en <https://presidencia.gva.es/es/guias-para-elaborar-informes-de-impacto-de-genero>

¹⁵⁷ “declara que el principio de cumplimiento normativo desde el diseño es de aplicación íntegramente al desarrollo de los entornos digitales y por ello los desarrollos científicos, tecnológicos y su despliegue contemplarán en la determinación de sus requerimientos un análisis sobre el cumplimiento de tal principio.”

cualquiera que fuera su origen, causa o naturaleza, en relación con las decisiones, uso de datos y procesos basados en inteligencia artificial” (apartado V). Y para el sector público se hace referencia a la “no discriminación, en particular por razón de sexo, de las tecnologías usadas por las Administraciones públicas” (XVIII. 3º) y, por lo que más interesa, que “Será necesaria una evaluación de impacto en los derechos digitales en el diseño de los algoritmos en el caso de adopción de decisiones automatizadas o semiautomatizadas.” (XVIII. 7º). Además, las evaluaciones han de alcanzar las leyes administrativas (apartado VII. 4º).

Como primera proyección el artículo 25. 1º Ley 15/2022, de 12 de julio afirma que “1. La protección frente a la discriminación obliga a la aplicación de métodos o instrumentos suficientes para su detección, la adopción de medidas preventivas”. Y el nuevo artículo 23 Ley 15/2022, de 12 de julio es el mejor exponente cuando afirma “los algoritmos involucrados en la toma de decisiones que se utilicen en las administraciones públicas tengan en cuenta criterios de minimización de sesgos, transparencia y rendición de cuentas, siempre que sea factible técnicamente. En estos mecanismos se incluirán su diseño y datos de entrenamiento, y abordarán su potencial impacto discriminatorio. Para lograr este fin, se promoverá la realización de evaluaciones de impacto que determinen el posible sesgo discriminatorio” (1º) y “priorizarán la transparencia en el diseño y la implementación y la capacidad de interpretación de las decisiones adoptadas por los mismos.” (2º).

La nueva ley emplea términos muy laxos (“favorecerán”, “tengan en cuenta”, “siempre que sea factible”, “se promoverá”, “priorizarán”, “promoverán”). Sin embargo, es muy posible que la necesidad de regular el uso de IA público que aquí se sostiene dé protagonismo práctico a los estudios de impacto que sí que son obligatorios. Dado que estos estudios de impacto normativo de género en el futuro van a ser sensibles e irán incluyendo estas cuestiones, especialmente para el sector público, es muy posible que generen un importante efecto preventivo de la discriminación algorítmica en la práctica. Bien es cierto que no se trataría estrictamente de una obligación hasta que haya una mejora regulatoria o hasta que las autoridades regulatorias competentes o judiciales hagan una interpretación más activista de las normas.

En todo caso, siempre *quedará el París* de la obligatoriedad de medidas proactivas por la protección de datos, con su interpretación claramente expansiva y proyectable para el sector público. Y también, todos los deberes del RIA para los usos públicos de la IA de alto riesgo.

11.4.- La necesidad de una visión social y colectiva y de integrar garantías de calidad de datos y de participación

Las medidas proactivas y preventivas no sólo deben tener un enfoque en el individuo, sino en la prevención de la igualdad y la no discriminación desde el interés de los colectivos, la sociedad y el sistema democrático en su conjunto. No olvidemos el impacto estructural y masivo del uso de estos sistemas. Además, muchos sistemas IA públicos se utilizan para apoyar la toma de decisiones generales políticas o normativas. En estos supuestos, aunque los sistemas no manejen datos personales, también habrán de darse exigencias de los principios de calidad y buena gobernanza y gestión respecto de los datos no personales. Asimismo, habrá de asegurarse la particular transparencia y la participación de la sociedad civil en estos contextos de uso. Ello puede articularse en el marco de procedimientos participativos, de audiencia pública que se adopten o su inserción en órganos colegiados. Así lo propuse para la Carta de Derechos Digitales en un apartado sobre “Garantías de impacto social”¹⁵⁸ y de nuevo resulta especialmente inspirador Mantelero¹⁵⁹. La Recomendación Unesco es llamativamente insistente en la necesidad de un enfoque colectivo y sobre todo de la inclusión de mecanismos de participación social respecto del uso de sistemas IA¹⁶⁰. Para la IA pública, también el European Law

¹⁵⁸ Ver mi aportación en RED DAIA (2020 b): *Aportación DAIA a consulta pública Carta de derechos digitales*, diciembre 2020 <https://bit.ly/3paF0H0>

¹⁵⁹ Además de la obra reseñada, por ejemplo, (2017): “From group privacy to collective privacy: towards a new dimension of privacy and data protection in the big data era”, en TAYLOR, L.; VAN DER SLOOT, B.; FLORIDI, L. (eds.) (2017), ob. cit. cap. 8.

¹⁶⁰ Así en el apartado específico de “Equidad y no discriminación” se menciona la “participación inclusiva de colectivos” (nº 28) y la “participación significativa” de la sociedad en la evitación de brechas a lo largo del ciclo de vida de la IA (nº

Institute hace interesantes propuestas de participación en estudios de impacto (art. 6. 2 c) ii), auditorías (art. 10) y participación pública (art. 11).

Recuerda Gutiérrez las Recomendaciones de la APDCAT de publicar la EIPD relativas a los tratamientos que son responsabilidad de las Administraciones públicas, pues “resulta de especial relevancia cuando el interés del tratamiento puede llegar a suponer un cierto impacto social o puede generar un estado de opinión en las personas los datos de las cuales pueden llegar a ser objeto de tratamiento”¹⁶¹.

12.- Y SI LOS REMEDIOS NO VALEN, ¿QUÉ PASA CUANDO SE DETECTA UNA DISCRIMINACIÓN ALGORÍTMICA PÚBLICA?

12.1.- *Obligaciones de mitigación, rectificación, corrección e incluso cese de uso del sistema de IA público*

Según se ha expuesto, de modo proactivo y con carácter general los proveedores o usuarios de sistemas de IA públicos han de evaluarlos con pruebas y mitigación de sesgos de forma continua y con supervisión organizativa. Las evaluaciones en su caso han de ser independientes. De modo reactivo, si se detecta un sistema de IA público sesgado o discriminatorio, se deben adoptar medidas de mitigación, modulación, corrección o eliminación de las causas. En su caso, habrá que bloquear o cesar el uso del sistema. Como ahora se sigue, estas obligaciones son claras desde la normativa de igualdad, de procedimiento Administrativo, de protección de datos y el RIA.

30). En los números 46 y 47 sobre “Gobernanza y colaboración adaptativas y de múltiples partes” se incide en que “la participación de las diferentes partes interesadas a lo largo del ciclo de vida de los sistemas de IA es necesaria para garantizar enfoques inclusivos de la gobernanza de la IA”. Y la participación ciudadana debe facilitarse respecto de la evaluación del impacto ético (nº 50, 52, 53), lo mismo que se reitera en los números 60 y 64 o en el nº 91 respecto la participación en el ámbito de actuación 6 de género.

¹⁶¹

AUTORIDAD CATALANA DE PROTECCIÓN DE DATOS (2017), *ob. cit.*

Desde la normativa de igualdad, cabe recordar que el artículo 25. 1º Ley 15/2022 además de las medidas preventivas y de detección, “obliga” a “la articulación de medidas adecuadas para el cese de las situaciones discriminatorias”, asimismo, “3. Ante un incidente discriminatorio, las autoridades encargadas de hacer cumplir esta ley tomarán las medidas oportunas para garantizar que los hechos no vuelvan a repetirse, especialmente en los casos en los que el agente discriminador sea una administración pública.” En el mismo sentido, se han de “adoptar las medidas oportunas y proporcionadas para su eliminación” (art. 31.1º). En razón del artículo 31 habrá de incoarse el procedimiento Administrativo correspondiente y la investigación. Deben recordarse además las particularidades que puedan darse en sectores públicos específicos regulados en los artículos 9 a 24, como en la educación (art. 13), atención sanitaria (art. 15), en la prestación de los servicios sociales (art. 16), en la oferta al público de bienes y servicios (art. 17), en el ámbito de la seguridad ciudadana (art. 18), en la administración de justicia (art. 19), o en culturales y deportivas (art. 23).

En el ámbito del procedimiento, si se considera que el sesgo del sistema de IA no tiene suficiente relevancia, la detección y corrección del sistema y de sus consecuencias podría articularse a través del artículo 109. 2º LPAC. Así, podría considerarse “rectificar en cualquier momento, de oficio o a instancia de los interesados, los errores materiales, de hecho o aritméticos existentes en sus actos”. También, ante la detección o denuncia de un sesgo o discriminación algorítmica en el procedimiento, incluso antes del inicio formal del mismo, no será extraña la adopción de medidas provisionales (art. 56 LPAC) e incluso la posibilidad de la suspensión de actos (art. 108 LPAC). La denuncia por el afectado o interesado del sistema de IA podrá implicar la suspensión de la ejecución del acto o actuación impugnada, por tratarse de una posible causa de nulidad de pleno por discriminación, bien por los perjuicios de imposible o difícil reparación (art. 117. 2º LPAC).

En el ámbito de protección de datos, los posibles incumplimientos bien pueden estar relacionados con la discriminación o sesgo algorítmicos. El responsable y el encargado del tratamiento tienen deberes de asegurar el cumplimiento de las obligaciones que les corresponden y de actuar en caso de que se detecten deficiencias. Y, obviamente, ante estas situaciones, denuncias de terceros o por la actuación misma de las autoridades de control, deberán cumplir las órdenes de las

autoridades (art. 58 RGPD). Ello puede suponer la paralización del sistema de IA público: las “medidas provisionales y de garantía de los derechos” pueden incluir el “bloqueo de los datos y la cesación de su tratamiento”, incluso “su inmovilización” (art. 69 Ley orgánica 3/2018), igualmente la resolución de una infracción incluirá las medidas “para que cese la conducta o se corrijan los efectos de la infracción” (art. 77.2º Ley orgánica 3/2018).

Bajo el RIA, los proveedores o usuarios de sistemas IA públicos de “alto riesgo” están obligados a adoptar “medidas correctoras” en cuanto detecten que el sistema no cumple con el RIA (art. 16); “investigarán inmediatamente, en su caso, las causas” (art. 21). También informarán a los organismos y autoridades competentes (art. 22). La autoridad de vigilancia del mercado que aprecie un riesgo para la igualdad del sistema de alto riesgo tendrá que evaluar el sistema y “exigirá sin demora indebida al agente económico correspondiente que adopte todas las medidas correctoras adecuadas”, en su caso, “todas las medidas provisionales adecuadas para prohibir o restringir la comercialización del sistema de IA en su mercado nacional, retirar el producto de ese mercado o recuperarlo”. También habrá de informar a las autoridades competentes, incluso advertir a la Comisión y otros Estados si el riesgo va más allá de España (art. 65 RIA).

12.2.- Nulidad, anulabilidad e inderogabilidad de los sistemas de IA público y sus efectos

En el caso de una discriminación algorítmica de relevancia, hay que partir de que “son nulos de pleno derecho las disposiciones, actos o cláusulas de los negocios jurídicos que constituyan o causen discriminación” en el caso de las discriminaciones especialmente prohibidas (art. 26, Ley 15/2022, de 12 de julio). En la misma línea y como principio, se trataría de causa de nulidad de pleno derecho del artículo 47. 1. a) LPAC, de 1 de octubre por cuanto los sistemas IA “lesionen los derechos y libertades susceptibles de amparo constitucional”¹⁶².

¹⁶² Sobre el tema, en general. GAMERO, E. y FERNÁNDEZ, S. (2016), *ob. cit.* pp. 532 y ss. y MELLADO RUIZ, L. (2017), “Supuestos de nulidad de pleno derecho de los actos administrativos”, en GAMERO CASADO, E. (dir.), *Tratado de Procedimiento Administrativo Común y Régimen Jurídico Básico del sector*

En el caso de discriminaciones y sesgos no vinculados con especiales causas de no discriminación, no hay que excluir la lesión del artículo 14 CE y la consiguiente nulidad radical. Según se ha visto, en no pocos supuestos el sistema de IA público ha de venir regulado en leyes o en su caso disposiciones generales. Si la causa de la discriminación algorítmica está establecida normativamente ello provocaría la nulidad de pleno Derecho de la norma (art. 47. 2º LPAC). También puede darse el caso de la nulidad de las resoluciones de adopción del sistema u otros instrumentos en los que se haya formalizado la causa de la discriminación. En los casos en los que las medidas y decisiones técnicas en el diseño, implantación o aplicación del sistema se separen de las normas que lo regulan, se dará la inderogabilidad singular del artículo 37 LPAC.

Por cuanto a los efectos que genere el sistema público de IA discriminatorio, en su caso y dependiendo de diversos factores, puede ser aplicable la anulabilidad del artículo 48 LPAC. Posiblemente se aplicará a supuestos de discriminación no especialmente prohibida, en estos casos por lo general la discriminación convergerá con vicios de arbitrariedad y no motivación, por lo que es posible que se canalicen en mayor medida como anulables. Asimismo, es bastante posible la anulabilidad en el caso de considerarse que hay una vulneración del artículo 18 CE, que rara vez se considera que genere nulidad y como Valero subraya desde siempre, no hay que obviar que el artículo 18. 4º CE no sólo es relativo a la protección de datos sino a otro tipo de amenazas de la “informática”¹⁶³.

público Vol. 2, Tirant lo Blanch, Valencia, pp. 1929-1983; JORDANO FRAGA, J. (1996): “Nulidad de los actos administrativos y derechos fundamentales, Marcial Pons, Madrid y en (2011): “La nulidad de los actos que lesionen derechos fundamentales: evolución jurisprudencial y desarrollo de la causa de nulidad, núcleo de las garantías en el Estado de Derechos Fundamentales”, en LÓPEZ MENUDO F. (coord.), *Derechos y garantías del ciudadano: estudios en homenaje al profesor Alfonso Pérez Moreno*, 2011, pp. 715-743.

¹⁶³ Se trata de un tema complejo, estudiado especialmente por Valero en diversos de sus estudios como VALERO TORRIJOS, J. (2013): *Derecho, innovación y administración electrónica*, Global Law = Derecho Global, Sevilla. O, FERNÁNDEZ SALMERÓN M. (2004): *La protección de los datos personales en las Administraciones Públicas*, Aranzadi, Cizur, pp. 465 a 470.

Es muy posible también que haya de aplicarse el artículo 49 LPAC en los casos en los que la IA pública sólo se utiliza en partes del procedimiento que puedan considerarse independientes. En consecuencia, la “nulidad o anulabilidad de un acto no implicará la de los sucesivos en el procedimiento que sean independientes del primero” (art. 49. 1º). En todo caso, habrá que valorar si la discriminación producida, esto es “la parte viciada sea de tal importancia que sin ella el acto administrativo no hubiera sido dictado” (art. 49. 2º). Además, la detección de un fallo en el sistema de IA público utilizado podría suponer una revisión de disposiciones y actos nulos de oficio o a solicitud de los interesados (art. 106 LPAC) o provocar una revocación de actos (art. 109. 1º) o una revisión de actos nulos del artículo 106.

12.3.- Los efectos de la discriminación algorítmica en el contexto de la contratación administrativa

Cabe remitir a esta obra y otros estudios respecto de la contratación por el sector público del desarrollo o uso de sistemas tecnológicos y de IA¹⁶⁴. En EEUU recientemente se recomiendan los desarrollos *in-house*¹⁶⁵. En cualquier caso, cabe señalar que la Administración contratante del sistema de IA cuenta con generales facultades de inspección (art. 190 Ley 9/2017, de 8 de noviembre). Habrá que estar al marco del contrato específico y su modalidad, pero si el sistema de IA

¹⁶⁴ Además del estudio en esta obra sobre el tema interesa LOZA CORERA, M. (2022): “Explotación y regulación del uso del big data e inteligencia artificial para los servicios públicos y la ciudad inteligente”, en COTINO HUESO, L. y TODOLÍ SIGNES, A. (coord.), *ob. cit.* pp. 125-148. Sobre la contratación de IA policial, JOH, E. E. (2017): “The Undue Influence of Surveillance Technology Companies on Policing” (27 febrero). *N.Y.U. L. Review Online*, pp. 18-20 <https://ssrn.com/abstract=2924620> Muchos de los problemas proyectables para la IA los aborda VALERO TORRIJOS, J. (2015): “Ciudades inteligentes y datos abiertos: implicaciones jurídicas para la protección de los datos de carácter personal”, *Istituzioni del federalismo: rivista di studi giuridici e politici*, n.º 4, 2015 (“Smart Cities” e amministrazioni intelligenti), pp. 1025-1047 y VESTRI, G. (2021): “Sistemas de inteligencia artificial en la contratación pública: entre códigos fuente y datos abiertos”, *Actualidad administrativa*, n.º 12. CONSEJO DE EUROPA (2018), *op.cit.*, pp. 57-58 insiste en la necesidad de participar desde una fase temprana en la adquisición de sistemas de IA público.

¹⁶⁵ FREEMAN ENGSTROM, D. y otros (2020), *op.cit.*, pp. 7, 18, 71 y ss.

desarrollado genera sesgos y discriminación puede darse un supuesto de incumplimiento parcial o cumplimiento defectuoso (art. 192). En el caso habitual de que se articule a través de un contrato de prestación de servicios, “2. El contratista será responsable de la calidad técnica de los trabajos que desarrolle y de las prestaciones y servicios realizados, así como de las consecuencias que se deduzcan para la Administración o para terceros de las omisiones, errores, métodos inadecuados o conclusiones incorrectas en la ejecución del contrato.” Si se considera que no se ajusta a lo que corresponde, se podrá rechazar o reclamar la subsanación de los mismos (art. 311). Si el uso del sistema de IA se inserta en un contrato de concesión de servicios, se puede considerar que del funcionamiento del sistema de IA “se derivase perturbación grave y no reparable por otros medios en el servicio” y, en consecuencia, podría intervenir o secuestrar el contrato (art. 293).

Que el sistema desarrollado genere sesgos o discriminación puede tener diversas consecuencias, como penalidades, daños y perjuicios al proveedor del sistema (art. 194). La exigencia de responsabilidad podrá incluir a los terceros, como lo serían los afectados por el sistema de IA (art. 196). La detección de sesgos y discriminación en su caso podría suscitar una modificación del contrato (arts. 203 y ss. en general), incluso podría darse la resolución del contrato en algunos casos, como por ejemplo por supresión del servicio por razones de interés público (art. 294 d).

12.4.- Responsabilidad por el uso de IA público

La discriminación causa responsabilidad patrimonial y la obligación de reparación del daño y cabe recordar que “Acreditada la discriminación se presumirá la existencia de daño moral” (art. 27 Ley 15/2022, de 12 de julio) y cabe acudir al régimen de la responsabilidad patrimonial de las Administraciones Públicas (arts. 32 y ss. LRJSP)¹⁶⁶. La normativa propia del sector público muy posiblemente habría de

¹⁶⁶ Cabe remitirse al estudio sobre la responsabilidad en esta obra. La literatura en general ya es muy abundante. No obstante, para el ámbito del sector público, una primera aproximación en ALCOLEA AZCÁRRAGA, C., “La responsabilidad patrimonial de la Administración y el uso de algoritmos”, *Revista General de Derecho Administrativo*, n.º. 59, 2022.

conjugarse con el régimen más general de responsabilidad por uso de IA. Cabe ahora simplemente recordar que se trata de un tema de extraordinaria complejidad y actualidad, al punto de que el 28 de septiembre de 2022 se ha lanzado la Propuesta de Directiva relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial (Directiva sobre responsabilidad en materia de IA)¹⁶⁷. La misma es especialmente destacable en materia de carga de la prueba y presunción de responsabilidad, en particular su artículo 4. En el mismo se tienen especialmente en cuenta el cumplimiento del deber de diligencia que sea aplicable, así como si el sistema es o no de alto riesgo en el marco del RIA.

13.- ¿QUIÉNES PUEDEN REACCIONAR ANTE UNA DISCRIMINACIÓN ALGORÍTMICA Y QUÉ AUTORIDADES DEBEN ACTUAR?

En el caso de la posible discriminación por un sistema de IA público, por una parte, los interesados y potenciales reclamantes serán los titulares del derecho de igualdad o los derechos y los intereses legítimos afectados por el uso del sistema de IA público (art. 4 LPAC). Ahora bien, por otra parte, lo más normal será que haya una “pluralidad indeterminada o de difícil determinación” de personas afectadas y, en consecuencia, se dará el presupuesto por el que serán interesados amplios sujetos de la sociedad civil, además, en la mayoría de los casos sin necesidad de autorización de los afectados individualmente. Ello es así debido a la previsión expresa del artículo 31. Ley 15/2022, de 12 de julio¹⁶⁸. Y en el contexto de la protección de datos por el

¹⁶⁷ <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52022PC0496>
https://ec.europa.eu/info/business-economy-euro/doing-business-eu/contract-rules/digital-contracts/liability-rules-artificial-intelligence_en

¹⁶⁸ “2. A los efectos de lo establecido en el artículo 4.2 de la LPAC, los sindicatos, las asociaciones profesionales de trabajadores autónomos, las organizaciones de personas consumidoras y usuarias y las asociaciones y organizaciones legalmente constituidas que tengan entre sus fines la defensa y promoción de los derechos humanos y cumplan los requisitos fijados en el artículo 29 de la presente ley, podrán tener la consideración de interesado en los procedimientos administrativos en los que la Administración tenga que pronunciarse en relación con una situación de discriminación prevista en esta ley, siempre que cuenten con la auto-

que se puede vehicular una acción, cabe recordar que además de los interesados son posibles las acciones colectivas en virtud del artículo 80 RGPD. Este precepto ha sido recientemente interpretado por la STJUE de 28 de abril de 2022, asunto C-319/20. Asimismo y en su caso, puede llegar a ser aplicable la Directiva (UE) 2020/1828, de 25 de noviembre que regula también las acciones colectiva y su transposición en España. En los supuestos en los que el sistema de IA público venga regulado en una disposición general, cualquier persona podrá impugnar directamente la misma (art. 26 LJCA).

Por cuanto a las autoridades que puedan actuar potencialmente, son muchas debido al tipo de Administración que use el sistema de IA, la finalidad de uso, efectos y muchos otros factores. En todo caso cabe destacar las autoridades de protección de datos, supervisión de algoritmos y de igualdad con las facultades de actuación cuando se detecta un incumplimiento normativo.

Dada la fácil conexión del uso de IA pública con la aplicación del régimen de protección de datos, será muy fácil que de oficio o por denuncia actúen las autoridades estatal o autonómicas. Asimismo, la Agencia Española de Supervisión de Inteligencia Artificial (AESIA) ya prevista legalmente¹⁶⁹ en principio está pensada para la asunción rápida de todas las obligaciones que imponga del RIA (artículo 59 RIA). Según su regulación definitiva, bien se le puede facultar para asumir competencias más allá de los sistemas de alto riesgo, como muchos de los sistemas de IA públicos.

También, la generación de sesgos y discriminaciones por el sector público será de natural el ámbito de actuación de la Autoridad Independiente para la Igualdad de Trato y la No Discriminación creada por el artículo 40 Ley 15/2022, de 12 de julio, con facultades de

rización de la persona o personas afectadas. No será necesaria esta autorización cuando las personas afectadas sean una pluralidad indeterminada o de difícil determinación, sin perjuicio de que quienes se consideren afectados puedan también participar en el procedimiento.”

¹⁶⁹ La Ley 22/2021, de 28 de diciembre, de Presupuestos Generales del Estado para el año 2022 autorizó al Gobierno a crear la AESIA para llevar a “cabo medidas destinadas a la minimización de riesgos significativos sobre la seguridad y salud de las personas, así como sobre sus derechos fundamentales, que puedan derivarse del uso de sistemas de inteligencia artificial”.

“c) Iniciar, de oficio o instancia de terceros, investigaciones sobre la existencia de posibles situaciones de discriminación”, “d) Ejercitar acciones judiciales”, “e) Interesar la actuación de la Administración del Estado para sancionar las acciones u omisiones”. Como recuerda el Consejo de Europa, los organismos de igualdad, entre otras posibilidades, deben exigir que el sector público se asegure de que cuenta con suficientes conocimientos jurídicos y técnicos para evaluar y supervisar los riesgos. Asimismo, exigir que se evalúen periódicamente si sus sistemas de IA tienen efectos discriminatorios¹⁷⁰.

BIBLIOGRAFÍA CITADA

- ABA CATOIRA, A. (2022), “Derechos de igualdad, personas con discapacidad y mayores en el entorno digital (VIII, XI Y XII)”, COTINO HUESO, L. (editor), *La Carta de Derechos Digitales*, Tirant Lo Blanch, Valencia, pp. 123-154.
- AEPD (2017), *Informe AEPD 385661/2017 de marzo relativo al anteproyecto de ley de sistema de alertas tempranas de la Comunidad Valenciana*. No publicado. He dipuesto un enlace en <https://www.dropbox.com/s/ut-zo5hrkf0b3cmi/aepdantifraude.pdf?dl=0>
- AEPD (2018): *Guía práctica para las evaluaciones de impacto en la protección de datos personales*, <https://www.lpdencastellon.com/wp-content/uploads/2018/02/Gu%C3%ADa-Evaluaci%C3%B3n-de-impacto-protecci%C3%B3n-de-datos.pdf>
- AEPD (2019): *Listas de tipos de tratamientos de datos que requieren EIPD (art 35.4)*, <https://www.aepd.es/media/criterios/listas-dpia-es-35-4.pdf>
- AEPD (2019): *Modelo de informe de Evaluación de Impacto en la Protección de Datos (EIPD) para Administraciones Públicas*, <https://www.aepd.es/es/node/816>
- AEPD (2020 a): *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*. <https://www.aepd.es/es/media/guias/adequacion-rgpd-ia.pdf>
- AEPD (2020 b): *Privacidad de grupo*, 19 de octubre, <https://www.aepd.es/es/prensa-y-comunicacion/blog/privacidad-de-grupo>
- AEPD (2021 a): *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021.

¹⁷⁰ CONSEJO DE EUROPA (2018), *op.cit.*, p. 57.

- AEPD (2021 b): *Gestión del riesgo y evaluación de impacto en tratamientos de datos personales*, <https://www.aepd.es/es/node/46578>
- ALBERTO GONZÁLEZ, P. (2017): “Responsabilidad proactiva en los tratamientos masivos de datos”, *Dilemata*, n° 24, pp. 115-129.
- ALCOLEA AZCÁRRAGA, C. (2022): “La responsabilidad patrimonial de la Administración y el uso de algoritmos”, *Revista General de Derecho Administrativo*, n°. 59.
- ALGEMENE REKENKAMER (Tribunal de Cuentas), (2020): *Digitaal Toetsingskader Algoritmes (Marco de evaluación digital de Algoritmos)*, Tribunal de Cuentas Excel, <https://www.rekenkamer.nl/binaries/rekenkamer/documenten/publicaties/2021/01/28/download-het-toetsingskader/Toetsingskader+algoritmes+v1.0.xlsx>
- AUTORIDAD CATALANA DE PROTECCIÓN DE DATOS (2017): *Guía práctica sobre la evaluación de impacto relativa a la protección de datos*, https://apdcat.gencat.cat/es/documentacio/guies_basiques/Guies-apdcat/Guia-sobre-la-evaluacion-de-impacto-relativa-a-la-proteccion-de-datos-en-el-RGPD/
- BALAGUER CALLEJÓN, M. L. (2018): “Acciones positivas para la igualdad efectiva de mujeres y hombres”, VENTURA FRANCH, A. y GARCÍA CAMPÁ, S. (coord.) *El derecho a la igualdad efectiva de mujeres y hombres: una evaluación del primer decenio de la Ley Orgánica 3/2007*, pp. 285-312.
- BAROCAS, S. y SELBST, A. D. (2016): “Big Data’s disparate impact”. 104 *Calif Law Rev* 671, <http://dx.doi.org/10.2139/ssrn.2477899>
- BARONA, S. (2021): *Algoritmización del derecho y de la justicia: De la inteligencia artificial a la Smart Justice*, Tirant lo Blanch, Valencia.
- BEN- SHAHAR, O. y PORAT, A. (2021): *Personalized law : different rules for different people*, Oxford University Press.
- BENÍTEZ PALMA, E. (2022): “Mucho ruido y pocas nueces: algunas consideraciones sobre los registros públicos de algoritmos”, VESTRI, G. (2022) *ob. cit.* pp. 83-96.
- BID- Universidad Adolfo Ibáñez (2022): *Formulación ética de proyectos de ciencia de datos*, Banco Interamericano de Desarrollo, División de Gobierno Digital y Universidad Adolfo Ibáñez, <https://goblab.uai.cl/guia-formulacion-etica-de-proyectos-de-ciencia-de-datos/>
- BOGEN, M. (2019): “All the Ways Hiring Algorithms Can Introduce Bias”. *Harvard Business Review*. Mayo. <https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias>
- BOIX PALOP, A. (2020): “Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la administración para la adopción de decisiones”, en

- Revista de Derecho Público: Teoría y Método*, Vol. 1, 2020. https://doi.org/10.37417/RPD/vol_1_2020_33
- CEPD (2022): *Guidelines 05/2022 on the use of facial recognition technology in the area of law enforcement*, Version 1.0, 12 mayo, https://edpb.europa.eu/our-work-tools/documents/public-consultations/2022/guidelines-052022-use-facial-recognition_en
- CERRILLO I MARTÍNEZ, A. (2022): “La personalización de servicios digitales”, en CERRILLO I MARTÍNEZ, A. (coord.) *La Administración Digital*, Dykinson, Madrid, pp. 311-370.
- CERRILLO MARTÍNEZ, A. (2020): “¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?”, *European Review of Digital Administration & Law* Volume 1 – Issue 1-2 – 2020, DOI: 10.4399/97888255389603, págs. 18-36
- CHRISTIAN, A. (2022): *Four Ways the TSA Is Making Flying Easier for Transgender People*. American Civil Liberties Union. <https://www.aclu.org/news/lgbtq-rights/four-ways-the-tsa-is-making-flying-easier-for-transgender-people>
- CITRON, D. K. (2008): “Technological Due Process”, 85 *Wash. U. L. Rev.* pp. 1249-1313, pp. 1254 y ss. p. 1286, http://openscholarship.wustl.edu/law_lawreview/vol85/iss6/2
- COBREROS MENDAZONA, E. (2007): “Discriminación por indiferenciación: estudio y propuesta”, *Revista española de derecho constitucional*, Año nº 27, Nº 81, 2007, págs. 71-114.
- COMISIÓN EUROPEA, RENDA A. (2021): *Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe. Final Report (D5)*, abril 2021. <https://op.europa.eu/es/publication-detail/-/publication/55538b70-a638-11eb-9585-01aa75ed71a1>
- CONSEJO DE EUROPA; ZUIDERVEEN BORGESIJUS F (2018): *Discrimination, artificial intelligence, and algorithmic decision-making*. Directorate General of Democracy, Consejo de Europa. <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>
- COTINO HUESO L. y CASTELLANOS CLARAMUNT J. (coords.), *Transparencia y explicabilidad de la inteligencia artificial*, Tirant lo Blanch, Valencia, 2022.
- COTINO HUESO, L. (2017): “Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales”, en *Dilemata. Revista Internacional de Éticas Aplicadas*, nº 24, 2017. pp. 131-150, <https://goo.gl/iERVha>
- COTINO HUESO, L. (2020 a): “SyRI, ¿a quién sanciono?” Garantías frente al uso de inteligencia artificial y decisiones automatizadas en el sector público

- blico y la sentencia holandesa de febrero de 2020, en *La Ley Privacidad*, Wolters Kluwer n° 4, mayo. Acceso en Academia
- COTINO HUESO, L. (2020 b): “El alcance e interacción del régimen jurídico de los datos personales y big data relacionado con salud y la investigación biomédica”, *Revista de derecho y genoma humano: genética, biotecnología y medicina avanzada*, n° 52, pp. 57-96, n°52 enero-junio 2020. acceso
- COTINO HUESO, L. (2021 a): “Hacia la transparencia 4.0: el uso de la inteligencia artificial y big data para la lucha contra el fraude y la corrupción y las (muchas) exigencias constitucionales”, en Carles RAMIÓ (coord.), *Repensando la administración digital y la innovación pública*, Instituto Nacional de Administración Pública (INAP), Madrid, 2021. <https://links.uv.es/FUW2pz6>
- COTINO HUESO, L. (2021 b): “El nuevo reglamento de Administración electrónica, que no innova en tiempos de transformación digital”, *Revista Catalana de Derecho Público* n° 63, 2021, pp. 118-136, DOI: <http://dx.doi.org/10.2436/rcdp.i63.2021.3672>
- COTINO HUESO, L. (2022 a): “Nuevo paradigma en la garantías de los derechos fundamentales y una nueva protección de datos frente al impacto social y colectivo de la inteligencia artificial”, en COTINO HUESO, Lorenzo (editor), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Thompson-Reuters Aranzadi, FIADI (Federación Iberoamericana de Asociaciones de Derecho e Informática), Cizur.
- COTINO HUESO, L. (2022 b): “Derechos ante la administración digital y la inteligencia artificial (XVIII Y XV)”, en COTINO HUESO, Lorenzo (editor), *La Carta de Derechos Digitales*, Tirant Lo Blanch, Valencia, 2022, pp. 251-284.
- COTINO HUESO, L. (editor) (2022), *La Carta de Derechos Digitales*, Tirant Lo Blanch, Valencia.
- COTINO HUESO, L. y otros (2021): “Un análisis crítico constructivo de la Propuesta de Reglamento de la Unión Europea por el que se establecen normas armonizadas sobre la Inteligencia Artificial (Artificial Intelligence Act)”, *Diario La Ley*, 2 de julio, Wolters Kluwer. <https://links.uv.es/2FK3xc4>
- CRAWFORD, K. Y SCHULTZ, J. (2014): “Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms”. *Boston College Law Review*, Vol. 55, No. 93 <https://ssrn.com/abstract=2325784>
- DE MIGUEL BERIAIN, I. (2018). “Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the Wisconsin v. Loomis ruling”, *Law, Probability and Risk*, Volume 17, Issue 1, 1 March 2018, pp. 45–53, <https://doi.org/10.1093/lpr/mgy001>

- DE LA SIERRA MORÓN, S. (2020): “Inteligencia artificial y justicia administrativa: una aproximación desde la teoría del control de la Administración Pública”, *Revista General de Derecho Administrativo*, n.º. 53, 2020, <https://laadministracionaldia.inap.es/noticia.asp?id=1510413>
- DELGADO GOBOY y otras (2003): *Guía de aplicación práctica, para la elaboración de informes de impacto de género de las disposiciones normativas que elabore el gobierno*, de acuerdo a la ley 30/2003, 2003, Instituto De La Mujer (Ministerio de Trabajo Y Asuntos Sociales), <https://www.inmujeres.gob.es/publicacioneselectronicas/documentacion/Documentos/DE0259.pdf>
- EUROPEAN INSTITUTE OF GENDER EQUALITY (EIGE), *Artificial intelligence, platform work and gender equality*, 2021, <https://eige.europa.eu/publications/artificial-intelligence-platform-work-and-gender-equality-report>
- EUROPEAN LAW INSTITUTE (2022): *Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration*, European Law Institute-University of Viena, , https://www.europeanlaw-institute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Model_Rules_on_Impact_Assessment_of_ADMSs_Used_by_Public_Administration.pdf
- EXECUTIVE OFFICE OF THE PRESIDENT (2016): *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights*. May, 2016. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf
- FEATHERS, T. (2021): “Major Universities Are Using Race as a “High Impact Predictor” of Student Success: Students, professors, and education experts worry that that’s pushing Black students in particular out of math and science”, en *The Markup*. Marzo <https://themarkup.org/machine-learning/2021/03/02/major-universities-are-using-race-as-a-high-impact-predictor-of-student-success>
- FRA (Agencia de la Unión Europea para los Derechos Fundamentales) (2018): UE. (2018). #BigData: *Discrimination in data-supported decision making*, European Union Agency for Fundamental Rights, Viena, <https://fra.europa.eu/en/publication/2018/big-data-discrimination>
- FRA, (2019): *Data quality and artificial intelligence- mitigating bias and error to protect fundamental rights*, Luxembourg, Publications Office, junio, <https://fra.europa.eu/en/publication/2019/artificial-intelligence-data-quality>
- FRA (2020): *Artificial Intelligence and Fundamental Rights*, Agencia de los Derechos Fundamentales de la Unión Europea, 2020, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf

- FRA (2021): *Getting the future right artificial intelligence and fundamental rights*, FRA, <https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights>
- FRA- CONSEJO DE EUROPA (2018): *Manual de legislación europea contra la discriminación*, revisión 2018 340 pp., Oficina de Publicaciones de la Unión Europea, https://www.echr.coe.int/documents/handbook_non_discrim_law_spa.pdf
- FREEMAN ENGSTROM, D. y otros (2020): “Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies. Report submitted to the Administrative conference of the United States”, *NYU School of Law, Public Law Research Paper No. 20-54*, 122 págs., <http://dx.doi.org/10.2139/ssrn.3551505>
- FTC, Comisión Federal del Comercio (2016). *Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues* (6 enero) . <https://www.ftc.gov/reports/big-data-tool-inclusion-or-exclusion-understanding-issues-ftc-report>
- GAMERO CASADO, E. (2021): “Necesidad de motivación e invalidez de los actos administrativos sustentados en inteligencia artificial o en algoritmos”, *Almacén de Derecho*, 4 febrero, <https://almacenederecho.org/necesidad-de-motivacion-e-invalidez-de-los-actos-administrativos-sustentados-en-inteligencia-artificial-o-en-algoritmos>
- GAMERO CASADO, E. y FERNÁNDEZ RAMOS, S. (2016): *Manual básico de Derecho Administrativo*, (13ª ed.), Tecnos, 2016
- GIMÉNEZ GLUCK, D. (2004): *Juicio de igualdad y Tribunal Constitucional*, J.M. Bosch Editor, Barcelona.
- GIMÉNEZ GLUCK, D. (1999): *Una manifestación polémica del principio de igualdad: acciones positivas moderadas y medidas de discriminación inversa*, Tirant lo Blanch, Valencia.
- GONZÁLEZ-TREVIJANO, P. (2020): *Los principios de igualdad y no discriminación, una perspectiva de Derecho Comparado*. España, EPRS Servicio de Estudios del Parlamento Europeo, PE 659.297, Octubre, [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/659297/EPRS_STU\(2020\)659297_ES.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/659297/EPRS_STU(2020)659297_ES.pdf)
- GOODMAN, B. W. (2016): “A step towards accountable algorithms?: Algorithmic discrimination and the European Union general data protection”, *29th Conference on Neural Information Processing Systems*, <http://www.mlandthelaw.org/>
- GRUPO DEL ARTÍCULO 29 (2017): *Directrices sobre la evaluación del impacto de la protección de datos (DPIA) y determinar si el procesamiento es “probable que genere un alto riesgo” a los efectos del Reglamento 2016/679*, WP248 rev.01, Revisadas por última vez y adoptadas

- el 4 de octubre de 2017, <https://www.aepd.es/sites/default/files/2019-09/wp248rev01-es.pdf>
- GRUPO DEL ARTÍCULO 29 (2018): *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*. 3 de octubre de 2017 y revisadas el 6 de febrero de 2018.
- GUTIÉRREZ DAVID, M.E., “Acceso al código fuente y a los algoritmos de las Administraciones inteligentes. Lecciones a partir de experiencias comparadas”, en COTINO HUESO L. y CASTELLANOS CLARAMUNT J. (coords.) (2022), *ob. cit.*, pp.135-167.
- HERNÁNDEZ PEÑA, J. C. (2022): *El marco jurídico de la inteligencia artificial. Principios, procedimientos y estructuras de gobernanza*, Aranzadi, Cizur.
- HERMSTRÜWER, Y. (2020): “Artificial Intelligence and Administrative Decisions Under Uncertainty”, en WISCHMEYER, T y RADEMACHER. T. (ed.), (2020), *ob. cit.* pp. 199-223.
- HLEG (Comisión Europea–Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial), (2019): *Directrices éticas para una IA fiable*, 2019, <https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>
- HUERGO LORA, A. J. (2020): “Una aproximación a los algoritmos desde el derecho administrativo”, en HUERGO LORA, A. J. (dir.), DÍAZ GONZÁLEZ, G. M. (coord.) *La regulación de los algoritmos*, Aranzadi Thomson Reuters, Cizur, 2020.
- HUERGO LORA, A. J. (2021): “Administraciones Públicas e inteligencia artificial: ¿más o menos discrecionalidad?”, *El Cronista del Estado Social y Democrático de Derecho*, n° 96-97, 2021, pp. 78-95, siglo por <https://laadministracionaldia.inap.es/noticia.asp?id=1512494>
- ISO (2006): *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*, ISO, Tech. Rep. ISO 3534-1:2006, <https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standards/04/01/40145.html>
- JAUME PALASÍ, L. (2020): “Cómo la inteligencia artificial está impactando en las sociedades”, CERRILLO I MARTÍNEZ, A. Y PEGUERA POCH, M. (coords.), (2020): *Retos jurídicos de la inteligencia artificial*, Aranzadi, Cizur, , pp. 27-39.
- JIMÉNEZ LÓPEZ, Jesús (2022): “Oscuridad algorítmica en el sector público”, en VESTRI, G. (2022) *ob. cit.* pp. 41-66.
- JOH, E. E. (2017): “The Undue Influence of Surveillance Technology Companies on Policing” (27 febrero). *N.Y.U. L. Review Online*. <https://ssrn.com/abstract=2924620>
- JOHNSON, C. (2022 a): *Flaws plague a tool meant to help low-risk federal prisoners win early release*. NPR. 26 enero, 2022 <https://www.npr>.

- org/2022/01/26/1075509175/flaws-plague-a-tool-meant-to-help-low-risk-federal-prisoners-win-early-release ;
- JOHNSON, C. (2022 b): *Justice Department works to curb racial bias in deciding who's released from prison*. NPR. 19 de abril, <https://www.npr.org/2022/04/19/1093538706/justice-department-works-to-curb-racial-bias-in-deciding-whos-released-from-pris>
- JOINT RESEARCH CENTRE, COMISIÓN EUROPEA (2022): *AI Watch, road to the adoption of artificial intelligence by the public sector: a handbook for policymakers, public administrations and relevant stakeholders*, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2760/288757>
- JOINT RESEARCH CENTRE, ESTÉVEZ ALMENZAR, M. y otros (2022): *Glossary of human-centric artificial intelligence*, Joint Research Centre, Unión Europea, Luxemburgo, <https://publications.jrc.ec.europa.eu/repository/handle/JRC129614> DOI:10.2760/860665
- JOINT RESEARCH CENTRE, MISURACA, G. (2020), *AI Watch. Artificial Intelligence in public services. Overview of the use and impact of AI in public services in the EU*, Joint Research Centre, Unión Europea. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC120399/jrc120399_misuraca-ai-watch-public-services_30062020_def.pdf
- JOINT RESEARCH CENTRE, TANGI, L. y otros (2022): *AI Watch European landscape on the use of Artificial Intelligence by the Public Sector*, JRC Science For Policy Report, Publications Office of the European Union, Luxembourg, 2022, doi:10.2760/39336, <https://publications.jrc.ec.europa.eu/repository/handle/JRC129301>
- JORDANO FRAGA, J. (1996): “Nulidad de los actos administrativos y derechos fundamentales”, Marcial Pons, Madrid.
- JORDANO FRAGA, J. (2011): “La nulidad de los actos que lesionen derechos fundamentales: evolución jurisprudencial y desarrollo de la causa de nulidad, nucleo de las garantías en el Estado de Derechos Fundamentales”, en LÓPEZ MENUDO F. (coord.), *Derechos y garantías del ciudadano: estudios en homenaje al profesor Alfonso Pérez Moreno*, 2011, pp. 715-743.
- KAMMOURIEH, L y otros, “Group privacy in the age of big data”, en TAYLOR, L.; VAN DER SLOOT, B.; FLORIDI, L. (eds.) (2017): *Group Privacy: New Challenges of Data Technologies*, Springer, Capítulo 3, pp. 48-83, <https://www.stiftung-nv.de/sites/default/files/group-privacy-2017-authors-draft-manuscript.pdf>
- LAMMERANT, H. y otros (2019): *Big data decision-making processes and shortcuts of discrimination*, NTM/NJCM-bull. 2018/1
- LIKADI (2018): *Guía para la evaluación de impacto de género de la normativa guía para la evaluación de impacto de género de la norma-*

- tiva*, Diputación Foral de Álava, 2018, <https://web.araba.eus/documents/105044/1393927/Gu%C3%ADapara+la+evaluaci%C3%B3n+de+impacto+de+g%C3%A9nero+%C3%81lava+castellano+navegable.pdf/74905ae0-fc50-7b15-a5c3-2ab4b2f16fa3?t=1557493191803>
- LOZA CORERA, M. (2022): “Explotación y regulación del uso del big data e inteligencia artificial para los servicios públicos y la ciudad inteligente”, en COTINO HUESO, L. y TODOLÍ SIGNES, A. (coord.), *ob. cit.* pp. 125-148.
- MANTELERO, A. (2017): “From group privacy to collective privacy: towards a new dimension of privacy and data protection in the big data era”, en TAYLOR, L.; VAN DER SLOOT, B.; FLORIDI, L. (eds.) *ob. cit.* Capítulo 8.
- MANTELERO, A. (2022): *Beyond Data. Human Rights, Ethical and Social Impact Assessment*, Springer, Information Technology and Law Series IT&LAW 36, , <https://link.springer.com/book/10.1007/978-94-6265-531-7>
- MARTÍN DELGADO, I. (2009): “Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada”, *Revista de administración pública*, nº 180, pp. 353-386.
- MARTÍNEZ GARAY, L. y MONTES SUAY, F. (2018): “El uso de valoraciones del riesgo de violencia en Derecho Penal: algunas cautelas necesarias”, *Indret: Revista para el Análisis del Derecho*, nº. 2,.
- MARTÍNEZ MARTÍNEZ, R. (2019): “Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo”, *Revista catalana de dret públic*, nº 58, pp. 64-81.
- MELLADO RUIZ, L. (2017), “Supuestos de nulidad de pleno derecho de los actos administrativos”, en GAMERO CASADO, E. (dir.), *Tratado de Procedimiento Administrativo Común y Régimen Jurídico Básico del sector público* Vol. 2, Tirant lo Blanch, Valencia, pp. 1929-1983.
- MINISTERIE VAN BINNENLANDSE ZAKEN; VAN DER SLOOT, B. y otros (2021): *Non-discriminatie by design*, 69 pp., encargo para (Ministerio del Interior), Tweede Kamer (Cámara de representantes), <https://www.tweedekamer.nl/downloads/document?id=2021D22772>
- MINISTERIE VAN JUSTITIE EN VEILIGHEID (Ministerio de Justicia y Seguridad), Rijksoverheid (Gobierno central) (2021), *Richtlijnen voor het toepassen van algoritmen door overheden en publieksvoorlichting over data-analyses* (*Pautas para la aplicación de algoritmos por parte de los gobiernos y educación pública sobre análisis de datos*), Directiva (Richtlijn), de 08-03-2021, <https://acortar.link/6C226N>
- MINISTRY OF THE INTERIOR AND KINGDOM (2022): *Impact Assessment. Fundamental rights and algorithms* (IAMA), Países Bajos, marzo. <https://www.government.nl/binaries/government/documenten/>

- reports/2022/03/31/impact-assessment-fundamental-rights-and-algorithms/Fundamental+Rights+and+Algorithms+Impact+Assessment.pdf
- MIRÓ LLINARES, F. (2022): “Inteligencia artificial, delito y control penal: nuevas reflexiones y algunas predicciones sobre su impacto en el derecho y la justicia penal”, en *El Cronista del Estado Social y Democrático de Derecho*, n°. 100, 2022, pp. 174-183.
- MORAL SORIANO, L. (2022): “Inteligencia artificial y filosofía del derecho decisiones automatizadas, derecho administrativo y argumentación jurídica”, en LLANO ALONSO F. H. (dir.), *Inteligencia artificial y filosofía del derecho*, Laborun, Madrid, 2022, pp. 475-498. https://idus.us.es/bitstream/handle/11441/137250/Inteligencia%20artificial_Llano%20Alonso.pdf?sequence=1&isAllowed=y
- NATIONAL INSTITUTE OF JUSTICE (2021): *Review and Revalidation of the First Step Act Risk Assessment Tool*. National Institute of Justice NCJ 303859. <https://www.ojp.gov/pdffiles1/nij/303859.pdf>
- NAVAS NAVARRO, S. (2017): “Derecho e inteligencia artificial desde el diseño. Aproximaciones”, NAVAS NAVARRO, S. (coord.). *Inteligencia artificial: tecnología, derecho*, Tirant lo Blanch, Valencia, pp. 23-72.
- NICOLAS, T. (2014) “Big Data Proxies and Health Privacy Exceptionalism”, en *Health Matrix* n° 24, pp. 65-108.
- NIST (2022): *AI Risk Management Framework: Initial Draft*, 17 de marzo, <https://www.nist.gov/document/ai-risk-management-framework-initial-draft>
- NIST; SCHWARTZ, R. y otros (2022): *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence, Special Publication 1270* (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, <https://doi.org/10.6028/NIST.SP.1270>
- NOORDT, C., y MISURACA, G. (2022), “Artificial intelligence for the public sector: results of landscaping the use of AI in government across the European Union”, *Government Information Quarterly*, 101714, <https://doi.org/10.1016/j.giq.2022.101714>
- O’NEIL, C. (2017): *Armas de Destrucción Matemática. Como el Big Data aumenta la desigualdad y amenaza la democracia*. Capitán Swing, Madrid.
- OBERMEYER, Z. y otros (2019), *Dissecting racial bias in an algorithm used to manage the health of populations*. Science. Vol. 366, No. 6464. Oct. 25, 2019 <https://www.science.org/doi/10.1126/science.aax2342>
- PALMA ORTIGOSA, A. (2021): *Régimen jurídico de la toma de decisiones automatizadas y el uso de sistemas de inteligencia artificial en el marco del derecho a la protección de datos personales*, Tesis doctoral Universidad de Valencia, 2021

- PARLAMENTO EUROPEO (2017): *Resolución de 14 de marzo de 2017, sobre las implicaciones de los macrodatos en los derechos fundamentales: privacidad, protección de datos, no discriminación, seguridad y aplicación de la ley* (2016/2225(INI)). Acceso en <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0076+0+DOC+XML+V0//ES>
- POLO ROCA, A. (2021): “Datos, datos, datos: el dato personal, el dato no personal, el dato personal compuesto, la anonimización, la pertenencia del dato y otras cuestiones sobre datos”. *Estudios de Deusto: revista de la Universidad de Deusto*, Vol. 69, n°. 1, pp. 165-194, DOI: 10.18543/ed-69(1)-2021
- RED DAIA (2020 a): *Aportación a consulta de derechos digitales de julio, 2020* <https://acortar.link/q1wl3i>
- RED DAIA (2020 b): *Aportación DAIA a consulta pública Carta de derechos digitales*, diciembre 2020 <https://bit.ly/3paF0H0>
- REY MARTÍNEZ, F. (2017): *Igualdad y prohibición de discriminación: de 1978 a 2018*, *Revista de derecho político*, N° 100, 2017, pp. 125-171.
- REY MARTÍNEZ, F. (2020): *Derecho antidiscriminatorio*, Aranzadi Thomson Reuters, Cízur.
- ROEHL, U. (2022): “Understanding Automated Decision-Making in the Public Sector: A Classification of Automated, Administrative Decision-Making”, en Juell-Skielse, G. y otros (eds.), *Service Automation in the Public Sector. Progress in IS*. Springer, Cham, pp. 35-63, DOI: 10.1007/978-3-030-92644-1_3,
- RUBIO LLORENTE, F. (1993): “Igualdad en la jurisprudencia del Tribunal Constitucional. Introducción”, en *La Forma del Poder (Estudios sobre la Constitución)*, CEC, Madrid, pp. 637-684;
- SANGEN, M. (2019), *Research on fair algorithms for policy*, 6-8-2019, <https://www.cbs.nl/nl-nl/corporate/2019/32/research-on-fair-algorithms-for-policy>.
- SIMMONS, R. (2016): “Quantifying Criminal Procedure: How to Unlock the Potential of Big Data in Our Criminal Justice System”, *Mich. St. L. Rev.* 947 (Ohio State Public Law Working Paper No. 362. pp. 977 y ss. <http://dx.doi.org/10.2139/ssrn.2816006>
- SIMÓ, P. y MAGRO, V. (2021): *Justicia cautelar e inteligencia artificial: la alternativa a los atávicos heurísticos judiciales*, J.M. Bosch Editor, Barcelona, 2021.
- SIMÓ, P. (2022): *La prisión algorítmica*, Tirant lo Blanch, Valencia.
- SORIANO ARNANZ, A. (2021 a): “Decisiones automatizadas y discriminación: aproximación y propuestas generales”, *Revista General de Derecho Administrativo* (Iustel, enero 2021) n° 56, acceso, <https://laadministraciondialia.inap.es/noticia.asp?id=1511706>

- SORIANO ARNANZ, A. (2021 b): “Decisiones automatizadas: problemas y soluciones jurídicas. Más allá de la protección de datos”, *Revista de Derecho Público: Teoría y Método*, Vol. 3, pp. 85-127, DOI:10.37417/RPD/vol_1_2021_535
- SORIANO ARNANZ, A. (2022 a): “El uso de sistemas automatizados por los poderes públicos. El caso del sistema empleado por el servicio público de empleo austriaco”, en COTINO HUESO, L. y TODOLÍ SIGNES, A. (coord.), *Explotación y regulación del uso del big data e inteligencia artificial para los servicios públicos y la ciudad inteligente*, Tirant lo Blanch, Valencia, pp. 192-216.
- SORIANO ARNANZ, A. (2022 b): “Discriminación algorítmica: garantías y protección jurídica”, en COTINO HUESO, L. (coord.), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Aranzadi, Cizur Menor, 2022, , págs. 139-169.
- TISCHBIREK, A. (2020): “Artificial Intelligence and Discrimination: Discriminating Against Discriminatory Systems”, en WISCHMEYER, T y RA-DEMACHER, T. (ed.), (2020), *ob. cit.* pp. 103-121.
- U.S. DEPARTMENT OF JUSTICE (2022): *Algorithms, Artificial Intelligence, and Disability Discrimination in Hiring*. Mayo, <https://beta.ada.gov/resources/ai-guidance/>
- U.S. EQUAL EMPLOYMENT OPPORTUNITY COMMISSION (2022): *The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees*. EEOC-NVTA-2022-2, <https://www.eeoc.gov/laws/guidance/americans-disabilities-act-and-use-software-algorithms-and-artificial-intelligence>
- U.S. TRANSPORTATION SECURITY ADMINISTRATION (2022): *Transgender/ Non Binary / Gender Nonconforming Passengers*. TSA, <https://www.tsa.gov/transgender-passengers>
- UNESCO (2020): *Artificial Intelligence and Gender Equality. Key findings of UNESCO’s Global Dialogue*, 2020, <https://unesdoc.unesco.org/ark:/48223/pf0000374174>
- VALERO TORRIJOS, J. (2007): *El régimen jurídico de la e-Administración. El uso de los medios informáticos y telemáticos en el procedimiento administrativo común*, 2.ª ed., Comares, Granada.
- VALERO TORRIJOS, J. (2013): *Derecho, innovación y administración electrónica*, Global Law = Derecho Global, Sevilla.
- VALERO TORRIJOS, J. (2015): “Ciudades inteligentes y datos abiertos: implicaciones jurídicas para la protección de los datos de carácter personal”, *Istituzioni del federalismo: rivista di studi giuridici e politici*, n.º. 4, 2015 (“Smart Cities” e amministrazioni intelligenti), pp. 1025-1047.

- VALLE ESCOLANO, R., “Transparencia en la inteligencia artificial y en el uso de algoritmos: una visión de género”, COTINO HUESO L. y CASTELLANOS CLARAMUNT J. (coords.) (2022), *ob. cit.*, pp. 85-110.
- VAN DER SLOOT, B. (2021): “Regulating non-personal data in the age of Big Data”, en TZANOY, M. (Ed.) (2021): *Health data privacy under the GDPR : Big Data challenges and regulatory responses*, Routledge pp. 85-105; p. 16 de la versión SSRN.
- VELASCO RICO, C. I. (2022): “Tecnologías disruptivas en la Administración Pública. Inteligencia artificial y Blockchain”, en CERRILLO I MARTÍNEZ, A. (coord.) *La Administración Digital*, Dykinson, Madrid, pp. 228-256.
- VESTRI, G. (2021): “Sistemas de inteligencia artificial en la contratación pública: entre códigos fuente y datos abiertos”, *Actualidad administrativa*, nº 12.
- VESTRI, G. (2022 a): “Denegación vs. Derecho de acceso al código fuente en los sistemas algorítmicos. Una perspectiva jurídico-administrativa”, COTINO HUESO L. y CASTELLANOS CLARAMUNT J. (coords.) (2022), *ob. cit.*, pp. 111-134.
- VESTRI, G. (2022 b): *Disrupción tecnológica en la administración pública retos*, Aranzadi, Cizur.
- WACHTER, S. (2022): “The Theory of Artificial Immutability: Protecting Algorithmic Groups under Anti-Discrimination Law”, 15 de febrero, *Tulane Law Review*, <http://dx.doi.org/10.2139/ssrn.4099100>
- WACHTER, S. y MITTELSTADT, B. D. (2019): “A right to reasonable inferences: re-thinking data protection law in the age of big data and AI”, *Columbia Business Law Review*, vol. 2019, No. 2, pp. 494-620, versión <https://ssrn.com/abstract=3248829>
- WACHTER, S., y otros, “Bias Preservation in Machine Learning: The Legality of Fairness Metrics Under EU Non-Discrimination Law”, *West Virginia Law Review*, Vol. 123, núm. 735, 2021, <https://x.gd/hKKKqS>
- WHITE HOUSE (2022): *The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People*, White House Office of Science and Technology Policy, Washington DC., octubre, <https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf>
- WISCHMEYER, T y RADEMACHER. T. (ed.), (2020): *Regulating Artificial Intelligence*, Springer, 2020, <https://doi.org/10.1007/978-3-030-32361-5>
- XENIDIS, R. y SENDEN, L. (2020): “EU Non-discrimination law in the era of artificial intelligence: mapping the challenges of algorithmic discrimination”, en BERNITZ, U. y otros, (eds.), *General Principles of EU Law and the EU Digital Order*, Alphen aan den Rijn, Kluwer Law International, 2020.

ZLOTNIK, A. (2019): “Inteligencia Artificial en las Administraciones Públicas: definiciones, evaluación de viabilidad de proyectos y áreas de aplicación”, *Boletic*, núm. 84, pp. 27-28, <http://apiscam.blogspot.com/2019/11/boletic-revista-de-astic-monografico-de.html>

Capítulo VIII

Las singularidades del tratamiento de datos de carácter personal en entornos de inteligencia artificial en el sector público

JULIÁN VALERO TORRIJOS

*Catedrático de Derecho Administrativo
Director de la Cátedra Fundación Integra
sobre Identidad y Derechos Digitales
Universidad de Murcia*

Resumen: El tratamiento de datos de carácter personal por parte de las Administraciones y entidades del sector público en el diseño e implementación de las herramientas de IA requiere tener en cuenta las garantías establecidas por la normativa europea y la legislación española que regulan este derecho fundamental. En este trabajo se analizan las condiciones de licitud que se han de respetar para asegurar que el uso de la IA es conforme con las exigencias establecidas por dicha considerando las singularidades existentes en el ámbito del sector público.

Abstract: The processing of personal data by Public Administrations when implementing AI tools involves taking into account the safeguards imposed by the European and Spanish legislation that rule this fundamental right. This work analyses the lawful conditions that must be complied with to guarantee that the use of AI fulfils the requirements established by those regulations, considering the singularities of the public sector.

Palabras clave: protección de datos personales, IA, algoritmos, actuación administrativa automatizada

Keywords: personal data protection, Artificial Intelligence, algorithms, automated administrative decisions

SUMARIO. - 1. EL USO DE DATOS PERSONALES POR LAS ADMINISTRACIONES PÚBLICAS Y LA INTELIGENCIA ARTIFICIAL.- 2. EL CONTEXTO TECNOLÓGICO Y REGULATORIO.- 2.1. La singularidad de la tecnología y su proyección sobre el objeto de estudio: el planteamiento del RGPD.- 2.2. La regulación sobre las garantías y medidas de seguridad que deben adoptarse con relación al uso de medios electrónicos en la actividad administrativa.- 3. LAS BASES JURÍDICAS PARA EL TRATAMIENTO DE LOS DATOS DE CARÁCTER PERSONAL.- 3.1. Las limitaciones del consentimiento como base jurídica para el tratamiento de datos personales en el sector público.- 3.2. Las singularidades de los datos de categoría especial: el caso de los datos biométricos.- 3.3. El sentido y alcance de la habilitación legal para el tratamiento de los datos personales sin consentimiento: una moneda de dos caras.- 4. LAS CONDICIONES PARA LA LICITUD DEL TRATAMIENTO.- 4.1. El alcance de los principios generales y su proyección en los tratamientos basados en IA por parte de las Administraciones Públicas.- 4.1.1.

La singular eficacia del principio de transparencia en la normativa sobre protección de datos personales.- 4.1.2. El alcance de las limitaciones impuestas por el principio de minimización.- 4.1.3. La prohibición de usos incompatibles y la limitación del tiempo de conservación de los datos.- 4.2. El efectivo cumplimiento de las normas de seguridad: la relevancia del análisis de riesgos y la evaluación de impacto.- 4.3. La externalización en la gestión de las herramientas basadas en el uso de IA: las condiciones jurídicas en el encargo del tratamiento de datos personales.- 5. LA INCIDENCIA DEL INCUMPLIMIENTO DE LA REGULACIÓN EN MATERIA DE PROTECCIÓN DE DATOS PERSONALES SOBRE LA VALIDEZ DE LA ACTUACIÓN ADMINISTRATIVA.- 6. REFLEXIÓN FINAL

1.- EL USO DE DATOS PERSONALES POR LAS ADMINISTRACIONES PÚBLICAS Y LA INTELIGENCIA ARTIFICIAL

Una de las principales condiciones para la viabilidad de las herramientas de IA que utilicen las Administraciones Públicas en el ejercicio de sus competencias se refiere a la necesidad de procesar un conjunto de datos adecuado, tanto desde la perspectiva cuantitativa como cualitativa¹. Se trata, por tanto, de una exigencia estrechamente vinculada a las condiciones en las que ha de tener lugar su diseño y su implantación práctica, lo que lleva a plantear la necesaria integración de la variable jurídica como presupuesto inexcusable para la legitimidad de la innovación tecnológica, en particular por lo que se refiere al respeto de los derechos fundamentales².

Aunque es necesario relativizar las implicaciones del derecho a la protección de los datos de carácter personal en el contexto del uso de algoritmos en la medida que, con frecuencia, su entrenamiento

¹ A este respecto, cfr. ZLOTNIK, A. (2019): “IA en las Administraciones Públicas: definiciones, evaluación de viabilidad de proyectos y áreas de aplicación”, *Boletín*, núm. 84, p. 27, así como HUERGO LORA, A. (2020): “Una aproximación a los algoritmos desde el Derecho Administrativo”, en HUERGO LORA, A. (dir.), *La regulación de los algoritmos*, Thomson-Aranzadi, Cizur Menor, p. 42, quien enfatiza que la disponibilidad de los datos estará relacionada con la permisividad de la legislación para permitir su uso.

² MARTÍNEZ MARTÍNEZ, R. (2019): “Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo”, *Revista catalana de Derecho Público*, núm. 58, pp. 75-78.

tendrá lugar con datos anonimizados³, dicha toma de postura inicial exige cuando menos dos importantes precisiones. Por una parte, la exclusión del régimen jurídico que regula este derecho fundamental requeriría que los datos fueran anónimos, condición que más allá de sus implicaciones estrictamente tecnológicas —especialmente desde la perspectiva de la reversibilidad de la anonimización—, tiene unos claros contornos jurídicos. En concreto, si tenemos en cuenta que, por lo que se refiere al cumplimiento normativo, la interpretación que llevan a cabo las autoridades de control⁴ adquiere una singular relevancia a la hora de aclarar cuándo nos encontramos realmente ante un supuesto de anonimización, al menos en tanto dicho criterio no haya sido objeto de revisión judicial.

Por otro lado, resulta imprescindible distinguir claramente entre anonimización y seudonimización ya que con frecuencia se confunden ambos términos y las consecuencias jurídicas son ciertamente relevantes⁵. Para ello hay que partir de la definición de datos de carácter personal, de modo que es precisamente la identificabilidad del titular la que justifica que en los casos de datos seudonimizados resulte aplicable la normativa sobre protección de datos personales tal y como se deriva del artículo 4.5 RGPD, lo que no resulta un obstáculo para su reutilización, según se ha demostrado con la disposición adicional decimoséptima LOPD, donde se admite el uso de datos personales seudonimizados sin necesidad de consentimiento para fines de investigación en salud y, en particular, biomédica siempre que se aseguren ciertas condiciones y límites fijados normativamente⁶.

La determinación del alcance de la normativa sobre protección de datos de carácter personal requiere plantear los escenarios en los que tendrá lugar un tratamiento de estos por parte de las Administraciones Públicas para el ejercicio de sus competencias. Así,

³ HUERGO LORA, A. (2020), *op.cit.*, pp. 53 y 54.

⁴ AEPD (2016), *Orientaciones y garantías en los procedimientos de anonimización de datos personales*.

⁵ A este respecto, véase AEPD (2021a), *10 malentendidos relacionados con la anonimización*.

⁶ COTINO HUESO, L. (2017): “Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales”, *Dilemata. Revista Internacional de Éticas Aplicadas*, núm. 24, p. 146.

en primer lugar, es posible que los algoritmos realicen tratamientos de datos que permitan identificar a sus titulares, bien de manera directa bien a través de información adicional en los términos antes referidos para la seudonimización. En estos casos la licitud del tratamiento obligaría, entre otras medidas y como analizaremos más adelante, a asegurar que el uso que se le pretende dar a los datos no sea incompatible con el que inicialmente justificara su recogida y tratamiento, así como informar al titular de que sus datos se están utilizando en otro contexto y para otro propósito. En consecuencia, resulta imprescindible que la Administración que realiza el nuevo tratamiento compruebe que efectivamente se respetan dichas garantías pues, de lo contrario, el uso de la información podría ser considerarse ilícito.

En segundo lugar, la utilización de IA puede suponer que se vaya a adoptar —o al menos que sirva de base a la misma— una decisión administrativa que afecta a una persona física identificada o identificable, incluso a través de la elaboración de perfiles. En este supuesto, aunque los datos que se hubieran utilizado para entrenar al algoritmo o adoptar la decisión estuvieran anonimizados, lo cierto es que la existencia de un tratamiento de información vinculado al ejercicio de una competencia administrativa se encontraría sometida al régimen de la protección de datos personales en la medida que se refiere a concretas personas físicas. En consecuencia, el propio procesamiento de los datos y la programación de la herramienta deberían respetar plenamente dicha regulación, de manera que constituirá en sí mismo un tratamiento sometido al RGPD, así como al resto de las disposiciones aplicables sobre procedimiento administrativo y régimen jurídico del sector público. Es más, no podría argumentarse que si el resultado del tratamiento fuera simplemente un instrumento de apoyo a la decisión administrativa el procesamiento previo no se encontraría sometido al RGPD, ya que esta regulación europea, a diferencia de la normativa estatal reguladora de la actuación administrativa automatizada, en modo alguno tiene por objeto establecer las condiciones para su licitud desde la perspectiva de la regularidad formal, procedimental y sustantiva, sino, por el contrario, asegurar que el tratamiento de la información necesaria para su adopción y/o preparación respeta una serie de garantías mínimas relativas al derecho fundamental consagrado en el artículo 18.4 CE. Aunque, como trataremos de justificar más

adelante y no podría ser de otra manera, la validez de la actuación administrativa también se encuentre condicionada al efectivo respeto de los derechos fundamentales y las libertades públicas.

En definitiva, las aplicaciones de IA utilizadas por las Administraciones Públicas para el ejercicio de sus competencias pueden procesar datos personales bien para entrenar sus sistemas de aprendizaje automático, esto es, para construir sus modelos algorítmicos, bien para hacer inferencias sobre concretas personas físicas⁷, tanto en las fases de entrenamiento, validación, despliegue efectivo y explotación⁸, lo que determina la aplicación del RGPD incluso si se tratase de datos seudonimizados.

2.- EL CONTEXTO TECNOLÓGICO Y REGULATORIO

2.1.- *La singularidad de la tecnología y su proyección sobre el objeto de estudio: el planteamiento del RGPD*

Resulta indiscutible que la generalización el uso de las tecnologías de la información y la comunicación ha tenido un enorme impacto en el Derecho, no sólo condicionando de manera decisiva las actividades que se realizan en los más diversos ámbitos sociales sino, sobre todo y por lo que ahora nos interesa, determinando drásticamente el enfoque y el alcance de la regulación normativa y la interpretación que llevan a cabo los operadores jurídicos. Además, supone además la incorporación de un elemento adicional de complejidad debido, entre otras razones, a la intermediación tecnológica de las relaciones jurídicas, en particular por lo que se refiere al estatuto de los sujetos. Aunque no procede en este momento analizar las diversas modalidades de IA y sus principales características⁹, sin embargo es necesario enfatizar que

⁷ SARTOR, G. (2020): *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*, European Parliamentary Research Service, Bruselas, p. I.

⁸ AEPD (2021b), *Gestión del riesgo y evaluación de impacto en tratamiento de datos personales*, p. 13.

⁹ Véase el trabajo de N. DÍAZ DÍAZ “¿Qué es la inteligencia artificial?” en el capítulo I de esta misma obra.

la propia configuración tecnológica de estas herramientas supone un reto cualitativo nada desdeñable respecto del objeto de este capítulo, en particular en el caso de los sistemas de aprendizaje automático¹⁰. A este respecto, las actuaciones automatizadas se caracterizan por la ausencia de la intervención directa de un empleado público en su adopción inmediata, si bien el control y la supervisión de la decisión pueden imputarse en última instancia a quien hubiere realizado la configuración técnica de la herramienta, que puede haber sido diseñada por el propio personal de la Administración Pública o, en su caso, de una empresa contratada al efecto.

Desde este planteamiento, la disciplina del artículo 41 LRJSP tiene un doble objetivo: por una parte, que la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente se lleve a cabo por un órgano administrativo; y, de otra, que exista seguridad jurídica en cuanto al órgano jurídicamente responsable de la decisión adoptada, en particular por lo que se refiere a su impugnación. Sin embargo, esta regulación resulta manifiestamente insuficiente en el caso de los sistemas de IA de aprendizaje automático ya que en estos casos no se trata únicamente de garantizar de una supervisión formal de la programación —lo que resulta manifiestamente exiguo— sino, en definitiva, de asegurar que la objetividad y legalidad de la decisión administrativa puedan ser justificadas a través de su efectiva explicabilidad más allá de la mera motivación en los términos del artículo 35 LPAC. Se trata de una implicación de enorme importancia desde la perspectiva del tratamiento de los datos personales, ya que la validez de la decisión administrativa no puede descansar ya en la

¹⁰ Incluso, como se ha destacado por las autoridades de control (AEPD-EDPS, 2022: *Malentendidos sobre el machine learning*, p. 7), estos sistemas “son capaces de identificar patrones en los datos personales que van más allá de los planteados explícitamente en el desarrollo del modelo, y que podrían ser desconocidos incluso para los propios afectados”, de manera que pueden verse condicionados por información que ni siquiera podrían anticipar y, de otra parte, cabría pensar que reciban información a ellos vinculada que podría tener un mayor impacto debido al contexto específico donde tiene lugar el tratamiento. Cfr. HUERGO LORA, A. (2020), *op.cit.*, p. 36, donde se enfatiza que las técnicas de *machine learning* puede llevar a que se planteen opciones que vayan más allá del modo de pensar o el marco conceptual comúnmente aceptado.

mera presunción consagrada legalmente sino, en última instancia, en el efectivo cumplimiento de las garantías formales que se analizarán en este trabajo, que además habrán de documentarse por exigencia del principio de responsabilidad proactiva, en particular por lo que se refiere a la definición inicial de los requisitos funcionales¹¹.

A este respecto, como se ha resaltado reiteradamente¹², el uso de la IA por parte de las Administraciones Públicas implica una reduplicada importancia de la transparencia en tanto que instrumento que, en última instancia, garantiza la adecuada defensa de los derechos e intereses legítimos. Este planteamiento resulta de especial trascendencia por lo que se refiere al tratamiento de datos personales para el desarrollo de actividades de interés público, donde la exigencia de consentimiento se ve desplazada a cambio de unas mayores garantías establecidas normativamente, entre las que destaca el refuerzo de la transparencia, en particular cuando se trate de actuaciones automatizadas.

El uso de medios electrónicos por parte de las Administraciones Públicas supone con frecuencia la intervención de otros sujetos que prestan servicios y que, por tanto, intermedian las relaciones entre aquellas y los ciudadanos, tratándose normalmente de empresas especializadas. En el caso de la IA se trata de una característica presente todavía en mayor medida¹³, lo que supone un riesgo adicional tanto por la pérdida de capacidad de control que implica para las entidades

¹¹ SANTOS MORÓN, M.J.: “Tratamiento de datos, sujetos implicados, responsabilidad proactiva”, en AA.VV.: *Protección de datos personales*, Tirant lo Blanch, Valencia, 2020, p. 49.

¹² Véase el trabajo de I. MARTÍN DELGADO “La aplicación del principio de transparencia a la actividad administrativa algorítmica” en el capítulo IV de esta misma obra; CERRILLO i MARTÍNEZ, A. (2019): “Com obrir les caixes negres de les administracions públiques? Transparència i rendició de comptes en l’ús dels algorismes”, *Revista Catalana de Dret Públic*, núm. 58. p. 18-20; así como VESTRI, G. (2021): “La IA ante el desafío de la transparencia algorítmica. Una aproximación desde la perspectiva jurídico-administrativa”, *Revista Aragonesa de Administración Pública*, núm. 56, pp. 384-387, donde se distingue su relevancia en función del momento previo a la adquisición de la tecnología o de su efectivo funcionamiento.

¹³ RAMÍO, C. (2018): *Inteligencia artificial y administración pública: Robots y humanos compartiendo el servicio público*, Catarata, Madrid, pp. 95-95 y, desde la perspectiva de su incidencia en el modelo democrático, pp. 127-128.

públicas como, por lo que ahora nos interesa, debido a las consecuencias que esta mediación podría tener a la hora de que el ciudadano ejerza sus derechos, sobre todo dado el potencial conflicto con la posición jurídica de la empresa que hubiese desarrollado la aplicación. Más aún, con frecuencia estas entidades prestan sus servicios desde la distancia geográfica, matiz importante desde la perspectiva de la aplicación de otros ordenamientos jurídicos y la competencia de autoridades pertenecientes a otros Estados. Se trata, pues, de una consecuencia relevante por lo que se refiere a la protección de datos personales, ya que cuando las herramientas de IA utilicen servicios basados en la denominada “computación en la nube”, podrían estar teniendo lugar transferencias internacionales y este tipo de tratamientos tienen una regulación específica que contempla las garantías que han de adoptarse¹⁴.

La digitalización de la actividad administrativa presenta también importantes respecto a las mayores posibilidades de conservación de la información. En concreto, la IA con frecuencia se basa en el análisis masivo de datos (*big data*)¹⁵, lo que supone la necesidad de tratar información que proviene de fuentes muy diversas y, sobre todo, que puede haber sido obtenida para concretas finalidades que, más tarde, se ven superadas al reutilizarse la información para otras muy distintas. Así pues, como analizaremos más adelante, la exigencia normativa de que los datos personales sólo se conserven durante el tiempo estrictamente necesario puede afectar a la validez de las decisiones administrativas que se adopten utilizando estas herramientas.

En definitiva, la caracterización del entorno digital en que se desenvuelve la IA plantea una singular incidencia sobre los derechos y libertades de los ciudadanos en sus relaciones con las Administra-

¹⁴ Como es sabido, se trata de una cuestión que ha generado una importante polémica desde el punto de vista a las garantías jurídicas aplicables, en particular por lo que se refiere a las redes sociales y las transferencias de datos a los Estados Unidos. A este respecto, véase la importante STJUE de 16 de julio de 2020, asunto C-311/18 Maximilian Schrems c. Data Protection Commissioner y Facebook Ireland Ltd.

¹⁵ Aunque no siempre está presente esta condición, hasta el punto de que las propias autoridades de control en materia de protección de datos han advertido específicamente que se trata de un malentendido frecuente. A este respecto, AEPD-EDPS (2022), *ob.cit.*, p. 4.

ciones Públicas, que se pueden ver afectados de manera reduplicada, sobre todo ante el riesgo de que las potestades unilaterales se vean reforzadas de manera determinante al proyectarse en el ámbito de la normativa sobre protección de datos personales, fundamentalmente por tres razones¹⁶:

- a) la excesiva generalidad de las normas aplicables, opción regulatoria que en última instancia resulta imprescindible para hacer frente a la evolución tecnológica y, en particular, de herramientas disruptivas como puede ser el caso de la IA;
- b) como consecuencia de lo anterior, la abundancia de conceptos jurídicos indeterminados y el recurso a los principales generales —singularmente los consagrados en el artículo 5 RGPD—, lo que supone una especial dificultad al proceder a su efectiva aplicación práctica¹⁷;
- c) la incorporación de conceptos tecnológicos en las propias normas jurídicas, circunstancia que obliga a incorporar un criterio singularmente especializado a la hora de valorar el alcance de las obligaciones a cumplir por las Administraciones Públicas como responsables del tratamiento de datos de carácter personal.

Teniendo en cuenta que, en aplicación de lo dispuesto en el artículo 39 LPAC, las decisiones adoptadas por las Administraciones Públicas se presumirán válidas, las notas características de la regulación que se acaban de exponer pueden suponer un refuerzo adicional a la hora de evaluar su conformidad a Derecho. En concreto, la eventual disconformidad jurídica del ciudadano como titular de los datos sometidos a tratamiento normalmente no podrá basarse, sin más, en la mera discrepancia interpretativa o en la simple contraposición de argumentos estrictamente jurídicos sino que, por el contrario, deberá fundamentarse en el incumplimiento de criterios técnicos o que,

¹⁶ Para una visión más amplia de este planteamiento, VALERO TORRIJOS, J. (2022): “Los derechos en la era digital”, en RODRÍGUEZ AYUSO, J.F. (Coord.), *Nuevos retos en materia de derechos digitales en un contexto perspectiva multidisciplinar*, Thomson Reuters-Aranzadi, Cizur Menor, pp. 32-37.

¹⁷ El RGPD abunda en cláusulas vagas y normas abiertas, cuya aplicación a menudo requiere un equilibrio de intereses en conflicto, tal y como ha destacado SARTOR, G. (2020), *op.cit.*, p. III.

al menos, exijan un elevado nivel de especialización; cualificación de la que no siempre dispondrá el ciudadano a la hora de defender su posición jurídica en tanto que titular de los datos. En consecuencia, las resoluciones y documentos elaborados por las autoridades de control —así como, en su caso, el criterio adoptado en el ejercicio de sus competencias, a la hora de tramitar procedimientos formalizados— adquieren una mayor relevancia ante una eventual discrepancia por parte de los ciudadanos con la actuación de la Administración responsable del tratamiento, por lo que dichas autoridades están llamadas a desempeñar una relevante y reduplicada función garantista de la posición jurídica de aquellos en el referido escenario de desequilibrio.

2.2.- La regulación sobre las garantías y medidas de seguridad que deben adoptarse con relación al uso de medios electrónicos en la actividad administrativa

Como se analizará más adelante, el tratamiento de datos de carácter personal por las Administraciones Públicas basado en el uso de instrumentos de IA no se puede sustentar, con carácter general, en el consentimiento de la persona titular de la información. En efecto, la desigualdad en la posición jurídica constituye por sí misma una razón para excluir el consentimiento más allá de los supuestos de servicios de valor añadido libremente aceptados, de manera que será necesario acudir principalmente al ejercicio de funciones públicas en los términos del artículo 6.1.e) RGPD. Esta ubicación sistemática de la base legitimadora del tratamiento de los datos supone, tal y como prevé el artículo 6.3 RGPD, que deba ser fijada normativamente, bien por el Derecho de la Unión Europea bien por el Derecho del correspondiente Estado miembro. En concreto,

Dicha base jurídica podrá contener disposiciones específicas para adaptar la aplicación de normas del presente Reglamento, entre otras: las condiciones generales que rigen la licitud del tratamiento por parte del responsable; los tipos de datos objeto de tratamiento; los interesados afectados; las entidades a las que se pueden comunicar datos personales y los fines de tal comunicación; la limitación de la finalidad; los plazos de conservación de los datos, así como las operaciones y los procedimientos del tratamiento, incluidas las medidas para garantizar un tratamiento lícito y equitativo

Más allá de las medidas específicas que pueda establecer la regulación sectorial aplicable a cada ámbito de actividad material, por lo que se refiere a los medios electrónicos utilizados por las Administraciones Públicas en el ejercicio de sus funciones, tras la reforma legislativa de 2015 se ha impuesto con carácter general la exigencia de su utilización tanto para la gestión documental como, asimismo, las relaciones con los ciudadanos. En consecuencia, las condiciones normativas establecidas por el legislador estatal constituyen, asimismo, premisas inexcusables para el uso de la IA por parte de las Administraciones Públicas y, en última instancia, para la efectividad de los derechos y libertades fundamentales¹⁸.

En concreto, partiendo de lo establecido en el artículo 32 RGPD acerca de la necesidad de adoptar las medidas de seguridad adecuadas en el tratamiento de los datos de carácter personal, el cumplimiento del Esquema Nacional de Seguridad se convierte en una exigencia inexcusable (artículos 17, 27 y 70 LPAC, artículos 46 y 156 LRJSP); obligación que ha sido reforzada de manera contundente por la legislación estatal sobre protección de datos de carácter personal al exigir a las entidades públicas que apliquen “a los tratamientos de datos personales las medidas de seguridad que correspondan de las previstas en el Esquema Nacional de Seguridad” (disposición adicional primera LOPD). Incluso, lo que presenta especial importancia en el supuesto de la IA, cuando un tercero preste un servicio a la entidad pública que implique el tratamiento de información de los ciudadanos se establece expresamente que “las medidas de seguridad se corresponderán con las de la Administración pública de origen y se ajustarán al Esquema Nacional de Seguridad”. Con relación a esta exigencia normativa se puede constatar en la práctica una preocupante falta de cumplimiento de las medidas de seguridad aludidas, lo que sin duda supone un grave problema en términos jurídicos ya que en el contexto digital el efectivo cumplimiento de las normas técnicas constituye precisamente la premisa elemental para el respeto de las garantías jurídicas.

¹⁸ CANALS AMETLLER, D.: “La seguridad en el entorno digital”, en CANALS, D. (dir.), *Ciberseguridad. Un nuevo reto para el Estado y los Gobiernos Locales*, Wolters Kluwer-El Consultor, Madrid, 2021, p. 76, quien destaca que el derecho más afectado es precisamente la protección de los datos de carácter personal.

Por otro lado, en la medida que el uso de la IA suponga la realización de actividades automatizadas, el cumplimiento de lo dispuesto en el artículo 41 LRSJP se convierte también en una condición inexcusable de licitud para el tratamiento de datos de carácter personal ya que afecta de manera sustancial a las garantías en que ha de tener lugar. Según este precepto resulta imprescindible “la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente”. Así pues, la ausencia de esta determinación por el órgano competente en la respectiva Administración Pública puede considerarse un incumplimiento de las “condiciones generales que rigen la licitud del tratamiento por parte del responsable” según han sido establecidas por el legislador estatal, al amparo de lo dispuesto en el artículo 6.1.e) RGPD, para los supuestos en que el tratamiento de los datos personales tuviera lugar bajo la modalidad de una actuación administrativa automatizada.

3.- LAS BASES JURÍDICAS PARA EL TRATAMIENTO DE LOS DATOS DE CARÁCTER PERSONAL

3.1.- *Las limitaciones del consentimiento como base jurídica para el tratamiento de datos personales en el sector público*

Según dispone el artículo 4.11 RGPD se considera consentimiento del interesado “toda manifestación de voluntad libre, específica, informada e inequívoca por la que el interesado acepta, ya sea mediante una declaración o una clara acción afirmativa, el tratamiento de datos personales que le conciernen”. Así pues, la libertad sería la primera de las condiciones necesarias para la validez del consentimiento y, desde la perspectiva que ahora interesa, el Considerando 43 RGPD advierte que el consentimiento no puede ser una base jurídica adecuada cuando “exista un desequilibrio claro entre el interesado y el responsable del tratamiento, en particular cuando dicho responsable sea una autoridad pública y sea por lo tanto improbable que el consentimiento se haya dado libremente en todas las circunstancias de dicha situación particular”. Partiendo de esta afirmación general, el

Grupo de Trabajo del Artículo 29 ha reconocido que, aun cuando no existirán alternativas realistas en la mayoría de los casos, no por ello se ha de excluir necesariamente el consentimiento para legitimar el tratamiento de datos personales por parte de las autoridades públicas en ciertas circunstancias¹⁹. En concreto, la AEPD ha mantenido de manera rotunda que el consentimiento no puede ser la base jurídica en las relaciones con la Administración “en aquellos supuestos en que existe una relación en la que no puede razonablemente predicarse que exista una situación de equilibrio”²⁰.

Por su parte, el Considerando 42 RGPD afirma que no puede entenderse que el consentimiento se ha prestado libremente “cuando el interesado no goza de verdadera o libre elección o no puede denegar o retirar su consentimiento sin sufrir perjuicio alguno”. En consecuencia, según ha quedado claramente establecido por el RGPD, no sólo bastaría con respetar tales condiciones al prestar el consentimiento, sino que, en cualquier momento, se debe permitir su retirada “sin sufrir perjuicio alguno” en la relación jurídica que se entable con la Administración.

En definitiva, el consentimiento sólo puede jugar un papel residual en el tratamiento de datos personales por parte de las Administraciones Públicas, siempre que se den las siguientes condiciones:

- a) cuando se trate de ofrecer servicios de valor añadido que vayan más allá del mero ejercicio competencial por parte de la respectiva entidad pública, como sería el caso, por ejemplo, de la prestación de servicios personalizados y proactivos que impliquen un perfilado previo de la persona que los solicita;
- b) que se trate de servicios y actuaciones administrativa que, en el caso de revocarse el consentimiento, pudieran prestarse o realizarse de manera satisfactoria incluso por medios electrónicos, aunque ello suponga que no se pueda obtener el beneficio que permitiría la modalidad que aporta el referido valor añadido.

¹⁹ ARTICLE 29 WORKING PARTY (2017), *Guidelines on consent under Regulation 2016/679*, p. 6.

²⁰ AEPD, *Informe 175/2018*, p. 4.

3.2.- Las singularidades de los datos de categoría especial: el caso de los datos biométricos

Una especial importancia merece desde la perspectiva de la IA el tratamiento de las categorías especiales de datos regulados por el artículo 9 RGPD, esto es, los que revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, y el tratamiento de datos genéticos, datos biométricos dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o datos relativos a la vida sexual o la orientación sexual de una persona física.

Sin duda el uso de datos biométricos requiere un análisis específico, en particular por lo que se refiere a su eventual uso como medio de identificación utilizando herramientas de IA. El artículo 4.14 RGPD los define como “datos personales obtenidos a partir de un tratamiento técnico específico, relativos a las características físicas, fisiológicas o conductuales de una persona física que permitan o confirmen la identificación única de dicha persona, como imágenes faciales o datos dactiloscópicos”.

Ahora bien, según precisa el Considerando 51 RGPD, mientras que “el tratamiento de fotografías no debe considerarse sistemáticamente tratamiento de categorías especiales de datos personales”, por el contrario, “se encuentran comprendidas en la definición de datos biométricos cuando el hecho de ser tratadas con medios técnicos específicos permita la identificación o la autenticación unívocas de una persona física”. En consecuencia, para que un tratamiento sea estimado de categoría especial por estar basado en el uso de datos biométricos son necesarias dos exigencias:

- que se utilicen medios técnicos específicos, no bastando por tanto meras actuaciones de comprobación por parte de otra persona física;
- que tales medios permitan la identificación o autenticación unívocas de una concreta persona física.

La concurrencia de estos dos requisitos determina que resulte de aplicación el artículo 9 RGPD, relativo a las condiciones de licitud del tratamiento de categorías especiales de datos personales, lo que

supone el desplazamiento del régimen general del artículo 6 RGPD. En concreto, el artículo 9.1 RGPD parte de la prohibición general del “uso de datos biométricos dirigidos a identificar de manera unívoca a una persona física”, salvo que concurren algunas de las excepciones contempladas en el apartado siguiente. Son dos las excepciones contempladas en el artículo 9.2 RGPD que podrían resultar de aplicación:

- a) el interesado dio su consentimiento explícito para el tratamiento de dichos datos personales con uno o más de los fines especificados, excepto cuando el Derecho de la Unión o de los Estados miembros establezca que la prohibición mencionada en el apartado 1 no puede ser levantada por el interesado;
- f) el tratamiento es necesario por razones de un interés público esencial, sobre la base del Derecho de la Unión o de los Estados miembros, que debe ser proporcional al objetivo perseguido, respetar en lo esencial el derecho a la protección de datos y establecer medidas adecuadas y específicas para proteger los intereses y derechos fundamentales del interesado.

Así pues, por lo que se refiere al interés público esencial resulta imprescindible la existencia de una regulación específica, que bien puede ser estatal o europea, pero que, en todo caso, ha de respetar una serie de garantías adicionales. En primer lugar, sus previsiones en cuanto al tratamiento de este tipo de datos han de ser proporcionales respecto del objetivo perseguido, lo que exige realizar una ponderación entre la necesidad de proceder al tratamiento de la concreta tipología de datos y el alcance del propio tratamiento que, además, debe llevarse a cabo a partir de una interpretación restrictiva de lo que deba entenderse por razones de interés público, puesto que dicha categorización parte de la base de que el interés no puede ser cualquiera al resultar preceptivo que sea “esencial”. En segundo lugar, la regulación ha de respetar los aspectos esenciales del derecho fundamental, lo que necesariamente supone tener en cuenta las demás previsiones que establece el RGPD, en particular por lo que se refiere a las condiciones jurídicas esenciales en que ha de tener lugar el tratamiento de los datos. Finalmente, la regulación ha de establecer medidas adecuadas y específicas para proteger la posición jurídica del titular de los datos, de manera que no resultarían suficientes cláusulas generalistas y estereotipadas y, por tanto, las que se establezcan han de tener en cuenta la singularidad de la finalidad de

interés general y las condiciones concretas en que tendrá lugar el tratamiento desde la perspectiva de la integridad de los derechos y libertades de las personas cuyos datos van a ser tratados. Más aún si tenemos en cuenta la interpretación sentada por el Tribunal Constitucional al reconocer que el RGPD establece “una obligación concreta de los Estados miembros de establecer tales garantías, en el caso de que habiliten para tratar los datos personales especialmente protegidos”²¹.

3.3.- El sentido y alcance de la habilitación legal para el tratamiento de los datos personales sin consentimiento: una moneda de dos caras

Como anteriormente se ha explicado, al margen de la categoría especial de los datos referidos en el precedente epígrafe, el artículo 6.1.e) RGPD habilita con carácter general el tratamiento de datos de carácter personal por parte de las Administraciones Públicas para “el cumplimiento de una misión realizada en interés público o en el ejercicio de poderes públicos conferidos al responsable del tratamiento”. Sin embargo, a diferencia de lo previsto para las categorías especiales del artículo 9, lo cierto es que la regulación ahora examinada no establece la necesidad de que la regulación estatal o europea establezcan garantías adicionales o específicas, ya que simplemente se trata de una posibilidad por cuanto se limita a contemplar con carácter facultativo (artículo 6.3.b. pár.2) que:

podrá contener disposiciones específicas para adaptar la aplicación de normas del presente Reglamento, entre otras: las condiciones generales que rigen la licitud del tratamiento por parte del responsable; los tipos de datos objeto de tratamiento; los interesados afectados; las entidades a las que se pueden comunicar datos personales y los fines de tal comunicación; la limitación de la finalidad; los plazos de conservación de los datos, así como las operaciones y los procedimientos del tratamiento, incluidas las medidas para garantizar un tratamiento lícito y equitativo.

A este respecto, el artículo 8.2 LOPD requiere que la utilización de esta base legitimadora del tratamiento —la ejecución de una misión realizada en interés público o en el ejercicio de poderes públicos— sólo

²¹ STC 76/2019, de 22 de mayo, f.j.4.

será admisible “cuando derive de una competencia atribuida por una norma con rango de ley”. Sin embargo, esta exigencia no supone que la disposición deba recoger necesariamente las condiciones en que ha de tener lugar el tratamiento de los datos, bastando simplemente con que el mismo sea inherente a la ejecución de la competencia atribuida legalmente. Ahora bien, el artículo 81 CE establece una reserva legal para los aspectos relativos al contenido esencial del derecho fundamental a la protección de los datos de carácter personal que implica la determinación de los presupuestos y condiciones de la intervención del legislador con previsibilidad y certeza, de manera que estén fijadas legalmente unas garantías adecuadas que han de manifestarse en reglas claras y precisas de conformidad con la interpretación realizada por el Tribunal Constitucional²².

En consecuencia, difícilmente cabría admitir que no existiera una mínima base legal que estableciera, específicamente para cada supuesto, las condiciones jurídicas básicas del tratamiento y, en particular, los sujetos cuyos datos van a ser tratados, las cesiones previstas y los plazos de conservación o, al menos, unos mínimos criterios para su determinación. En todo caso y más allá de su regulación normativa, se trata de extremos que necesariamente han de difundirse a través del Registro de Actividades de Tratamiento, instrumento de obligatoria existencia para las Administraciones Públicas. En última instancia, nos encontramos ante una manifestación del principio de transparencia que ha de regir el tratamiento de los datos personales y, en particular, de una condición para su validez de singular relevancia, sobre todo dado que se hará sin contar con el consentimiento del titular.

En todo caso, la adaptación de la herramienta tecnológica —en este caso la IA— al efectivo cumplimiento de las exigencias derivadas del RGPD no se ha de establecer necesariamente por vía normativa, pudiendo concretarse tales extremos a través de las decisiones de los

²² A este respecto, véase el sintético y claro resumen de la jurisprudencia constitucional que se expone en la STC 76/2019, de 22 de mayo, ff.jj. 5 y 6. Más aún, según se afirma expresamente en el f.j.8 respecto de los datos a que se refiere el artículo 9 RGPD, si la regulación estatal “no prevé esas garantías adecuadas, sino que, todo lo más, se remite implícitamente a las garantías generales contenidas en el Reglamento general de protección de datos, no puede considerarse que haya llevado a cabo la tarea normativa que aquel le exige”.

órganos competentes de la respectiva entidad. Sin embargo, a falta de una regulación específica de carácter general sobre el uso de la IA por parte de las Administraciones Públicas, lo cierto es que ya existen algunas medidas legislativas que han de tenerse en cuenta en función del ámbito y modalidad de la actuación que se vaya a llevar a cabo. Estas disposiciones establecen ciertas garantías de obligado cumplimiento que, en última instancia, se erigen en premisa inexcusable para el válido tratamiento de los datos personales al amparo de lo dispuesto en el artículo 6.1.e) RGPD ya que, en definitiva, forman parte de la habilitación legal que allí se contempla.

Por una parte, el artículo 41 LRJSP concreta una serie de condiciones de validez de las actuaciones administrativas automatizadas, de manera que si el algoritmo utilizado implica un tratamiento automatizado, esto es, sin intervención directa de un empleado público, que se lleve a cabo íntegramente a través de medios electrónicos y en el marco de un procedimiento administrativo, previamente se habrá de determinar cuál sería el órgano competente para “la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente”. En consecuencia, el tratamiento de los datos personales sería inválido tanto si dicho órgano no se hubiese establecido como, asimismo, cuando, a pesar de haberse determinado, no hubiese adoptado las decisiones correspondientes en ejercicio de su competencia. Por lo que se refiere a las obligaciones que contempla el RGPD exigiría, cuando menos y con el alcance que se analizará en el epígrafe 5 de este trabajo, haber puesto en marcha las medidas técnicas y organizativas adecuadas (artículo 24), implantar las medidas de seguridad apropiadas (artículo 32), asegurar el cumplimiento de la privacidad por diseño y por defecto (artículo 25) y del resto de principios generales (artículo 5), que se lleve a cabo la evaluación de impacto cuando sea preceptiva (artículo 35) y, en el caso de externalización del diseño de la herramienta de IA, que la entidad elegida cumpla con las garantías propias del encargado de tratamiento si el servicio a prestar conlleva el manejo de datos personales (artículo 28).

Por otro lado, cuando la herramienta de IA se pretenda utilizar en el ámbito tributario, la regulación sectorial aplicable también establece importantes medidas a tener en cuenta. En concreto, el artículo 96.4 Ley 58/2003, de 17 de diciembre, General Tributaria (LGT),

exige que los programas y aplicaciones que se utilicen para el ejercicio de las potestades tributarias sean previamente aprobados por la respectiva Administración, lo que necesariamente incluye tener en cuenta los extremos indicados en el párrafo anterior cuando la herramienta implique el tratamiento de datos de carácter personal, esto es, vengan referidos a personas físicas, de manera que su licitud estaría condicionada al efectivo cumplimiento de las previsiones del RGPD. En todo caso, el artículo 95 LGT ha establecido una serie de medidas adicionales específicas para el ámbito tributario cuyo incumplimiento supone la imposibilidad de utilizar los datos a que se refiere dicho precepto legal²³. Este precepto principalmente establece una serie de reglas destinadas a garantizar la licitud de las cesiones de datos desde la perspectiva del principio de finalidad (artículo 5 RGPD) pero, en última instancia, no exime a la Administración tributaria de la aplicación del régimen general, sin perjuicio de las singularidades que puedan existir bien en el propio RGPD bien en la legislación tributaria.

Del mismo modo, las importantes novedades regulatorias en el ámbito de las sanciones en el orden social auspiciadas por la reforma que tuvo lugar a través del Real Decreto-ley 2/2021, de 26 de enero, de refuerzo y consolidación de medidas sociales en defensa del empleo, suponen una importante limitación a la hora de tratar datos de carácter personal utilizando medios tecnológicos que han de tenerse en cuenta necesariamente, así como lo dispuesto en el Real Decreto 688/2021, de 3 de agosto, cuando se pretendan adoptar actuaciones automatizadas en el ejercicio de la función inspectora en el ámbito del Trabajo y la Seguridad Social²⁴.

²³ De gran importancia a estos efectos es la reciente STS 1045/2022, de 20 de julio, Sección 4ª, Sala Contencioso-Administrativo (ECLI: ES:TS:2022:3065), en la que se rechaza la utilización de datos obtenidos al amparo de la normativa tributaria para el ejercicio de otras potestades administrativas por cuanto supone exceder los límites del artículo 95 LGT.

²⁴ Al margen del tratamiento de los datos de los trabajadores, desde el punto de vista de la empresa es necesario tener en cuenta que los tratamientos de información vinculados a los autónomos también se encontrarían sometidos a este condicionante, tal y como recuerda TODOLÍ SIGNES, A. (2022): “Inspección de trabajo, algoritmos, inteligencia artificial y su uso en la planificación inspectora”, en COTINO, L. y TODOLÍ SIGNES, A. (coords.), *Explotación y regulación del uso del big data e inteligencia artificial para los servicios públicos y la ciudad inteligente*, Tirant lo Blanch, Valencia, pp. 239-240.

En definitiva, la existencia de una previsión legal que autorice el tratamiento de datos personales para el desarrollo de una misión de servicio público no sólo implica la correspondiente habilitación para la entidad pública sino, asimismo, el necesario cumplimiento de las garantías adicionales que se hubieran previsto como inexcusable condición de validez de la actuación administrativa.

4.- LAS CONDICIONES PARA LA LICITUD DEL TRATAMIENTO

4.1.- El alcance de los principios generales y su proyección en los tratamientos basados en IA por parte de las Administraciones Públicas

La regulación vigente sobre protección de datos personales se encuentra basada, en gran medida, en el reconocimiento de una serie de principios que, por tanto, han de ser aplicados teniendo en cuenta la singularidad de la tecnología que se vaya a utilizar y que, en última instancia, son concreción de otros a los que debe someterse con carácter general la actuación administrativa, como pueden ser el principio de legalidad o el de proporcionalidad²⁵. Su toma en consideración ha de tener lugar con carácter previo a la puesta en marcha de la herramienta de IA dadas las exigencias de la privacidad por diseño y por defecto (artículo 25 RGPD), lo que exige que sean tenidos en cuenta desde la determinación inicial de los requisitos funcionales²⁶ y en el resto de las fases del ciclo de vida del sistema de IA²⁷; todo ello a pesar

²⁵ IZQUIERDO CARRASCO, M. (2022): “Actuaciones automatizadas en la sanción de irregularidades en el orden social”, en COTINO, L. y TODOLÍ, A. (coords.), *Explotación y regulación del uso del big data e inteligencia artificial para los servicios públicos y la ciudad inteligente*, Tirant lo Blanch, Valencia, p. 279.

²⁶ Como ha enfatizado SORIANO, el alcance de estos principios se proyecta en gran medida sobre los datos que alimentan el sistema de IA. A este respecto, SORIANO ARNAIZ, A. (2021a): *Data protection for the prevention of algorithmic discrimination*, Thomson Reuters-Aranzadi, Cizur Menor, p. 122.

²⁷ Aunque los sistemas suelen tener varias fases bien definidas (concepción y análisis, desarrollo, explotación y retirada), no sólo hay que centrar el cumplimiento del RGPD en el componente de IA, ya que normalmente formará parte de

de las evidentes dificultades que puede suponer su proyección en un entorno singularmente complejo, en particular por lo que se refiere a la multiplicidad de sujetos que suelen intervenir en este tipo de desarrollos tecnológicos²⁸. En última instancia, corresponde al órgano competente de la Administración responsable comprobar si efectivamente se respetan estos principios ya que su vulneración puede determinar la invalidez del tratamiento de datos personales realizado por la herramienta de IA²⁹. Además, según la obligación de responsabilidad proactiva (artículo 5.2 RGPD), su aplicación y la justificación de las decisiones adoptadas en base a los mismos debe poder acreditarse, de manera que será imprescindible documentar las razones que han llevado a adoptar cada una de las decisiones relevantes en cumplimiento del RGPD y, además, hacerlo de manera continuada y no sólo en el momento inicial³⁰.

un tratamiento más amplio que hay que considerar como un todo. Sobre este planteamiento, véase AEPD (2020), *Adecuación al RGPD de tratamientos que incorporan IA*, p. 12.

²⁸ POULLET, Y. (2020): *Le RGPD face aux défis de l'intelligence artificielle*, Lar-cier, Bruselas, p. 133.

²⁹ Por lo que se refiere al alcance de tales principios en el ámbito del uso de sistemas automatizados de vigilancia en el sector público, resulta especialmente significativo el caso de la legislación valenciana sobre la herramienta SALER y la interpretación realizada por la AEPD. En particular desde la perspectiva de la proporcionalidad del tratamiento por una excesiva indeterminación del conjunto de datos, el incumplimiento del principio de finalidad, el excesivo plazo de conservación y, en definitiva, la insuficiencia de las garantías específicamente contempladas en la propuesta legislativa. A este respecto, véase el análisis de CO-TINO HUESO, L.: “Hacia la transparencia 4.0: el uso de la inteligencia artificial y el *big data* para la lucha contra el fraude y la corrupción y las (muchas) exigencias consuetudinarias”, en RAMÍO, C., coord., (2021), *Repensando la Administración digital y la innovación pública*, Instituto Nacional de Administración Pública, Madrid, pp. 186-189.

³⁰ MARTÍNEZ MARTÍNEZ, R. (2019b): “El principio de responsabilidad proactiva y la protección de datos desde el diseño y por defecto”, en GARCÍA MAHAMUT, R. y TOMÁS MALLÉN, B. (eds.), *El Reglamento General de Protección de Datos. Un enfoque nacional y comparado*, Tirant lo Blanch, p. 339. En definitiva, cuando se trata de actuaciones automatizadas no sólo será necesario anticipar el efectivo cumplimiento de las garantías dada su inmediatez sino, asimismo, revisar a posteriori el funcionamiento del sistema. Sobre esta idea, véase ROIG, A. (2020): *Las garantías frente a las decisiones automatizadas. Del Reglamento General de Protección de Datos a la gobernanza algorítmica*, J.M. Bosch, Barcelona, 2020, pp. 196-197.

¿Cuáles son los principios de mayor relevancia y qué incidencia pueden tener en el tratamiento de datos personales por las Administraciones Públicas utilizando herramientas de IA? Para ello debemos partir de lo dispuesto en el artículo 5 RGPD³¹, sin perjuicio de precisar el alcance de los principios allí previstos en función de otras previsiones aplicables que pueden ser de gran utilidad.

4.1.1.- La singular eficacia del principio de transparencia en la normativa sobre protección de datos personales

Aun cuando en el sector público el consentimiento no constituya en general una base jurídica adecuada para legitimar el tratamiento de los datos personales, no por ello deja de ser exigible que se proporcione la información necesaria a los sujetos afectados en los términos a que se refieren los artículos 13 y 14 RGPD, así como el artículo 11 LOPD. Además, de conformidad con lo dispuesto en el artículo 30 RGPD, cada entidad pública habrá de disponer de un registro de las actividades de tratamiento que se lleven a cabo bajo su responsabilidad, que tendrá que estar accesible por medios electrónicos por exigencia del artículo 33 LOPD. En este sentido, tanto si los datos se obtienen directamente o no de la persona titular de los mismos, la Administración habrá de informar acerca de “la existencia de decisiones automatizadas, incluida la elaboración de perfiles”, ofreciendo “información significativa sobre la lógica aplicada, así como la importancia y las consecuencias previstas de dicho tratamiento para el interesado”³². Incluso, en el caso de las entidades públicas se trata de una obligación reforzada por las exigencias de explicabilidad³³ de las

³¹ No sin advertir, con carácter general, acerca de las dificultades de proceder a su efectiva aplicación en el contexto de la IA, tal y como se ha destacado certeramente en COTINO HUESO, L. (2022): “Nuevo paradigma en las garantías de los derechos fundamentales y una nueva protección de datos frente al impacto social y colectivo de la inteligencia artificial”, en COTINO HUESO, L. (dir.), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Thomson Reuters-Araznadi, Cizur Menor, p. 84.

³² Sobre el alcance del carácter significativo de la información, véase AEPD (2020), *ob.cit.*, p. 24.

³³ En definitiva, se trata de “garantizar que la toma de decisiones públicas automatizadas es capaz de producir justificaciones aceptables dentro del marco norma-

decisiones que adopten utilizando componentes de IA ante el riesgo reduplicado que la opacidad supone en este ámbito³⁴, en particular cuando se trate de actuaciones automatizadas³⁵. En consecuencia, desde la perspectiva de la regulación ahora analizada, existe una clara e inequívoca obligación de información que, sin embargo, no está fijada con carácter general para el resto de los supuestos en que la actuación administrativa automatizada no implique el tratamiento de datos de carácter personal³⁶, al menos en el régimen general que contempla el artículo 41 LRJSP.

Adicionalmente, la garantía establecida por el artículo 22 RGPD supone, al menos indirectamente, una medida de transparencia que afecta sustancialmente a la tecnología utilizada. En concreto, según dicho precepto cualquier persona física “tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar”³⁷, de manera que la efectividad de este derecho requiere, en última instancia, la

tivo actual. De lo contrario, deberemos aceptar un cambio de paradigma en las exigencias de motivación de los actos administrativos, aceptando explicaciones no intuitivas”, en cuyo caso “se deberá garantizar el control de los sistemas automatizados empleados en el ámbito público a través de otra clase de mecanismos” (SORIANO ARNAIZ, A., 2021b: “Decisiones automatizadas: problemas y soluciones jurídicas. Más allá de la protección de datos”, *Revista de Derecho Público: Teoría y Método*, vol. 3, p. 102).

³⁴ OROFINO, A.G. (2020): “The Implementation of the Transparency Principle in the Development of Electronic Administration”, *European Review of Digital Administration & Law*, vol. 1, núm. 1-2, 2020, pp. 123-142.

³⁵ Se ha destacado una triple dimensión por lo que se refiere específicamente a la transparencia en el ámbito del sector público: el conocimiento de la existencia del sistema automatizado, de su funcionamiento y su comprensibilidad. Sobre esta idea, VESTRI, G. (2021), *ob.cit.*, p. 388.

³⁶ Con relación a este matiz, cfr. OLIVARES OLIVARES, B. (2020): “Transparencia y aplicaciones informáticas en la Administración tributaria”, *Crónica Tributaria*, núm. 174, p. 108.

³⁷ Para una visión general acerca de la interpretación de las autoridades de control sobre los requisitos que establece este precepto, en particular por lo que se refiere a la automatización y a las singularidades en el ámbito de la Administración tributaria desde una perspectiva comparada, véase OLIVARES OLIVARES, B. (2022a): “La toma de decisión automatizadas y el derecho fundamental a la protección de los datos de carácter personal”, *Quincena Fiscal*, núm. 15-16, versión electrónica.

posibilidad de conocer que dicho tratamiento tendría lugar. Ahora bien, de conformidad con el precepto citado, este derecho no aplicaría en el caso de que estuviere excluido normativamente, siempre que la regulación que lo establezca adopte las medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, lo que necesariamente obliga partir de las singularidades establecidas por la legislación administrativa, tanto de carácter general como, en su caso, sectorial que anteriormente fueron analizadas en el epígrafe 3.3. En todo caso, es importante enfatizar que las garantías del artículo 22.3 RGPD —esto es, obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión— no resultarán de aplicación en el ámbito de la actividad que desarrollan las Administraciones Públicas, tal y como se deduce con claridad de la dicción literal del propio precepto.

4.1.2.- El alcance de las limitaciones impuestas por el principio de minimización

El principio de minimización supone que tanto en el propio tratamiento que lleve a cabo la herramienta de IA como, asimismo, en la fase de entrenamiento previa sólo podrán utilizarse datos de carácter personal en la medida que resulte estrictamente necesario para la finalidad perseguida y, asimismo, se respeten el resto de las garantías examinadas, en particular desde las exigencias de la proporcionalidad³⁸. A este respecto, la aplicación del principio conlleva relevantes restricciones que han de tenerse en cuenta al diseñar el componente de IA, en concreto por lo que se refiere a la tipología de datos que se van a utilizar en las diversas fases, a la granularidad de la información y su antigüedad, al número de personas respecto de las cuales se procesarán los datos o, entre otros, a las categorías de datos que van a ser objeto de tratamiento³⁹.

³⁸ En este sentido, sobre todo en los sistemas de aprendizaje automático, “no es una práctica proporcional aumentar sustancialmente la cantidad de datos personales en el conjunto de datos de entrenamiento para tener únicamente una ligera mejora en el rendimiento de los sistemas”, hasta el punto de que “más datos podrían implicar una generación de más sesgo en el modelo”, según se afirma en AEPD-EDPS (2022), *10 malentendidos...*, *ob.cit.*, p. 3.

³⁹ A este respecto, AEPD, 2020, *ob.cit.*, p. 39.

En particular, en el caso de no resultar adecuado el uso de datos anonimizados en función de la singularidad y exigencias de la tecnología utilizada, la Administración Pública responsable del tratamiento deberá justificar por qué no sería suficiente con el uso de datos seudonimizados, ya que se trata de un criterio general que, cuando menos, debería valorarse al amparo de lo dispuesto en los artículos 25 y 32 RGPD. En todo caso, la seudonimización puede estar contemplada expresamente en la regulación aplicable como una condición esencial para el tratamiento de los datos⁴⁰, de manera que en estos casos nos encontraríamos ante una exigencia adicional de cumplimiento necesario si se pretende basar la licitud del tratamiento en dicha regulación, no pudiendo considerarse que constituya una mera recomendación u orientación a tener en cuenta a la hora de valorar las medidas de seguridad a adoptar por el responsable del tratamiento.

4.1.3.- La prohibición de usos incompatibles y la limitación del tiempo de conservación de los datos

Como antes se ha indicado, la base legitimadora del tratamiento supone un importante condicionamiento por lo que se refiere a la limitación del origen de los datos que vayan a ser objeto de tratamiento. En efecto, según el artículo 5 RGPD sólo podrán ser utilizados para finalidades que no resulten incompatibles⁴¹ con la que justificara inicialmente la recogida y tratamiento de los datos, lo que obliga a hacer un análisis caso por caso para determinar si procede o

⁴⁰ Este sería el caso, por ejemplo, de la investigación en salud pública y biométrica realizada al amparo de lo previsto en la disposición adicional decimoséptima LOPD, tal y como se adelantó al comienzo de este trabajo.

⁴¹ Como se ha destacado con acierto, se trata de un concepto que corresponde concretar en cada caso al responsable del tratamiento, lo que le otorga una relevante capacidad de decisión y, asimismo, una amplia flexibilidad (SORIANO ARANAIZ, A., 2021a, *ob.cit.*, p. 120). En el caso de las entidades públicas esta posición se encuentra reforzada por la presunción de legalidad de sus actos, que puede ser discutida no sólo a través de la impugnación formal de la decisión adoptada —en el caso de que exista— o del resultado de la actuación material, sino que también cabría acudir ante la autoridad de control competente, cuya decisión sería de obligado cumplimiento incluso para la Administración Pública responsable del tratamiento salvo que fuera objeto de la correspondiente impugnación y su pretensión fuera estimada.

no utilizar cada conjunto de datos en función del marco normativo que resulte de aplicación al contexto de donde proviniera inicialmente la información.

Todo ello, claro está, con las matizaciones necesarias cuando se trate de actividades de vigilancia y persecución de infracciones penales en los términos y con las garantías que establezca la normativa aplicable a la actividad que desarrollen en el ámbito judicial y policial o, en su caso, para la salvaguarda de la seguridad pública, la seguridad del Estado y la defensa. En concreto, a estos efectos debe partirse de la exigencia del artículo 23 RGPD que, para limitar el alcance de los principios del artículo 5 en conexión con las garantías de los artículos 12 a 22, requiere que existan condiciones y garantías adecuadas como contrapeso inexcusable, medidas que habrán de ser establecidas normativamente, ya por el Derecho de la Unión Europea ya de los Estados miembros. En todo caso, debe tratarse de una decisión necesaria y proporcionada en una sociedad democrática con el fin de salvaguardar “objetivos importantes de interés público” o la seguridad del Estado.

Por otra parte, aunque expresamente se contempla que la conservación de los datos con finalidades de archivo en interés público no se considera incompatible —artículo 5.1.b) RGPD—, sin embargo, dicha excepción se refiere exclusivamente a los fines que inicialmente justificaron el tratamiento de los datos. En consecuencia, si se pretendieran utilizar datos que obren en poder de otras entidades sería necesario justificar la licitud de la comunicación, lo que puede suponer una importante limitación en los entornos basados en modelos *big data*, salvo que previamente se hubiera procedido a su anonimización. Este mismo problema se dará desde la perspectiva del tiempo de conservación de los datos a tenor de lo dispuesto en el artículo 5.1.b) RGPD que, en definitiva, impide el uso de aquella información que se hubiese mantenido vinculada a personas físicas durante más tiempo del necesario para los fines de su tratamiento inicial. Se trata de una exigencia que resultará imprescindible valorar a la hora de que la Administración responsable de la herramienta de IA —o la empresa que lo haga en su lugar— proceda a su diseño y, en su caso, a plantear el análisis de riesgos y realizar la evaluación de impacto. En definitiva, no podrán utilizarse datos personales, ni si quiera seudonimizados, si

previamente no se ha comprobado la licitud de su uso⁴², en concreto por lo que se refiere al eventual incumplimiento del tiempo máximo de su conservación.

4.2.- El efectivo cumplimiento de las normas de seguridad: la relevancia del análisis de riesgos y la evaluación de impacto

La licitud del tratamiento de datos personales utilizando medios electrónicos requiere inexcusablemente el respeto de unas mínimas condiciones de seguridad, especialmente si, como sucede con frecuencia en el caso de las Administraciones Públicas, el titular de los datos ni siquiera puede negarse al tratamiento al llevarse a cabo en el ejercicio de una misión de servicio público o ser inherente a una potestad pública. A este respecto, la efectividad de las garantías jurídicas en el actual contexto de complejidad de los entornos digitales radica sustancialmente en el efectivo respeto a las reglas técnicas que se establezcan, exigencia que ha de observarse al menos desde una doble consideración. Así, en primer lugar, las normas jurídicas deben establecer un nivel adecuado de protección desde la perspectiva de los estándares tecnológicos en función de la naturaleza de los riesgos existentes, lo que podría suponer una dificultad añadida en el caso de ciertos sistemas de IA. Como afirmamos en otro lugar, “de nada sirve que la norma fije estándares muy garantistas si en la práctica su incumplimiento no conlleva consecuencias jurídicas lo suficientemente contundentes como para desincentivar las conductas contrarias a las previsiones normativas. Y no se trata de establecer sanciones más o menos elevadas sino, sobre todo, de tratar de impedir —ya cautelarmente o, llegado el caso, de manera definitiva— la utilización de la información cuando no se hubiesen cumplido realmente las exigencias derivadas de la seguridad tecnológica”⁴³.

⁴² En definitiva, la Administración como responsable del tratamiento deberá mostrar la debida diligencia en la comprobación de la legitimidad de la fuente de datos utilizada. Sobre esta exigencia, AEPD (2020), *ob.cit.*, p. 22.

⁴³ VALERO TORRIJOS, J. (2017): “Seguridad, tecnología y Administración Pública electrónica: la necesaria reconfiguración del alcance de las garantías jurídicas ante la innovación tecnológica”, en GIL, J.R. y CRIADO, J.I. (eds.), *Tecnologías*

A este respecto, el artículo 32 RGPD establece la obligación inequívoca de que las Administraciones Públicas, en su condición de responsables del tratamiento de datos personales, adopten las “medidas técnicas y organizativas apropiadas para garantizar un nivel de seguridad adecuado al riesgo”, para lo cual habrán de tener en cuenta, entre otros extremos, la naturaleza y el contexto tecnológico del tratamiento, junto con los riesgos de probabilidad y gravedad para los derechos y libertades. Así pues, el órgano competente de la respectiva entidad deberá realizar una evaluación de tales circunstancias a la hora de decidir la implantación de cada concreta solución tecnológica basada en IA, ponderando las anteriores circunstancias en función de la singularidad del sistema que se pretenda implantar. Se trata, en fin, de una condición jurídica esencial para la validez del tratamiento según la configuración normativa del derecho fundamental analizado. De manera que bajo ningún concepto podría admitirse la simple concurrencia de una base legítima adecuada según los artículos 6 y 9 RGPD y el respeto a los principios que establece el artículo 5 RGPD —con las obligaciones derivadas de este precepto, en los términos antes analizados— para justificar la licitud del tratamiento de los datos puesto que se estaría ante la infracción de una condición esencial de carácter sustantivo que supondría, en definitiva, la vulneración del derecho fundamental garantizado por el artículo 18.4 CE. Adicionalmente, por exigencia del legislador estatal, idéntica consecuencia se daría cuando se ignoren las medidas de seguridad establecidas en el Esquema Nacional de Seguridad, instrumento normativo que por exigencia de la disposición adicional primera LOPD resulta de obligado cumplimiento para las entidades del sector público cuando utilicen medios electrónicos en el desarrollo de su actividad.

La obligación de adoptar efectivamente estas medidas no puede convertirse, por tanto, en un mero trámite formal desprovisto de un previo análisis material que, en definitiva, exige la ponderación de los riesgos a partir de las concretas circunstancias en que tiene lugar el tratamiento. De ahí que resulte imprescindible partir de la singularidad de la herramienta tecnológica utilizada en cada caso, ya que

no sería irrelevante que el tratamiento de los datos se realice por un algoritmo predictivo o que se empleen sistemas de aprendizaje autónomo o basados en redes neuronales. Tampoco se trata de una obligación que pueda diferirse temporalmente sino que, por el contrario, según las exigencias del principio de privacidad por diseño y por defecto (artículo 25 RGPD), su cumplimiento debería verificarse ya desde el momento inicial al precisar los requisitos funcionales de la herramienta tecnológica utilizada. Asimismo, se ha de garantizar la participación del Delegado de Protección de Datos de la correspondiente entidad en los términos previstos por el artículo 38 RGPD.

El análisis de los riesgos y la determinación de las medidas a adoptar requiere tener en cuenta el alcance de la regulación del artículo 35 RGPD sobre la evaluación de impacto relativa a la protección de datos. En concreto, según este precepto, su realización será obligatoria antes de iniciar el tratamiento cuando su naturaleza, alcance, contexto o fines, especialmente si utiliza nuevas tecnologías, entrañe un alto riesgo para los derechos y libertades de las personas físicas. A este respecto, el artículo 35.4 RGPD requiere que las autoridades de control determinen en qué supuestos dicha evaluación será preceptiva, previsión que ha dado lugar a la publicación del correspondiente listado por parte de la AEPD⁴⁴. En concreto, por lo que se refiere al objeto de este trabajo, entrarían dentro del ámbito de la obligación aquellos instrumentos que utilicen soluciones de IA que cumplan al menos con dos de los siguientes criterios:

- supongan la elaboración de perfiles,
- impliquen la adopción de decisiones automatizadas o contribuyan en gran medida a las mismas,
- determinen la observación, monitorización, supervisión, geolocalización o control de personas de forma sistemática y exhaustiva,
- impliquen el uso de categorías especiales de datos a las que se refiere el artículo 9.1 RGPD, en particular el uso de datos bio-

⁴⁴ Disponible en <https://www.aepd.es/sites/default/files/2019-09/listas-dpia-es-35-4.pdf> (última visita: 15/09/2022).

métricos con el propósito de identificar de manera única a una persona física,

- conlleven el uso de datos a gran escala,
- impliquen la asociación, combinación o enlace de registros de bases de datos de dos o más tratamientos con finalidades diferentes o por responsables distintos,
- afecten a sujetos de sujetos vulnerables o en riesgo de exclusión social,
- supongan la utilización de nuevas tecnologías o un uso innovador de tecnologías consolidadas, lo que sin duda afecta al uso de sistemas de aprendizaje automático y redes neuronales.

Este listado debe completarse con el que la propia AEPD ha publicado al amparo de lo dispuesto en el artículo 35.5 RGPD, relativo a los supuestos en los que no sería necesaria la evaluación⁴⁵, que precisamente incluye una previsión específicamente dirigida a las Administraciones Públicas. En concreto, se considera que no se aplicaría esta obligación si el tratamiento es necesario para la realización de una misión de interés público o el ejercicio de potestades públicas si en el propio mandato legal no se obliga a realizarla y “siempre y cuando ya se haya realizado una EIPD completa”. Así pues, aun cuando podría considerarse que las entidades del sector público estarían exoneradas de dicha obligación cuando no lo contemple la legislación que le hubiese otorgado la potestad pública o encargado la realización de la actividad de interés público, lo cierto es que el inciso final entrecomillado ha de interpretarse según lo dispuesto en el artículo 35.10 RGPD que, en definitiva, contempla que la exoneración sólo se daría cuando “ya se haya realizado una evaluación de impacto relativa a la protección de datos como parte de una evaluación de impacto general en el contexto de la adopción de dicha base jurídica completa”. En definitiva, si la evaluación hubiese tenido lugar durante el proceso de tramitación de la normativa estatal o europea que sirva de base

⁴⁵ Disponible en <https://www.aepd.es/sites/default/files/2019-12/ListasDPIA-35.51.pdf> (última visita: 15/09/2022).

jurídica para realizar el tratamiento al amparo de lo dispuesto en el artículo 6.1.e) RGPD⁴⁶.

La evaluación de los riesgos y la determinación de las medidas de seguridad que deban adoptarse ha de llevarse a cabo a través de una metodología adecuada, lo que en la práctica supone utilizar los criterios y estándares que haya establecido la autoridad de control⁴⁷. Ahora bien, a diferencia de lo que sucede con el Esquema Nacional de Seguridad, dado que no se trata de auténticas normas jurídicas en sentido estricto su incumplimiento no puede llevar a la conclusión de que la evaluación sea contraria a Derecho. Sin embargo, serviría al menos para establecer un estándar de referencia cuyo incumplimiento podría dar lugar a la imposición de una sanción o, en su caso, la adopción de una medida cautelar consistente en la suspensión del tratamiento que se esté realizando sin adoptar las medidas de seguridad adecuadas según el criterio del regulador, ya sea europeo o nacional, al amparo de los poderes correctivos que contempla el artículo 58.2 RGPD.

Finalmente, como parte del análisis que debe realizarse al llevar a cabo una evaluación de impacto, el artículo 35.7 RGPD requiere que se evalúe de manera específica la necesidad y proporcionalidad de las operaciones de tratamiento con respecto a su finalidad, lo que en definitiva implica que también se tenga en cuenta la idoneidad de la solución tecnológica de IA elegida y, en particular, su naturaleza y configuración. A este respecto, lejos de ser admisibles meras afirmaciones genéricas, deberá demostrarse que la concreta herramienta supera el filtro que supone el cumplimiento material de tales exigencias⁴⁸. En definitiva, el empleo de soluciones basadas en el uso de IA sitúa a la

⁴⁶ En este sentido, véase RODRÍGUEZ LÓPEZ DE LEMUS, P.: “¿Están los organismos públicos exceptuados de realizar EIPD?”, opinión publicada en el siguiente enlace de su blog: <https://lopezdelemus.com/estan-los-organismos-publicos-exceptuados-de-realizar-eipd/> (última visita: 15/09/2022).

⁴⁷ AEPD (2021b), *ob.cit.*

⁴⁸ A este respecto, la AEPD (2020), *ob.cit.*, p. 44, sostiene que “la disponibilidad o novedad de una tecnología no justifica, por sí misma, su utilización, sino que debe ser objeto de ponderación, realizando un análisis de si el tratamiento, en la forma en que se plantea, es equilibrado porque se derivan más beneficios y ventajas concretas para el interés general y la sociedad en su conjunto que perjuicios, entendidos estos como los riesgos sobre los derechos y libertades de los sujetos cuyos datos son objeto de tratamiento”.

Administración Pública en tanto que responsable del tratamiento de datos personales en un escenario en el que el tratamiento, por sus características y, en particular, por la tecnología elegida, puede conllevar un alto nivel de riesgo. Por lo tanto, debería valorarse si el objetivo del tratamiento no puede ser conseguido utilizando otro tipo de solución que alcance la misma funcionalidad, con un margen de rendimiento aceptable y un nivel de riesgo menor⁴⁹. En definitiva, más allá del mero cumplimiento de un trámite de gran relevancia, la evaluación de impacto se convierte en una oportunidad inestimable para integrar el cumplimiento de este derecho fundamental en la planificación estratégica y la gobernanza de la correspondiente Administración⁵⁰, debiendo garantizarse, en su caso, que en la realización del trámite se han habilitado los cauces de participación oportunos según establece el artículo 35.9 RGPD⁵¹.

4.3.- La externalización en la gestión de las herramientas basadas en el uso de IA: las condiciones jurídicas en el encargo del tratamiento de datos personales

Con frecuencia la Administración precisará de una entidad privada que le preste un servicio vinculado al diseño y/o utilización de la herramienta de IA⁵² que, en consecuencia, en la medida que

⁴⁹ Ibídem.

⁵⁰ OLIVARES OLIVARES, B. (2022b): “Implicaciones de la normativa sobre protección de datos en el desarrollo de la inteligencia artificial por la administración tributaria: la gobernanza de los datos”, en MORENO GONZÁLEZ, S. (Dir.), *Nuevas tecnologías disruptivas y tributación*, Thomson Reuters-Aranzadi, Cizur Menor, p. 109.

⁵¹ En todo caso, más allá de las garantías que específicamente contempla la normativa sobre protección de datos personales, la participación social en la toma de decisiones relativa al uso de la IA por parte de las entidades públicas constituye una elemental exigencia del principio democrático. Sobre esa idea, véase TERRÓN SANTOS, D. y DOMÍNGUEZ ÁLVAREZ, J.L. (2022): *i-Administración Pública, sistemas algorítmicos y protección de datos*, Iustel, Madrid, pp. 180-181.

⁵² Como ha enfatizado al comienzo de este trabajo, en el caso de la IA la dependencia de empresas privadas por parte de la Administración Pública es incluso más intensa, lo que supone una dificultad a la hora de que esta última asuma sus obligaciones como responsable del tratamiento, en particular cuando se trate de exigir a las empresas el cumplimiento de las suyas como encargada. Esta dificul-

requiera el tratamiento de datos personales, tendrá la consideración de encargada del tratamiento a los efectos de lo dispuesto en el artículo 28 RGPD y el artículo 33 LOPD⁵³. En estos casos, según dispone la normativa europea, la empresa adjudicataria del servicio deberá ofrecer “garantías suficientes para aplicar medidas técnicas y organizativas apropiadas, de manera que el tratamiento sea conforme con los requisitos del presente Reglamento y garantice la protección de los derechos del interesado”⁵⁴. Asimismo, hay que tener en cuenta las previsiones específicas que establece la legislación sobre contratación del sector público con relación a las medidas de seguridad que deben adoptarse⁵⁵, debiéndose advertir que no siempre coincidirá el órgano de contratación y el que ha de considerarse encargado del tratamiento⁵⁶.

A este respecto, sería necesaria la celebración de un contrato en el que la Administración, como responsable del tratamiento, establezca

tad puede tener consecuencias muy relevantes desde la perspectiva de la privacidad por diseño y por defecto o, en su caso, la supervisión de la realización de la evaluación de impacto de las herramientas tecnológicas desarrolladas cuando sea obligatoria.

⁵³ Para una visión de las obligaciones que corresponden al encargado del tratamiento específicamente en estos entornos, véase AEPD (2020), *ob.cit.*, pp. 17-19.

⁵⁴ Más aún, “es de vital importancia que [...] se definan adecuadamente todas las exigencias legales a las que debe ajustarse el sistema, se prevea la presencia de juristas en los equipos de trabajo y se instauren mecanismos de verificación del cumplimiento de tales requerimientos con ocasión de la ejecución de los contratos”, tal y como ha enfatizado GAMERO CASADO, E. (2022): “Supervisión, auditoría y control jurídico en la contratación pública de soluciones de robotización e inteligencia artificial para soporte a la toma de decisiones”, *Observatorio de la Contratación del Sector Público*, 11/10/2022, versión electrónica accesible en <https://www.obcp.es/opiniones/supervision-auditoria-y-control-juridico-en-la-contratacion-publica-de-soluciones-de> (última visita: 14/10/2022).

⁵⁵ CANALS AMETLLER, D. (2021): “La seguridad digital en las pequeñas y medianas entidades locales: hacia una gestión municipal colaborativa”, en FONDEVILA, J. (Dir.), *Transformación digital en las medidas y pequeñas entidades locales*, Wolters Kluwer-El Consultor, Madrid, pp.251 y 252.

⁵⁶ Por lo que se refiere a la forma de resolver los problemas que plantea esta dualidad, véase FARRÉ TOUS, S. (2021): “El encargado del tratamiento y los contratos de encargado del tratamiento”, en TRONCOSO REIGADA, A. (Dir.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, tomo I, Thomson Reuters-Civitas, Madrid, pp. 1864-1865.

el objeto, la duración, la naturaleza y la finalidad del mismo, el tipo de datos personales y categorías de interesados, así como las obligaciones y derechos de ambas partes. En concreto, y al margen del resto de las previsiones del artículo 28 RGPD, dicho contrato debería contemplar las siguientes obligaciones de la empresa adjudicataria. En primer lugar, exigirle la aplicación de las medidas técnicas y organizativas apropiadas teniendo en cuenta los especiales riesgos a que se encuentren sometidos los datos, que habrán de revisarse y actualizarse cuando sea preciso. Dichas medidas han de permitir garantizar y demostrar que el tratamiento que hace de los datos personales es conforme con el RGPD. En particular, la Administración debe comprobar que la empresa adopta las medidas necesarias adecuadas, según el artículo 32 RGPD, para afrontar los riesgos derivados de un eventual problema de seguridad que implique el compromiso de los datos personales.

Por lo que se refiere a la seguridad de la información, la empresa se ha de comprometer y será capaz de demostrar de manera fehaciente ante la Administración —y, en su caso, ante la entidad que esta designe para la realización de inspecciones o auditorías— que respeta las exigencias sobre seguridad de los tratamientos a que se refiere el artículo 32 RGPD. Asimismo, según la disposición adicional primera LOPD, habrá de adoptar las medidas que fija el Esquema Nacional de Seguridad. No basta, por tanto, con la mera presentación de un documento genérico sino, por el contrario, se han de contemplar obligaciones precisas que permitan a la Administración Pública comprobar de manera exhaustiva el cumplimiento de dichas exigencias.

En todo caso, los datos a los que acceda la empresa para prestar el servicio sólo podrán utilizarse para la finalidad indicada, no estando permitido su uso para cualquier otra y mucho menos su cesión a terceros. Asimismo, deberá garantizarse que solo se pueda acceder a los datos por parte de aquellas personas de la empresa contratista, o de sus aplicaciones, que lo precisen conforme a las exigencias del principio de minimización. Igualmente, deberá demostrar ante la Administración que los empleados que vayan a manejar la información se han comprometido a respetar la confidencialidad y han recibido la formación oportuna en relación con el cumplimiento del marco normativo sobre protección de datos. En el supuesto de que la entidad que gestionara la aplicación de IA contemplara la colaboración de otra en el tratamiento de la información, resultará imprescindible

contar con la autorización previa por escrito de la Administración tal y como requiere el artículo 28.2 RGPD. Este tipo de supuestos se dará con frecuencia cuando el funcionamiento de la herramienta se articule a través de entornos de gestión en la nube⁵⁷, de manera que los datos podrían estarían almacenados en los servidores de otra entidad que, de estar ubicados en un espacio geográfico ajeno a la Unión Europea, exigirán el cumplimiento de las normas relativas a las transferencias internacionales de datos y la información a las personas afectadas en los términos de los artículos 13 y 14 RGPD. A este respecto ha de tenerse en cuenta la necesidad de que se respeten las normas relativas a las transferencias internacionales de datos⁵⁸, que no sólo vendrán determinadas por las previsiones del RGPD sino que, además, están establecidas en la regulación sobre procedimiento administrativo común por lo que se refiere a los tratamientos que afecten a ciertos sistemas de identificación⁵⁹.

El tratamiento de los datos terminará con la ejecución del contrato. En ese momento, la empresa tendrá que poner a disposición de la Administración toda la información que conserve conforme a las instrucciones que se le indiquen, entendiendo que esta obligación se refiere a la versión actualizada de la información y, en su caso, al histórico de la misma que pueda verse afectada por el cumplimiento de obligaciones legales. En consecuencia, una vez finalice la prestación del servicio, la empresa se debe comprometer a eliminar de forma segura toda la información correspondiente a personas físicas identificadas o identificables, debiendo facilitar que por parte de la Administración se hagan las comprobaciones oportunas o, en su caso, demostrando que efectivamente ha procedido a dar cumplimiento a esta obligación.

⁵⁷ En estos supuestos uno de los riesgos principales se refiere a la posibilidad de que las autoridades de terceros países puedan acceder a la información, tal y como ha destacado NAVARRO CASTRO, M. (2020): “El *cloud computing* como forma de prestación de servicios de tratamiento de datos”, en AA.VV.: *Protección de datos personales*, Tirant lo Blanch, Valencia, pp. 653-654, de manera que resulta de especial importancia establecer una condiciones jurídicas adecuadas para la prestación del servicio a través de los correspondientes pliegos.

⁵⁸ AEPD (2018), *Orientaciones para prestadores de servicios de cloud computing*.

⁵⁹ Véase el artículo 9 LPAC.

5.- LA INCIDENCIA DEL INCUMPLIMIENTO DE LA REGULACIÓN EN MATERIA DE PROTECCIÓN DE DATOS PERSONALES SOBRE LA VALIDEZ DE LA ACTUACIÓN ADMINISTRATIVA

Dada la ausencia de una disciplina de ámbito general relativa al uso de la IA no es sorprendente que la regulación sobre protección de datos personales se haya mostrado como la primera respuesta jurídica antes los riesgos que suscita esta tecnología ya que, en definitiva, los datos son un elemento imprescindible para su funcionamiento⁶⁰. Ahora bien, se trata de un marco normativo que plantea evidentes insuficiencias para este fin⁶¹, especialmente por lo que respecta al sector público, sobre todo porque la utilización del *big data* y la IA tienen un impacto jurídico que va más allá de los derechos fundamentales en su dimensión individual: en última instancia, suscitan una incidencia sistémica que potencialmente afecta al conjunto de la sociedad⁶². Así pues, el incumplimiento de una obligación en principio formal, como la relativa a la realización de una evaluación de impacto, adquiere un significado más profundo que la mera omisión de un trámite subsanable. Por esta razón, más allá de las limitaciones que presenta la configuración normativa por el legislador español de la potestad sancionadora de la AEPD frente a los incumplimientos por parte de las entidades públicas, resulta de especial importancia llevar a cabo un análisis doctrinal en profundidad no sólo del régimen jurídico general del RGPD sino, en concreto, de las consecuencias de su aplicación práctica en el ámbito de las Administraciones Públicas, sobre todo teniendo en cuenta la incidencia sobre un derecho fundamental como es el garantizado por el artículo 18.4 del Texto Constitucional.

Por lo que se refiere al impacto que puede tener el incumplimiento de la regulación sobre protección de datos de carácter personal en la validez de las decisiones administrativas es necesario partir de la consagración legal —artículo 47.1.a) LPAC— de una causa de nulidad de pleno derecho para los actos administrativos que vulneren los dere-

⁶⁰ HUERGO, A. (2020), *ob.cit.*, p. 52.

⁶¹ SORIANO ARANAIZ, A. (2021b), *ob.cit.*, pp. 107-109.

⁶² COTINO HUESO, L. (2022), *ob.cit.*, p. 78.

chos fundamentales⁶³. Aunque con carácter general se haya planteado la anulabilidad como respuesta a los problemas derivados de la programación en las actuaciones administrativas automatizadas⁶⁴, la incidencia del incumplimiento de las garantías relativas a la protección de datos personales no puede limitarse sin más a un problema de programación. A este respecto es importante advertir que, tras la reforma de 1999, la regulación sobre procedimiento administrativo común ya no exige que la vulneración se refiera al contenido esencial del derecho fundamental para entender aplicable la causa de nulidad de pleno derecho, de manera que la indeterminación legal se ha de concretar a la vista del principio de proporcionalidad partiendo de la exigencia jurisprudencial de que la lesión sea real, inequívoca y cierta⁶⁵.

Así, dado su carácter de garantías sustanciales para la efectividad del derecho, se podría considerar que los actos administrativos basados en un sistema de IA son nulos de pleno derecho cuando en la adopción de la decisión:

- a) no se hubiere respetado las limitaciones que establece el artículo 5 RGPD sobre el principio de finalidad y el plazo de conservación en la recopilación de los datos utilizados para entrenar los algoritmos y dicha omisión fuere relevante para el contenido del acto administrativo afectado;
- b) no se hubiesen respetado las normas de seguridad legalmente exigibles y, en particular, las que contempla el Esquema Nacional de Seguridad, puesto que el tratamiento se habrá llevado a cabo sin las garantías necesarias legalmente establecidas;
- c) se careciera de una base jurídica adecuada en función de la naturaleza y alcance del tratamiento, así como en el supuesto de que la Administración responsable del tratamiento no ejerciera un título competencial adecuado para llevarlo a cabo;

⁶³ VALERO TORRIJOS, J. y LÓPEZ PELLICER, J.A. (2011): “Algunas consideraciones sobre el derecho a la protección de datos personales en la actividad administrativa”, *Revista Vasca de Administración Pública*, núm. 59, p. 273.

⁶⁴ BERNING PRIETO, A.D. (2019): *Validez e invalidez de los actos administrativos en soporte electrónico*, Thomson Reuters-Aranzadi, Cizur Menor, p. 268.

⁶⁵ PICÓN ARRANZ, A. (2022), *Las causas de nulidad de pleno derecho del acto administrativo. Configuración legal y aplicación práctica*, Thomson Reuters-Aranzadi, Cizur Menor, pp. 112-114 y 118-119.

- d) hubiere tenido lugar un tratamiento completamente automatizado y no se cumplieran las exigencias del artículo 22 RGPD, ya que se reconoce expresamente un derecho a no verse sometido a decisiones automatizadas salvo, con carácter excepcional, si se respetan las condiciones allí previstas;
- e) no se hubiese llevado a cabo la evaluación de impacto si resultara obligatoria, ya que más que un simple trámite formal se trata de una herramienta cuya finalidad es hacer efectivas las garantías sustanciales más allá de la dimensión estrictamente individual del derecho fundamental, si bien en este caso la nulidad se derivaría del artículo 47.1e) LPAC.

En consecuencia, siguiendo la disciplina legal —artículo 48 LPAC—, el resto de las vulneraciones podrían dar lugar a la anulabilidad, como sucedería, por ejemplo, con el mero incumplimiento del principio de proactividad, salvo que se demostrara que las medidas sustantivas hubieran sido efectivamente adoptadas. Por su parte, la falta de participación del Delegado de Protección de Datos en el proyecto, aun siendo preceptiva para las entidades del sector público, no supondría por sí misma una causa de invalidez, ya que esta ausencia no determinará necesariamente el incumplimiento de la regulación; aunque en el caso de la adjudicación de un contrato a una entidad que carezca de esta figura sí que podría, en cambio, determinar su invalidez cuando se tratara de un requisito esencial para justificar la capacidad y solvencia técnica del contratista. En cambio, la falta de precisión en la información adecuada sobre las condiciones del tratamiento podría considerarse una irregularidad susceptible de ser convalidada, salvo que pueda generar indefensión al ni siquiera ofrecer la información elemental para que el ciudadano pueda, en su caso, ejercer sus derechos o, como antes se indicaba, en el supuesto del artículo 22 RGPD.

Ahora bien, la inexistencia de una obligación legal para que la opción de utilizar una solución de IA se adopte por la Administración Pública previa la tramitación de un procedimiento adecuado y, por tanto, a través de una actuación formalizada, determina una especial complejidad a la hora de impugnar formalmente la decisión administrativa. En todo caso, a pesar de la dificultad de precisar el origen de la actuación administrativa, la causa de invalidez se proyectaría

sobre su aplicación práctica, en concreto su manifestación formalizada mediante el correspondiente acto administrativo que afecte a la persona interesada, bien porque sus datos hubiesen sido objeto de un tratamiento ilícito bien porque se incumplieran las obligaciones normativas en los términos antes indicados.

6.- REFLEXIÓN FINAL

Más allá del enfoque puramente reactivo es imprescindible adoptar un planteamiento más general que se sustente en el liderazgo del poder público en el cumplimiento normativo para el impulso de un modelo de IA basado en el respeto a los derechos y libertades fundamentales y, en definitiva, la dignidad humana⁶⁶. Se ha destacado la necesidad de que el debate sobre los efectos que están generando los modelos algorítmicos y sistemas de IA se base en evidencias, sobre todo para orientar de manera más certera la evolución del uso de dichas tecnologías en el sector público⁶⁷. Y, precisamente, dada la mínima e insuficiente regulación en el ámbito de las Administraciones Públicas, la normativa sobre protección de datos se ha convertido en la “piedra de toque” para determinar hasta qué punto las garantías jurídicas propias de nuestro modelo jurídico-político europeo se han convertido realmente en un obstáculo para la innovación o, por el contrario, constituyen un refuerzo del carácter democrático del uso de la tecnología por parte del sector público.

Para ello resulta esencial que se minimice el impacto de los riesgos y los daños que las tecnologías más disruptivas puedan tener sobre los derechos y libertades, de ahí que resulte especialmente importante la adopción de modelos de gobernanza adecuados que, en última instancia, se han de inspirar en la regulación del RGPD⁶⁸. Ante todo, el cumplimiento efectivo de la normativa sobre protección de datos

⁶⁶ TERRÓN SANTOS, D. y DOMÍNGUEZ ÁLVAREZ, J.L. (2022), *ob.cit.*, pp. 56-62.

⁶⁷ CRIADO, J.I. (2021): “IA (y Administración Pública)”, *Eunomía. Revista en Cultura de la Legalidad*, núm. 20, p. 369.

⁶⁸ VIDA FERNÁNDEZ, J. (2022): “La gobernanza de los riesgos digitales: desafíos y avances en la regulación de la IA”, *Cuadernos de Derecho Transnacional*, vol. 14, núm. 1, pp. 495.

personales requiere un cambio importante en las Administraciones y entidades del sector público, tanto desde la perspectiva organizativa como funcional⁶⁹, de manera que las garantías del derecho fundamental se inserten, de manera natural y desde el primer momento, a la hora de valorar a la puesta en marcha y la ejecución de instrumentos de decisión basados en el uso de IA. Y a tal efecto, más allá de modas y superficiales pátinas de modernidad, la decisión de utilizar esta tecnología debería partir de la constatación de una efectiva necesidad según las exigencias del principio de proporcionalidad, en particular teniendo en cuenta la finalidad perseguida. Esta ponderación se tendría que llevar a cabo en el marco de un procedimiento administrativo formalizado donde instrumentos como la evaluación de impacto, la participación social y la intervención del Delegado de Protección de Datos adquieren su más profundo sentido en tanto que instrumentos esenciales para la efectiva garantía del derecho fundamental analizado en este capítulo.

BIBLIOGRAFÍA Y OTRAS REFERENCIAS CITADAS

- AEPD-EDPS (2022), *Malentendidos sobre el machine learning*,
 AEPD (2021a), *10 malentendidos relacionados con la anonimización*.
 — (2021b), *Gestión del riesgo y evaluación de impacto en tratamiento de datos personales*.
 — (2020), *Adecuación al RGPD de tratamientos que incorporan IA*.
 — (2018), *Orientaciones para prestadores de servicios de cloud computing*.
 — (2016), *Orientaciones y garantías en los procedimientos de anonimización de datos personales*.
 ARTICLE 29 WORKING PARTY (2017), *Guidelines on consent under Regulation 2016/679*.

⁶⁹ En palabras de OLIVARES OLIVARES, B. (2022a), si la Administración “no determina previamente qué es información personal, no analiza qué es una decisión automatizada ni cuándo se produce, no categoriza las aplicaciones en función del riesgo ante un tratamiento masivo de datos de carácter personal, ni tiene establecido un protocolo riguroso para informar a los interesados o ser capaz de dar respuesta con solvencia al ejercicio del derecho de acceso, me temo que podría estar vulnerándose el contenido esencial del derecho a la protección de datos de carácter personal” (versión electrónica).

- BERNING PRIETO, A.D. (2019): *Validez e invalidez de los actos administrativos en soporte electrónico*, Thomson Reuters-Aranzadi, Cizur Menor.
- CANALS AMETLLER, D. (2021a): “La seguridad digital en las pequeñas y medianas entidades locales: hacia una gestión municipal colaborativa”, en FONDEVILA ANTOLÍN, J. (Dir.), *Transformación digital en las medianas y pequeñas entidades locales*, Wolters Kluwer-El Consultor, Madrid, pp. 241-267.
- CANALS AMETLLER, D. (2021b): “La seguridad en el entorno digital”, en CANALS AMETLLER, D. (dir.), *Ciberseguridad. Un nuevo reto para el Estado y los Gobiernos Locales*, Wolters Kluwer-El Consultor, Madrid, 2021, pp. 63-88.
- CERRILLO i MARTÍNEZ, A. (2019): “Com obrir les caixes negres de les administracions públiques? Transparència i rendició de comptes en l’ús dels algorismes”, *Revista Catalana de Dret Públic*, núm. 58. p. 17.
- COTINO HUESO, L. (2022): “Nuevo paradigma en las garantías de los derechos fundamentales y una nueva protección de datos frente al impacto social y colectivo de la inteligencia artificial”, en COTINO HUESO, L. (dir.), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Thomson Reuters-Araznadi, Cizur Menor, pp. 67-105.
- COTINO HUESO, L. (2021): “Hacia la transparencia 4.0: el uso de la inteligencia artificial y el *big data* para la lucha contra el fraude y la corrupción y las (muchas) exigencias consitucionales”, en RAMIÓ, C. (coord.), *Repensando la Administración digital y la innovación pública*, Instituto Nacional de Administración Pública, Madrid, pp. 169-196.
- COTINO HUESO, L. (2017): “Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales”, *Dilemata. Revista Internacional de Éticas Aplicadas*, núm. 24, pp. 131-150.
- CRIADO, J.I. (2021): “Inteligencia artificial (y Administración Pública)”, *Eunomía. Revista en Cultura de la Legalidad*, núm. 20, pp. 348-372.
- FARRÉ TOUS, S. (2021): “El encargado del tratamiento y los contratos de encargado del tratamiento”, en TRONCOSO REIGADA, A. (Dir.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, tomo I, Thomson Reuters-Civitas, Madrid, pp. 1845-1873.
- GAMERO CASADO, E. (2022): “Supervisión, auditoría y control jurídico en la contratación pública de soluciones de robotización e inteligencia artificial para soporte a la toma de decisiones”, *Observatorio de la Contratación del Sector Público*, 11/10/2022. Accesible en <https://www.obcp.es/opiniones/supervision-auditoria-y-control-juridico-en-la-contratacion-publica-de-soluciones-de> (última visita: 14/10/2022).

- HUERGO LORA, A. (2020): “Una aproximación a los algoritmos desde el Derecho Administrativo”, en HUERGO LORA, A. (dir.), *La regulación de los algoritmos*, Thomson-Aranzadi, Cizur Menor, pp. 23-87.
- IZQUIERDO CARRASCO, M. (2022): “Actuaciones automatizadas en la sanción de irregularidades en el orden social”, en COTINO HUESO, L. y TODOLÍ SIGNES, A. (coords.), *Explotación y regulación del uso del big data e inteligencia artificial para los servicios públicos y la ciudad inteligente*, Tirant lo Blanch, Valencia, pp. 261-288.
- MARTÍN DELGADO, I. (2020): “Innovación tecnológica e innovación administrativa en la contratación pública”, en MARTÍN DELGADO, I. y MORENO MOLINA, J.A. (dirs.), *Administración electrónica, transparencia y contratación pública*, Iustel, Madrid, pp. 19-54.
- MARTÍNEZ MARTÍNEZ, R. (2019a): “Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo”, *Revista catalana de Derecho Público*, núm. 58, pp. 64-81.
- MARTÍNEZ MARTÍNEZ, R. (2019b): “El principio de responsabilidad proactiva y la protección de datos desde el diseño y por defecto”, GARCÍA MAHAMUT, R. y TOMÁS MALLÉN, B. (eds.), *El Reglamento General de Protección de Datos. Un enfoque nacional y comparado*, Tirant lo Blanch, Valencia, pp. 310-341.
- OLIVARES OLIVARES, B. (2022a): “La toma de decisión automatizadas y el derecho fundamental a la protección de los datos de carácter personal”, *Quincena Fiscal*, núm. 15-16, versión electrónica.
- OLIVARES OLIVARES, B. (2022b): “Implicaciones de la normativa sobre protección de datos en el desarrollo de la inteligencia artificial por la administración tributaria: la gobernanza de los datos”, en MORENO GONZÁLEZ, S. (Dir.), *Nuevas tecnologías disruptivas y tributación*, Thomson Reuters-Aranzadi, Cizur Menor, pp. 89-112.
- OLIVARES OLIVARES, B. (2020): “Transparencia y aplicaciones informáticas en la Administración tributaria”, *Crónica Tributaria*, núm. 174, p. 89-111.
- OROFINO, A.G. (2020): “The Implementation of the Transparency Principle in the Development of Electronic Administration”, *European Review of Digital Administration & Law*, vol. 1, núm. 1-2, pp. 123-142.
- PICÓN ARRANZ, A. (2022): *Las causas de nulidad de pleno derecho del acto administrativo. Configuración legal y aplicación práctica*, Thomson Reuters-Aranzadi, Cizur Menor.
- RODRÍGUEZ LÓPEZ DE LEMUS, P.: “¿Están los organismos públicos exceptuados de realizar EIPD?”, accesible en <https://lopezdelemus.com/estan-los-organismos-publicos-exceptuados-de-realizar-eipd/> (última visita: 15/09/2022)

- POULLET, Y. (2020): *Le RGPD face aux défis de l'intelligence artificielle*, Larcier, Bruselas.
- RAMÍO, C. (2018): *Inteligencia artificial y administración pública: Robots y humanos compartiendo el servicio público*, Catarata, Madrid.
- ROIG, A. (2020): *Las garantías frente a las decisiones automatizadas. Del Reglamento General de Protección de Datos a la gobernanza algorítmica*, J.M. Bosch, Barcelona.
- SANTOS MORÓN, M.J. (2020): “Tratamiento de datos, sujetos implicados, responsabilidad proactiva”, en AA.VV.: *Protección de datos personales*, Tirant lo Blanch, Valencia.
- SARTOR, G. (2020): *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*, European Parliamentary Research Service, Bruselas.
- SORIANO ARANAIZ, A. (2021a): *Data protection for the prevention of algorithmic discrimination*, Thomson Reuters-Aranzadi, Cizur Menor.
- SORIANO ARANAIZ, A. (2021b): “Decisiones automatizadas: problemas y soluciones jurídicas. Más allá de la protección de datos”, *Revista de Derecho Público: Teoría y Método*, vol. 3, pp. 85-127.
- TERRÓN SANTOS, D. y DOMÍNGUEZ ÁLVAREZ, J.L. (2022): *i-Administración Pública, sistemas algorítmicos y protección de datos*, Iustel, Madrid.
- TODOLÍ SIGNES, A. (2022): “Inspección de trabajo, algoritmos, inteligencia artificial y su uso en la planificación inspectora”, en COTINO HUESO, L. y TODOLÍ SIGNES, A. (coords.), *Explotación y regulación del uso del big data e inteligencia artificial para los servicios públicos y la ciudad inteligente*, Tirant lo Blanch, Valencia, pp. 236-259.
- VALERO TORRIJOS, J. (2022): “Los derechos en la era digital”, en RODRÍGUEZ AYUSO, J.F. (Coord.), *Nuevos retos en materia de derechos digitales en un contexto perspectiva multidisciplinar*, Thomson Reuters-Aranzadi, Cizur Menor, pp. 22-37.
- VALERO TORRIJOS, J. (2017): “Seguridad, tecnología y Administración Pública electrónica: la necesaria reconfiguración del alcance de las garantías jurídicas ante la innovación tecnológica”, en GIL, J.R. y CRIADO, J.I. (eds.), *Tecnologías de información y comunicación en la Administración Pública: conceptos, enfoques, aplicaciones y resultados*, Infotec, México, pp. 135-159.
- VALERO TORRIJOS, J. y LÓPEZ PELLICER, J.A. (2011): “Algunas consideraciones sobre el derecho a la protección de datos personales en la actividad administrativa”, *Revista Vasca de Administración Pública*, núm. 59, pp. 255-286.
- VESTRI, G. (2021): “La Inteligencia artificial ante el desafío de la transparencia algorítmica. Una aproximación desde la perspectiva jurídico-admi-

- nistrativa”, *Revista Aragonesa de Administración Pública*, núm. 56, pp. 368-398.
- VIDA FERNÁNDEZ, J. (2022): “La gobernanza de los riesgos digitales: desafíos y avances en la regulación de la Inteligencia artificial”, *Cuadernos de Derecho Transnacional*, vol. 14, núm. 1, pp. 489-503.
- ZLOTNIK, A. (2019): “Inteligencia artificial en las Administraciones Públicas: definiciones, evaluación de viabilidad de proyectos y áreas de aplicación”, *Boletic*, núm. 84, pp. 24-32.

Capítulo IX

Las garantías de régimen jurídico del sector público y del procedimiento administrativo común frente a la actividad automatizada y la inteligencia artificial

EDUARDO GAMERO-CASADO¹

*Catedrático de Derecho Administrativo
Universidad Pablo de Olavide, de Sevilla*

Resumen: La implantación de soluciones de IA ha de respetar las garantías establecidas en el Derecho administrativo para resguardar la posición de los ciudadanos ante las decisiones que afecten a sus derechos e intereses legítimos. Esto tiene particular trascendencia en relación con el procedimiento administrativo, una de cuyas funciones es garantizar los derechos de la ciudadanía, a cuyo fin establece una serie de trámites mediante los que preservar el principio contradictorio, evitando que se produzca indefensión. Asimismo, debe respetar ciertas previsiones relativas al régimen jurídico del sector público, establecidas en la Ley 40/2015 en relación con la actividad administrativa automatizada. En este capítulo se analizan estas garantías, y la posibilidad de modularlas –pero no de ignorarlas– cuando se implantan sistemas de IA. Si se vulneran estas garantías, generando indefensión, las decisiones adoptadas mediante IA son inválidas.

Abstract: The implementation of AI solutions has to respect the guarantees established in administrative law to safeguard the position of citizens with regard to decisions affecting their rights and legitimate interests. This is particularly important in relation to the administrative procedure, one of its functions is to guarantee the rights of citizens, to which end it establishes a set of steps to safeguard the adversarial principle, avoiding the possibility of defencelessness. It must also respect certain provisions relating to the legal regime of the public sector, established in Law 40/2015 in relation to automated administrative decision-making. This chapter discusses these guarantees, and the possibility of modulating - but not disregarding - them when implementing AI systems. If these guarantees are violated, generating defencelessness, the decisions taken through IA are invalid.

Palabras-clave: inteligencia artificial, procedimiento administrativo, motivación, alegaciones, tutela judicial, indefensión, actividad administrativa automatizada

Key-words: artificial intelligence, administrative procedure, motivation, right to be heard, defencelessness, effective judicial review, automated decision-making

¹ ORCID: 0000-0002-6734-8672.

SUMARIO.- 1.- INTRODUCCIÓN. 1.1.- Acotación material: sistemas de IA para soporte a la toma de decisiones. 1.2.- Las especialidades de la IA en el sector público y su sujeción al Derecho administrativo general sin necesidad de disposiciones que lo establezcan expresamente. 2.- ACTIVIDAD ADMINISTRATIVA AUTOMATIZADA: PREVISIONES EN LA LEY 40/2015, DE RÉGIMEN JURÍDICO DEL SECTOR PÚBLICO. 2.1. El concepto de “actuación administrativa automatizada” y los grados de automatización. 2.2. La aprobación de la implantación del sistema. 2.3. El órgano competente para implementar los sistemas de IA y atribuirle su actividad. Firma electrónica de decisiones automatizadas. 2.4. La supervisión, control de calidad y auditoría del sistema de información y de su código fuente. 2.4.1.- Encuadre general. El deber legal de implantar estos controles de la IA. 2.4.2.- Control de calidad y auditoría del sistema y de su código fuente. 2.4.3.- Supervisión humana. La necesidad de *compliance* jurídico. 2.4.4.- Incumplimiento del requisito de mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema: invalidez de la actuación automatizada por vulneración del art.48.2 LPAC. 3.- INTELIGENCIA ARTIFICIAL Y PROCEDIMIENTO ADMINISTRATIVO: GARANTÍAS DEL PROCEDIMIENTO ADMINISTRATIVO COMÚN. 3.1.- Encuadre: la pérdida de derechos ante la administración digital. 3.2.- El derecho a subsanar la presentación de escritos. 3.3.- El derecho de audiencia. 3.4.- La motivación de los actos administrativos adoptados mediante sistemas de IA. 3.4.1.- La motivación como requisito de validez de los actos dictados mediante IA. 3.4.2.- Alcance y medio de la motivación: ¿reglas especiales en relación con la IA? 3.4.3.- La omisión de motivación genera indefensión y determina la invalidez de los actos dictados con sistemas de IA. Jurisprudencia nacional, comparada y del TJUE. 3.4.4.- Consecuencia: prohibición de algoritmos de caja negra en la toma de decisiones por el sector público. ¿Posibilidad de articular en el futuro una manera distinta de motivar, cuando se trata de decisiones sustentadas en IA? 3.5.- Inteligencia artificial y discrecionalidad administrativa. 4.- FINAL: LA INACEPTABLE TENSIÓN ENTRE PROGRESO TECNOLÓGICO Y GARANTÍAS DE LA CIUDADANÍA; LUCES Y SOMBRAS DE LA REGULACIÓN ESPECÍFICA DE LA IA EN EL SECTOR PÚBLICO. BIBLIOGRAFÍA CITADA

1.- INTRODUCCIÓN

1.1.- Acotación material: sistemas de IA para soporte a la toma de decisiones

La robotización y la inteligencia artificial (IA) se están incorporando al quehacer ordinario del sector público, como herramientas que agilizan la gestión, contribuyen a mejorar las decisiones y con ello también los servicios públicos, por lo que redundan igualmente en beneficios para la ciudadanía. Son indudables las ventajas que representa la IA para la mejora del sector público y de las actividades

y servicios que presta. No en vano, se estima que la tercera parte del sector público español ya aplicaba soluciones de IA en 2022².

Son muy diversos los usos que puede hacerse de la IA en el sector público: en el Capítulo II se ofrecen nutridos ejemplos. Ahora bien, existen sustanciales diferencias de régimen jurídico según sea el ámbito en el que se implanten estas soluciones. Existe un primer universo de sistemas orientados a la mejora de las prestaciones y servicios públicos sin impacto directo en los derechos e intereses legítimos de los ciudadanos: por ejemplo, la implantación de sistemas de información administrativa (telefónicos o vía chat) automatizados; el análisis de hábitos o necesidades en un determinado ámbito para identificar medidas de mejora de servicio (así, en el sector turístico); el uso de asistentes en cirugía; la regulación y flujo del tráfico o el trazado de las líneas de transporte urbano; la ubicación idónea de centros públicos (educativos, deportivos, sanitarios...); etc. Estos ámbitos de aplicación de IA no quedan al margen del Derecho, pero desde luego su régimen jurídico es mínimo y tiene que ver esencialmente con la protección de datos (los que se requieren para nutrir y entrenar el sistema).

Por el contrario, la situación es bien diferente cuando se trata de implantar sistemas de IA como soporte a la toma de decisiones que afectan a derechos e intereses legítimos de la ciudadanía. Y son ya muchos los campos en los que esto viene sucediendo, utilizándose sistemas de IA, por ejemplo, para elegir el universo de sujetos objeto de inspecciones; para incoar un procedimiento sancionador; para otorgar o denegar subvenciones; para conceder el acceso a servicios públicos, como el ingreso en un centro de educación o una prestación sanitaria; para la gestión del empleo público – reconocimiento de complementos salariales, asignación de tareas o destinos–; etc. En este segundo ámbito la cuestión es más delicada, porque resulta exponente de la actuación administrativa formalizada, y exige el respeto, no solo de los principios de actuación del sector público, sino también de las garantías y derechos de los ciudadanos en la tramitación de los

² Véase el informe promovido por Microsoft y elaborado por EY, *Inteligencia artificial en el sector público. España. Perspectivas europeas para 2020 y años siguientes*. El informe señala, no obstante, que los desarrollos están lejos todavía de extraer el máximo provecho de esta tecnología.

procedimientos administrativos. El presente capítulo se centra en el régimen jurídico de este segundo tipo de sistemas de IA.

1.2.- Las especialidades de la IA en el sector público y su sujeción al Derecho administrativo general sin necesidad de disposiciones que lo establezcan expresamente

No existe aun un marco normativo *específico* que regule el uso de estas tecnologías con el suficiente nivel de desarrollo, esto es: no se ha promulgado una norma de cabecera sobre IA, ni en el sector público ni en el privado. Tan solo se han aprobado, hasta la fecha, algunos preceptos deslavazados y dispersos en diferentes normas. Pero es obvio que su empleo impacta en relevantes derechos y, por lo que concretamente se refiere a su uso por el sector público, se encuentra además sujeto a las reglas que específicamente rigen su actuación. En este sentido, el Joint Research Centre ha destacado la necesidad de considerar reglas y principios específicos aplicables al sector público cuando se implantan soluciones de IA en este ámbito³, y el Parlamento Europeo ha puesto de manifiesto las diferencias que existen en su propuesta de Reglamento entre la aplicación de la IA al sector público y al privado⁴.

En este trabajo pretendemos sistematizar y analizar los principios, reglas y preceptos que vinculan al sector público cuando implementan sistemas de IA, en dos ámbitos normativos específicos: el régimen jurídico básico del sector público y el procedimiento administrativo común. En unos casos –los menos– se trata de disposiciones específicas sobre la robotización o la IA; en otros, se trata de garantías generales, aplicables a cualquier actuación realizada por el sector público.

Parecería ocioso recordar que las Leyes 39/2015, de 1 de octubre, del procedimiento administrativo común de las Administraciones públicas (LPAC) y 40/2015, de la misma fecha, de régimen jurídico del

³ Véase MANZONI M., MEDAGLIA R., TANGI L., VAN NOORDT C., VACCARI L., y GATTWINKEL D. (2022): *AI Watch. Road to the Adoption of Artificial Intelligence by the Public Sector*, JRC-European Commission, Luxembourg.

⁴ En GEORGIEVA, I.; TIMAN, T. y HOEKSTRA, M. (2022): *Regulatory divergences in the draft AI Act. Differences in public and private sector obligations*, European Parliamentary Research Service, Brussels.

sector público (LRJSP) deben respetarse cualquiera que sea el medio o soporte de la actuación administrativa –en el marco de la acotación material que realizamos a continuación–, y, por tanto, que a la hora de cumplir con determinaciones tan elementales como el principio contradictorio –materializado en una secuencia de trámites o requisitos que le dan soporte instrumental, como la subsanación, las alegaciones, la audiencia o la motivación de los actos administrativos–, es indiferente que el procedimiento se gestione exclusivamente por humanos –incluso tecleando en una máquina mecanográfica los sucesivos trámites del procedimiento–, o con procedimientos robotizados, o mediante sistemas de IA.

En efecto, no es necesario que ninguna norma establezca que la automatización o la IA deben respetar los derechos y garantías del régimen jurídico del sector público y del procedimiento administrativo común, como tampoco es necesario que lo hagan ante cualquier otro avance tecnológico o soporte a la tramitación de los procedimientos administrativos. Son normas básicas/comunes precisamente porque establecen garantías unitarias de los ciudadanos ante todas las Administraciones públicas (*vid.*, esencialmente, SSTC 55/2018, de 24 de mayo, FJ 1; y 50/1999, de 6 de abril, FJ 3) y con independencia del concreto medio o soporte que se utilice en cada caso específico.

No podemos interpretar la omisión de regulación en relación con la IA como una patente de corso para prescindir por completo de los principios de actuación del sector público y de los derechos y garantías de los ciudadanos que establecen respectivamente la LRJSP y la LPAC. Dicho en otras palabras: el hecho de que estas normas no aludan una y otra vez a la IA no significa que cuando se utilicen estos sistemas queden al margen del Ordenamiento jurídico. Todo lo contrario. La LRJSP y la LPAC han de respetarse sea cual sea el medio de actuación o soporte que se utilice en la actividad administrativa formalizada. Lo que resulta necesario es integrar la laguna determinando de qué concreta manera deben salvaguardarse los principios de actuación del sector público y las garantías de los ciudadanos cuando se utiliza la IA como soporte a la toma de decisiones que les afectan.

Es indudable que esa labor hermenéutica puede generar dudas, en particular, cuando el avance tecnológico supone un cambio radical de

paradigma y las verdades conocidas se tambalean. El entorno jurídico-público se mueve con cierto confort en los medios formalizados, que van dejando su *traza* en una secuencia de trámites, respaldados por un soporte físico –ya sea en papel o digital–, e inteligible en lenguaje natural por la práctica totalidad de la ciudadanía. El problema se presenta cuando esas evidencias se diluyen en un éter vaporoso, en una solución de IA cuya lógica de funcionamiento escapa al entendimiento común y exige saberes especializados. Y sobre todo, cuando se disuelven los trámites ritualistas y formalizados del procedimiento administrativo, demudándose en una especie de oráculo final –revenir las decisiones automatizadas es uno de los sesgos humanos que nos resalta PONCE SOLÉ más atrás– cuyo mecanismo no solo desconocen los profanos por más información que se les dé, sino también los expertos a quienes se les deniegue la información necesaria para calibrar su funcionamiento. Más aún: es eventualmente imposible saber la lógica que ha seguido el sistema cuando se utilizan determinadas tecnologías de IA (los sistemas de aprendizaje automático, redes neuronales, etc.).

Estas dificultades de interpretación generan inseguridad jurídica. Y por ello es deseable que se adopten disposiciones específicas que regulen de qué concreta manera se debe satisfacer el Derecho básico/común cuando se aplica la IA como soporte a la toma de decisiones. No porque en ausencia de un régimen específico esté al margen, sino para aclarar de qué manera lo está. En omisión de esos preceptos, nos corresponde a los operadores jurídicos determinar el marco normativo, el Derecho a aplicar, la manera de salvaguardar los principios de actuación del sector público y las garantías de los ciudadanos cuando se utilizan soluciones de IA como soporte a la toma de decisiones. Ese es el propósito que alienta este capítulo.

2.- ACTIVIDAD ADMINISTRATIVA AUTOMATIZADA: PREVISIONES EN LA LEY 40/2015, DE RÉGIMEN JURÍDICO DEL SECTOR PÚBLICO⁵

-
- ⁵ Sobre la actuación administrativa automatizada, véanse especialmente los trabajos de CERRILLO MARTÍNEZ, A. (2021), “Robots, asistentes virtuales y automatización de las administraciones públicas”, en *Revista Galega de Administración Pública*, núm. 61, págs. 271 ss. (se puede consultar en <https://egap.xunta.gal/revistas/REGAP/article/view/4791>., última visita 28 de marzo de 2023); CERRILLO MARTÍNEZ, A., “Automatización e inteligencia artificial”, en I. MARTÍN DELGADO (Dir.), *El procedimiento administrativo desde la perspectiva de la innovación tecnológica*, Iustel-IVAP, Madrid, 2020, págs. 389 y ss.; CERRILLO MARTÍNEZ, A., “Actividad administrativa automatizada y utilización de algoritmos”, en F. CASTILLO BLANCO, S. FERNÁNDEZ RAMOS Y J. M. PÉREZ MONGUIÓ (Dirs.), *Las políticas de buen Gobierno en Andalucía (I): Digitalización y transparencia*, Instituto Andaluz de Administración Pública, Sevilla, 2022, págs. 259 y ss. (se puede consultar en <https://www.juntadeandalucia.es/institutodeadministracionpublica/publico/libros/politicassBuenGobierno/> última visita 28 de marzo de 2023); MARTÍN DELGADO, I. (2009), “Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada”, en *Revista de Administración Pública*, núm. 180, págs. 353 y ss. (se puede consultar en <https://dialnet.unirioja.es/descarga/articulo/3209548.pdf>, última visita 28 de marzo de 2023); URIOS APARISI, X. y ALAMILLO DOMINGO, I. (2011): *La actuación administrativa automatizada en el ámbito de las Administraciones públicas. Análisis jurídico y metodológico para la construcción y explotación de trámites automáticos*, Escola d’Administració Pública de Catalunya, Barcelona (se puede consultar en https://eapc.gencat.cat/web/.content/home/publicacions/col_leccio_estudis_de_recerca_digital/3_actuacio_administrativa_automatitzada/alamillo_ursos_castellano.pdf, última visita 28 de marzo de 2023); BAUZÁ MARTORELL, FJ (2017): “Identificación, autenticación y actuación automatizada de las Administraciones públicas”, en E. GAMERO CASADO (Dir.), *Tratado de procedimiento administrativo común y régimen jurídico básico del sector público*, vol. 1, Tirant lo Blanch, Valencia, págs. 787 y ss.; GÓMEZ PADILLA, R. (2016): “Comentario al art. 41 LRJSP”, en RECUERDA GIRELA, M.A. (Dir.), *Régimen jurídico del sector público y procedimiento administrativo común*, Aranzadi-Thomson-Reuters, Cizur Menor, págs.1.222 y ss.; TERRÓN SANTOS, D. y DOMÍNGUEZ ÁLVAREZ, J.L. (2022): *i-Administración pública, sistemas algorítmicos y protección de datos*, Iustel, Madrid, págs. 107 ss.; TARRES VIVES, M. (2012): “La actuación administrativa automatizada”, en J. TORNOS MAS, *La Ley 26/2010, del 3 de agosto, de régimen jurídico y procedimiento de las Administraciones públicas de Cataluña*, Iustel, Madrid; y ALAMILLO DOMINGO, I. (2017): Comentarios a los artículos 41 y 42, en CAMPOS ACUÑA, C. (Dir.), *Comentarios a la Ley*

No toda actuación automatizada utiliza la IA. Pero todo desarrollo de IA sí que se sustenta de uno u otro modo en la actuación automatizada. Así pues, las previsiones que establece nuestro ordenamiento jurídico en relación con la actuación administrativa automatizada son de aplicación a las soluciones de IA.

El régimen jurídico de la actuación administrativa automatizada se recoge esencialmente en los arts.41 y 42 LRJSP, tributarios de lo que en su día estableció la Ley 11/2007, de 22 de junio, y ésta, a su vez, de la Ley 58/2003, de 17 de diciembre, General Tributaria; aunque entonces resultase pionero, y aun siéndolo hoy desde la perspectiva del Derecho comparado en muchos aspectos, lo cierto es que el grado de madurez que va adquiriendo la tecnología y los problemas jurídicos que evidencian su aplicación revelan que estos preceptos son ya insuficientes⁶. A continuación analizamos sus elementos principales.

2.1.- El concepto de “actuación administrativa automatizada” y los grados de automatización

El concepto de “actuación administrativa automatizada” previsto en el art. 41.1 LRJSP está tomado del Anexo de la Ley 11/2007, que contenía definiciones, entre las que se encontraba ésta. La única diferencia es que la Ley de 2007 se refería a personas naturales, mientras que la Ley 40/2015 se refiere a empleados públicos. Resaltemos algunos aspectos.

40/2015, de Régimen Jurídico del Sector Público, Wolters Kluwer, Madrid, págs. 222 y ss.

Sobre aspectos particulares de la actuación automatizada, véanse SORIANO ARNANZ, A. (2021): “Decisiones automatizadas y discriminación: aproximación y propuestas generales”, en *Revista General de Derecho Administrativo*, núm. 56, (también en línea, INAP en <https://laadministracionaldia.inap.es/noticia.asp?id=1511706>, última visita 28 de marzo de 2023); de la misma autora (2021), “Decisiones automatizadas: problemas y soluciones jurídicas. Más allá de la protección de datos”, en *Revista de Derecho Público: Teoría y Método*, vol. 3, págs. 85 y ss. (se puede consultar en <http://www.revistasmarcialpons.es/revistaderchopublico/article/view/535>, última visita 28 de marzo de 2023); y CAPDEFERRO VILLAGRASA, O. (2023): “Las obligaciones sujetas a control administrativo automatizado”, *Revista de Internet, Derecho y Política*, n.º 37.

⁶ Así lo entiende CERRILLO MARTÍNEZ, A. (2022), “Actividad administrativa automatizada...”, *op.cit.*, pág.261.

En general, aunque pueda parecer vaga, la definición se considera bastante apropiada⁷. En particular, el legislador es consciente de que es imposible adoptar una actuación administrativa absolutamente desprovista de intervención humana, al menos en el diseño y gestión del sistema de información. Por lo tanto, basta con que dicha intervención humana sea indirecta.

En segundo lugar, el concepto también está delimitado por el contexto en el que se aplica el sistema automatizado de actuación administrativa: ha de tratarse necesariamente de actuación realizada “en el marco de un procedimiento administrativo”. Por lo tanto, este régimen jurídico no se aplica a los sistemas automatizados, o sistemas de IA, utilizados para la mejora de la prestación material de los servicios públicos. Solo se refiere a la actividad administrativa formalizada, que es precisamente la que hemos delimitado como objeto de nuestra atención⁸.

Es importante destacar que esta definición no solo se refiere a las decisiones automatizadas (actos administrativos definitivos) en sentido estricto, sino también a los algoritmos, sistemas de IA o actuaciones automatizadas que apoyan la toma de decisiones, es decir, que se utilizan en el marco del procedimiento administrativo, ya sea como acto de trámite, ya como un *input* a tener en consideración para la decisión final en el proceso decisorio, es decir, a modo de *informe* o de actuación que se inserta en el procedimiento para aportar elementos valiosos que han de tomarse en consideración al tomar la decisión final, esto es, el acto administrativo definitivo –*resolutorio*– del procedimiento. Por tanto, cuando, con ocasión de la tramitación de un procedimiento administrativo, se utilice un sistema de IA para obtener información que sirva de base o apoyo para dictar la resolución definitiva del procedimiento, ese sistema de actuación automatizada

⁷ Véase, en particular, MARTÍN DELGADO, I. (2009): “Naturaleza, concepto...”, *op. cit.*, págs. 367 y ss.

⁸ Se ha reclamado la necesidad de regular también la actuación automatizada en la actividad administrativa prestacional, así, VALERO TORRIJOS, J. (2019): “Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración”, *Revista catalana de dret públic*, n.º 58, pág.85 (<https://doi.org/10.2436/rcdp.i58.2019.3307> , última visita 28/3/2023).

también ha de reunir y cumplir las exigencias legales. Esta aclaración es muy importante, porque amplía considerablemente su ámbito de aplicación, impidiendo que la Administración limite la aplicación de esta disposición y la exigencia de sus requisitos legales solo a los casos en que la decisión final se adopte de forma automatizada.

Por último, debe notarse que este concepto de actividad automatizada permite acoger los tres grados o niveles de automatización que comúnmente se definen en el marco de los procedimientos administrativos: 1) Para iniciar el procedimiento: por ejemplo, cuando un procedimiento sancionador se incoa como consecuencia de la información resultante de un sistema de IA que reúne evidencias de la comisión de una infracción, procediéndose a emitir de manera automatizada un acuerdo de iniciación y a notificarlo al interesado; 2) Como herramienta que se utiliza en un procedimiento a fin de reunir información útil para resolver, pero que exige posteriormente la intervención humana a fin de adoptar la decisión o acto administrativo definitivo del procedimiento; y 3) Actuación automáticamente automatizada, que permite adoptar actos administrativos sin intervención humana alguna.

Esta clasificación es interesante, puesto que, en función de los rasgos y circunstancias confluientes en un procedimiento, cabrá implantar un grado mayor o menor de automatización. Por ejemplo, en un procedimiento de otorgamiento de subvenciones, sería lícito implantar un sistema automatizado que verifique el cumplimiento de todas las exigencias formales y sustantivas por parte del solicitante, y proceda a emitir el acto de otorgamiento de manera automatizada. En cambio, como veremos en el siguiente bloque de cuestiones de este Capítulo, si en el procedimiento es preciso articular un trámite de subsanación (porque no se disponga de toda la información necesaria para resolver) o un trámite de alegaciones o de audiencia, es mucho más discutible que se pueda tramitar la totalidad del procedimiento de manera automatizada.

También debe notarse que en un mismo procedimiento pueden combinarse diferentes grados de automatización. Por ejemplo, en procedimientos susceptibles de producir actos favorables o desfavorables según las circunstancias: así, cuando la decisión automatizada es favorable puede producirse sin más, pero si es desfavorable requiere

previamente una revisión humana antes de adoptarse la decisión definitiva⁹.

2.2.- La aprobación de la implantación del sistema

El sistema automatizado de actuación administrativa debe ser establecido expresamente, lo que constituye un requisito sumamente importante, ya que impide la implementación de sistemas *de facto*¹⁰.

El instrumento jurídico mediante el cual acordar la implantación es muy variado: el sistema automatizado de actuación administrativa puede ser establecido mediante un acto administrativo, mediante un reglamento e incluso en una norma con rango de ley. Así, el artículo 13.2 del Real Decreto 203/2021, de 30 de marzo, por el que se aprueba el Reglamento de actuación y funcionamiento del sector público

⁹ Véase, en este sentido, la exposición que hace Oriol MIR PUIGPELAT del Sistema Europeo de Información y Autorización de Viajes (SEIAV, más conocido por sus siglas en inglés ETIAS) en el Capítulo XVI de esta obra: cuando el sistema concede automáticamente la autorización, no se somete a revisión; pero cuando la respuesta automatizada es denegatoria, se obliga a su revisión por un humano, que no está vinculado por la que adoptó el sistema.

¹⁰ A decir de TERRÓN SANTOS, D. y DOMÍNGUEZ ÁLVAREZ, J.L. (2022): *i-Administración pública...*, *op.cit.*, págs. 111 ss., la implantación de un sistema de actuación automatizada supone, *de facto*, la atribución de la *competencia* para adoptar tales decisiones, por lo que debe acordarse expresamente. Consideran que si el sistema de IA sustituye a la decisión humana (no siendo un mero soporte o apoyo a la posterior toma de decisión por una persona, sino reemplazando completamente la intervención humana), supondrían una alteración o traslación competencial y deben aprobarse como normas atributivas de competencia (*ibidem*, págs. 119 ss.). El art.33.1 de la Ley 11/2007, de 22 de junio, recordaba claramente que “La gestión electrónica de la actividad administrativa respetará la titularidad y el ejercicio de la competencia por la Administración Pública, órgano o entidad que la tenga atribuida”; no existe equivalente a este precepto en las Leyes siamesas, 39 y 40/2015, de 1 de octubre. A mi juicio, en la medida que sea el propio órgano competente para adoptar la decisión quien implante el sistema no sería necesaria norma alguna. Sí haría falta en el caso de que el órgano a quien se impute la actuación sea distinto del que resultase de los reglamentos orgánicos vigentes. Por otra parte, como se explica en el texto a continuación, ocasionalmente los sistemas de IA son implantados mediante normas legales o reglamentarias, por lo que en tales casos se cumpliría la exigencia de atribuirla mediante disposiciones normativas, siempre que fuesen de igual o superior rango que la norma organizativa preexistente.

por medios electrónicos (RAFESP), establece que “En el ámbito estatal la determinación de una actuación administrativa como automatizada se autorizará por resolución del titular del órgano administrativo competente”, es decir, mediante un acto administrativo¹¹. Pero también se han implementado actuaciones administrativas automatizadas a través de normas con rango de ley: por ejemplo, el Real Decreto-ley 2/2021, de 26 de enero, de refuerzo y consolidación de medidas sociales en defensa del empleo, ha regulado la iniciación de procedimientos administrativos en el ámbito de la Seguridad Social mediante actas de infracción automatizadas.

La exigencia de aprobación expresa de implantación del sistema es muy adecuada. Pero debe completarse con otras garantías, especialmente en relación con los derechos de la ciudadanía. Así, sin entrar en la cuestión de que un algoritmo –o un sistema de actuación automatizada o de IA– sea o no un reglamento¹², este acto o acuerdo de implantación del sistema debe publicarse en el diario oficial correspondiente, para general conocimiento. Y correlativamente, debería ser objeto de medidas de transparencia, que permitan verificar las exigencias legales que pesan sobre el diseño y funcionamiento de estos sistemas¹³.

Debe destacarse, en este sentido, lo dispuesto por el art.11.1 i) RAFESP, que obliga a publicar en cada sede electrónica la “Relación actualizada de las actuaciones administrativas automatizadas vinculadas a los servicios, procedimientos y trámites descritos en la letra anterior. Cada una se acompañará de la descripción de su diseño y funcionamiento, los mecanismos de rendición de cuentas y transparencia, así como los datos utilizados en su configuración y aprendizaje”. El precepto tiene embargo carácter básico. No supera todas

¹¹ Esta norma solo se aplica a la Administración General del Estado, no a las comunidades autónomas ni a los municipios. Por ofrecer algún ejemplo autonómico de implantación de sistemas de actuación administrativa automatizada, puede verse la Resolución de 24 de noviembre de 2022, de la Agencia Andaluza de la Energía, por la que se aprueba la automatización de actos en el marco de las convocatorias de incentivos que se citan (BOJA de 29 de noviembre de 2022).

¹² Problemática que, por otra parte, se analiza por Antonio BERNING PRIETO en el Capítulo III de esta obra.

¹³ Sobre la cuestión remito a lo que expone Isaac MARTÍN DELGADO en el Capítulo IV.

las exigencias resaltadas en el Capítulo IV de esta obra, pero supone un avance respecto de la situación anterior. Ahora bien, una cata por diversas sedes electrónicas en búsqueda de esta información ha resultado infructuosa, aun cuando he accedido a órganos que aplican sistemas de IA. Cabe pensar que la aplicación del precepto se ha entendido circunscrita a los sistemas de actuación automatizada que actúen precisamente en la sede, no incluyendo al total de los que opera el órgano en cuestión (al margen de la sede). De ser esta la interpretación que se está haciendo del art.11.1 i) RAFESP, sus consecuencias prácticas serán verdaderamente pobres, y además, se producirá una clara e incomprensible diferenciación de regímenes jurídicos entre las actuaciones automatizadas que se realicen en sede y las que no.

2.3.- El órgano competente para implementar los sistemas de IA y atribuirle su actividad. Firma electrónica de decisiones automatizadas

El art. 41.2 LRJSP establece que el órgano administrativo responsable del sistema, y al que se imputa la actuación a efectos de recurso, debe determinarse con carácter previo a su puesta en funcionamiento. Se trata de un requisito de la máxima relevancia.

En concreto, el ya mencionado art. 13.2 RAFESP –aplicable solo a la Administración estatal–, dispone que la resolución por la que se establezcan los ámbitos de actuación administrativa automatizada “expresará los recursos que procedan contra la actuación, el órgano administrativo o judicial, en su caso, ante el que hubieran de presentarse y plazo para interponerlos”. Como se observa, se trata de la garantía conocida como *pie de recurso*, en cuya virtud –art.88.3 LPAC–, esta misma información debe contenerse en todas las resoluciones que pongan fin al procedimiento y a los actos administrativos. Se evidencia así la asimilación que la normativa lleva a cabo entre actuación automatizada y no automatizada, sometiénolas a garantías homogéneas.

Además, hay que destacar que estas disposiciones descartan la posibilidad de que la actuación administrativa automatizada pueda considerarse auténticamente “autónoma”, ya que siempre estará atribuida a un órgano administrativo concreto.

Enlaza con todo ello lo dispuesto en el art.42 LRJSP, en cuya virtud se podrán implantar los siguientes sistemas de firma electrónica en el ejercicio de la competencia en la actuación administrativa automatizada¹⁴:

“a) Sello electrónico de Administración Pública, órgano, organismo público o entidad de derecho público, basado en certificado electrónico reconocido o cualificado que reúna los requisitos exigidos por la legislación de firma electrónica.

b) Código seguro de verificación vinculado a la Administración Pública, órgano, organismo público o entidad de Derecho Público, en los términos y condiciones establecidos, permitiéndose en todo caso la comprobación de la integridad del documento mediante el acceso a la sede electrónica correspondiente”.

La principal consecuencia de esta norma es que las decisiones adoptadas mediante actuación administrativa automatizada incorporarán un sistema de firma electrónica, lo que abunda en la asimilación entre la actuación administrativa automatizada y la tradicional. Como apunta ALAMILLO, el art.8 LRJSP determina que siempre se deba determinar a qué concreto órgano se atribuye el ejercicio de la competencia mediante actuación administrativa automatizada, sin que pueda utilizarse en este caso el sello genérico de la Administración actuante¹⁵. Ello permite identificar de manera clara e inequívoca el órgano al que se atribuye la actuación, lo que es asimismo relevante a efectos de recurso.

¹⁴ Entiende CERRILLO MARTÍNEZ, A (2021).: “Actividad administrativa automatizada...”, *op.cit.*, pág.209, que el precepto admite tanto la implantación de estos dos tipos de firma, como también que una Administración decida no exigir firma para la actuación automatizada. En mi opinión, la discrecionalidad que otorga el precepto es la de elegir necesariamente entre uno de los dos sistemas de firma, pero no admite que una actuación automatizada carezca de firma. Tengamos presente que la actuación automatizada regulada por la LRJSPO es la formalizada, y por tanto, que se trata esencialmente de dictar actos administrativos, aunque sean de trámite. En tales circunstancias, a mi juicio, el art.26.3 LPAC, en conexión con el art.26.2 e) de la misma Ley, determinan que estas actuaciones integrantes de un procedimiento administrativo, y por tanto, también del expediente, deban firmarse en todo caso.

¹⁵ Véase ALAMILLO DOMINGO, I. (2017): Comentarios a los artículos..., *op.cit.*, pág. 224.

2.4.- La supervisión, control de calidad y auditoría del sistema y de su código fuente

2.4.1.- Encuadre general. El deber legal de implantar estos controles de la IA

El art.41.2 LRJSP dispone, como hemos comprobado, que, al implantar una actuación administrativa automatizada, “deberá establecerse previamente el órgano u órganos competentes, según los casos, para la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente”. El precepto comporta unas consecuencias más intensas de lo que parece a primera vista.

En efecto, el artículo no establece *directamente* que deba existir una supervisión, control de calidad y, en su caso, auditoría del sistema, pero sí lo impone *indirectamente*, pues en la medida que se debe pre-determinar el órgano competente para definir todas estas cuestiones, también se está estableciendo que tales previsiones deben producirse con ocasión de la implantación de actuación administrativa automatizada, si se toma en consideración la finalidad de la norma. Se trata, por tanto, de un deber legal implícito, del que existen otros ejemplos en nuestro Ordenamiento jurídico, incluso en materia específica de administración electrónica, como es el deber de cada Administración pública de crear un Punto de Acceso General electrónico¹⁶.

Sentado esto, resulta preciso determinar el alcance del precepto, lo que no es sencillo, por dos razones. La primera, porque no queda claro el sentido que debe darse a la locución “en su caso”, que luce

¹⁶ En efecto, el art. 13.a) LPAC establece que los ciudadanos tienen derecho “A comunicarse con las Administraciones Públicas a través de un Punto de Acceso General electrónico de la Administración”. Ninguna norma regula este Punto de Acceso General, ni en la propia Ley 39/2015, ni tampoco en la 40/2015; pero se ha interpretado como un punto que integra todos los servicios y procedimientos de cada Administración pública, facilitando así las relaciones de los ciudadanos por medios electrónicos, y cuya creación, por consiguiente, es preceptiva para todas las Administraciones públicas. El Tribunal Constitucional, en su sentencia 55/2018, de 24 de mayo, considera que el legislador nacional puede obligar a todas las Administraciones públicas a disponer de dicho Punto de Acceso General electrónico.

antes de la mención a la auditoría del sistema. Dos preguntas se nos plantean: *i)* ¿Qué función desempeña esta expresión en la frase? Si nos atenemos a su naturaleza gramatical, se trata de una locución conjuntiva condicional, es decir, que supedita el efecto de la acción expresada en la frase al acaecimiento de una determinada circunstancia. El problema reside en que no se nos ofrece pista alguna acerca de cual sea la circunstancia en cuestión, y, por tanto, no parece haber manera de concretar cuándo se da el caso al que alude el precepto, configurándola, aparentemente, como una suerte de condición fantasma; *ii)* ¿Cuál es su proyección al resto de elementos de la frase? Se describen varias acciones –definición de las especificaciones, programación, mantenimiento, supervisión, control de calidad y auditoría del sistema de información y de su código fuente–, sin que del tenor literal de las palabras pueda determinarse si todas ellas quedan supeditadas a la condición fantasma, o tan solo las que figuran tras la locución conjuntiva condicional, esto es, la auditoría del sistema de información y de su código fuente. A mi juicio, y como respuesta a ambas preguntas *i)* y *ii)*, lo que ha pretendido el precepto es describir o relacionar el conjunto de medidas de supervisión y control que podrían existir sobre un sistema de actuación administrativa automatizada, compilándolos todos, a fin de *que se establezcan de manera efectiva tantos como sea pertinente aplicar en función de las características del sistema*. El condicional –*en su caso*– no aludiría entonces a una concreta circunstancia o requisito –puesto que no se indica cuál es en ninguna parte–, y, por tanto, nos gustaría decir que *todos y cada uno de estos controles deben imponerse sobre el sistema siempre que sea posible en función de las características del mismo*.

Aclarado este extremo nos queda desarrollar la segunda cuestión que no es fácil de exponer: el contenido material de todas y cada una de las obligaciones que se imponen. La cuestión no merece mayor dificultad –ni atención– por lo que se refiere a las especificaciones, programación y mantenimiento, por cuanto que se trata de conceptos jurídicos indeterminados suficientemente expresivos en las ciencias remitidas. Ahora bien, ¿qué debemos entender por supervisión, control de calidad y auditoría del sistema de información y de su código fuente? Vayamos por partes.

2.4.2.- Control de calidad y auditoría del sistema de información y de su código fuente

En efecto, la IA es una tecnología relativamente reciente y en rapidísima expansión, lo que no ha permitido desplegar todavía herramientas adecuadas para verificar todos estos extremos, llegándose a afirmar que los criterios para desplegar estos controles, y especialmente la auditoría, aún no están claros¹⁷. Ahora bien, como veremos en este apartado, existen multitud de iniciativas así como estudios comparados entre ellas¹⁸. La variedad de enfoques, tanto en el ámbito sobre el que debe desplegarse el control como en la metodología aplicable para realizarlo, complican la sistematización del material existente.

Hay un ámbito concreto en el que se han desarrollado algunos métodos: la protección de datos personales. Por lo que respecta a la legislación española, hay que citar la aportación de MARTÍNEZ MARTÍNEZ¹⁹ y los documentos elaborados por la Agencia Española de Protección de Datos: en particular, las directrices “Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción” (2020), y “Requisitos para auditorías de tratamientos

¹⁷ Véase BONE, J. (2020): “Auditoría de Inteligencia Artificial”, en *International Dynamic Advisors*, 31 de julio de 2020, (se puede consultar en <https://www.intedya.com/internacional/2764/noticia-auditoria-de-inteligencia-artificial.html>, última visita 28 de marzo de 2023); BERGHOFF, C. y otros (2021): *Towards Autitable AI Systems. Current status and future directions*, Whitepaper, Federal Office for Information Security-Fraunhofer Forum, pág.2.

¹⁸ Estudios comparados de diferentes metodologías desplegadas para el control de la IA pueden verse en JOSHI, D. y otros (2021): *Algorithmic accountability for the public sector. Learning from the first wave of policy implementation*, Ada Lovelace Institute, AI Now Institute and Open Government Partnership, accessible en <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/> (última visita: 28/03/2023); EZEANI, G.; KOENE, A.; KUMAR, R.; SANTIAGO, N.; y WRIGHT, D. (2021): *A survey of artificial intelligence risk assessment methodologies. The global state of play and leading practices identified*, Ernst&Young-Trilateral Research; y BERGHOFF, C. y otros (2021): *Towards Autitable AI Systems...*, *op.ult.cit.*

¹⁹ Véase MARTÍNEZ MARTÍNEZ, R. (2019): “inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo”, *Revista Catalana de Derecho Público*, n.º 58, págs. 64 ss.

que incluyan inteligencia artificial” (2021)²⁰. En el Reino Unido se ha elaborado también material al respecto²¹.

Con un enfoque más amplio –que suele incluir la protección de datos pero abarcar también otros extremos– se han elaborado otras herramientas²² e incluso viene hablándose de la necesidad de un análisis de impacto (*impact assessment*) previo a la implantación de soluciones de IA en el sector público: es el caso de la primaria aportación de REISMAN y otros²³, y de los ya más prolijos trabajos de MOSS y otros (2020)²⁴, de KAZIM y otros (2021)²⁵ o de MANTELERO²⁶.

²⁰ Los documentos citados se encuentran en <https://www.aepd.es/media/guias/adecuacion-rgpd-ia.pdf> y <https://www.aepd.es/media/guias/requisitos-auditorias-tratamientos-incluyan-ia.pdf>, respectivamente (última visita 29 de marzo de 2023). Sobre protección de datos y algoritmos e IA, véase lo que expone VALE-RO en el Capítulo VIII de esta obra.

²¹ El Information Commissioner ha creado una sección de su sitio web en la que facilita diferentes herramientas para la implantación de la IA conforme al RGPD; la documentación se encuentra accesible en <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/> (última visita: 29/03/2023).

²² Por ejemplo, LESLIE, D. (2019): *Understanding artificial intelligence ethics and safety. A guide for the responsible design and implementation of AI*, The Alan Turing Institute (<https://doi.org/10.5281/zenodo.3240529>, última visita 29 de marzo de 2023). Véase también el Informe del Grupo Independiente de Expertos de Alto Nivel Sobre Inteligencia Artificial de la Comisión Europea *Directrices éticas para una IA fiable*, 2019; y el amplio comentario del documento que realiza COTINO HUESO, L. (2022): “Transparencia y explicabilidad de la inteligencia artificial y *compañía* (comunicación, interpretabilidad, inteligibilidad, auditabilidad, testabilidad, comprobabilidad, simulabilidad...). Para qué, para quién y cuánta”, en L. Cotino Hueso y J. Castellanos Claramunt, *Transparencia y explicabilidad de la inteligencia artificial*, Tirant lo Blanch, Valencia, y especialmente, en relación con la auditabilidad, págs.58 ss.

²³ REISMAN, D.; SCHULTZ, J.; CRAWFORD, K.; y WHITAKER, M. (2018): *Algorithmic Impact Assessments: a practical framework for public agency accountability*, AINow Institute (University of New York).

²⁴ MOSS, E.A. WATKINS, J. METCALF y ELISH, M.C. (2020): “Governing with Algorithmic Impact Assessments: Six Observations”, 24/4/2020, <https://dx.doi.org/10.2139/ssrn.3584818>, última visita 28 de marzo de 2023.

²⁵ En este caso, relativo al Reino Unido; KAZIM, E., MENDES, D., DENNY, T. y KOSHIYAMA, A. (2021): “AI auditing and impact assessment: according to the UK information commissioner’s office”, *AI Ethics*, <https://doi.org/10.1007/s43681-021-00039-2>, última visita 28 de marzo de 2023.

²⁶ MANTELERO, A. (2022): *Beyond Data. Human Rights, Ethical and Social Impact Assessment in Artificial Intelligence*, Asser Press–Springer, The Hague.

También deben destacarse las reglas-modelo elaboradas por un prestigioso panel de juristas en el Instituto de Derecho Europeo (ELI): se trata del informe (2022) *Model Rules on Impact Assessment of Algorithmic Decision-Making System Used by Public Administration*, un documento excelente, y diseñado de manera específica para el sector público, que puede aplicarse en cualquier Ordenamiento jurídico²⁷. El ELI también nos ofrece los *Guiding principles* elaborados por DE LAS HERAS en el seno de esta organización²⁸.

Asimismo, han surgido iniciativas que articulan un análisis de impacto para verificar la conformidad de los sistemas de IA con el reglamento europeo²⁹.

Además de otras iniciativas que podrían mencionarse³⁰, el Gobierno de Canadá ya ha implantado una herramienta (la *Algorithmic Impact Assessment Tool*³¹), en el marco de su Directiva de Decisiones Automatizadas³². El Reino Unido ha publicado una *Guía para el uso*

²⁷ El equipo de trabajo, coordinado por M. WIERZBOWSKI (Universidad de Varsovia), estuvo integrado principalmente por M. CLÉMENT (magistrado francés), P. CRAIG (Universidad de Oxford) y J.P. SCHNEIDER (Universidad de Friburgo), contando, además, con otros colaboradores. Accesible en https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Model_Rules_on_Impact_Assessment_of_ADMSs_Used_by_Public_Administration.pdf; última visita 28 de marzo de 2023.

²⁸ DE LAS HERAS BALLELL, T. (2022): *Guiding Principles for Automated Decision-Making in the EU*, European Law Institute, Vienna.

²⁹ Así, FLORIDI, L. y otros (2022): *capAI. A procedure for conducting conformity assessment of AI systems in line with the EU Artificial Intelligence Act*, University of Oxford, Oxford. El documento anuncia su actualización periódica.

³⁰ Como las que articulan un mecanismo de certificación de conformidad, véase por ejemplo GALÁN, C. (2019): “The certification as a mechanism for control of artificial intelligence in Europe”, accesible en <https://ec.europa.eu/futurium/en/european-ai-alliance/certification-mechanism-control-artificial-intelligence-europe-0.html>; o la aportación de MERCHÁN ARRIBAS, M. (2020): *Guía de uso de la Inteligencia Artificial en el Sector Público*, accesible en <https://www.clubdeinnovacion.es/guia-uso-la-ia-sector-publico/> (última visita: 28/03/2023).

³¹ Accesible en <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html>, última visita 28 de marzo de 2023.

³² De la que conviene consultar muy especialmente su Anexo C, a efectos de nuestro objeto de interés. Puede consultarse en <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592> última visita 28 de marzo de 2023.

de la inteligencia artificial en el sector público³³. Malta ha implantado un sistema voluntario de certificación de IA³⁴; entre nosotros, algo parecido viene reclamando Alexander ZLOTNIK³⁵. Y el NIST ha elaborado modelos para identificar y gestionar los sesgos en la IA³⁶.

2.4.3.- Supervisión humana. La necesidad de *compliance* jurídico

Por lo que respecta a la alusión a la supervisión, a mi juicio, apela a la exigencia de supervisión *humana*, ampliamente demandada tanto en el Derecho español como en el Derecho comparado³⁷.

En mi opinión, el problema no es solo que deba haber una supervisión humana. Además, esa supervisión debe ser multidisciplinar, debe estar formada por especialistas de diferentes campos. En particular, debe incluir la supervisión *jurídica*. Este es un punto en el que he insistido especialmente³⁸: dada la existencia de un marco jurídico que respetar cuando se acometen desarrollos de IA, es necesario que, en

³³ Es el documento *A guide to using artificial intelligence in the public sector*, Office for Artificial Intelligence-Government Digital Service, London, 2020.

³⁴ Si bien centrado en protección de datos. Véase <https://malta.ai/>

³⁵ En ZLOTNIK, A. (2019): “Inteligencia Artificial en las Administraciones Públicas: definiciones, evaluación de viabilidad de proyectos y áreas de aplicación”, *Boletín*, n.º 84, 2019, pp.24 ss, https://www.astic.es/sites/default/files/articulos-boletic/b84_monografico4_alexander_zlotnik_v2-2.pdf, última visita 28 de marzo de 2023.

³⁶ Se trata de SCHWARTZ, R. y otros (2022): *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence*, National Institute of Standards and Technology (United States Department of Commerce), accessible en <https://doi.org/10.6028/NIST.SP.1270> (última visita: 28/03/2023).

³⁷ Véase el Capítulo V de esta misma obra, del que es autor Juli PONCE SOLÉ, que versa, precisamente, sobre “Seres humanos e inteligencia artificial: discrecionalidad artificial, reserva de humanidad y supervisión humana”.

³⁸ Véase GAMERO CASADO, E. (2021): “*Compliance* (o cumplimiento normativo) de desarrollos de Inteligencia Artificial para la toma de decisiones administrativas”, en *Diario La Ley* (Wolters Kluwer), núm. 50, Sección Ciberderecho, 19/04/2021 (se puede consultar en https://diariolaley.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAEAMtMSbF1CTEAAmMjUyMzC7Wy1KLizPw8WyMDI0MDE0NTtbz8lNQqF2fb0ryU1LTMvNQKJLMtEqX_OSQyoJU27TEEnOJUtdSk_PxsFJPiYSYAAPsOzuljAAAAWKE, última visita 28 de marzo de 2023).

el equipo que lleve a cabo la supervisión, haya expertos jurídicos que garanticen el cumplimiento de la legislación. Téngase presente que, exista o no un régimen jurídico específico de la IA, hay derechos y garantías comunes de la ciudadanía sea cual sea el medio o herramienta que se utilice para la toma de decisiones, y que debido, precisamente, a la complejidad y la especificidad que presenta la IA, es imprescindible que haya juristas conocedores de la materia supervisando los sistemas desde su diseño, prueba y puesta en servicio, y periódicamente hasta el fin de su ciclo de vida.

En la mayoría de los casos, el sector público carece de recursos para llevar a cabo desarrollos de IA por sus propios medios y tiene que recurrir a la contratación de estas soluciones en el mercado. Sin embargo, un examen de las licitaciones públicas para sistemas de IA en España muestra que no prevén la inclusión de expertos jurídicos en los equipos de trabajo. Hay ingenieros, matemáticos, informáticos... incluso expertos en comunicación (para campañas publicitarias para dar a conocer la implantación del sistema), pero los juristas brillan por su ausencia.

Considero que esto constituye un grave problema, porque todos estos especialistas son necesarios, y desarrollarán su trabajo magníficamente bien, pero sin tener a la vista los –intensos– condicionamientos jurídicos que rodean la implantación de estas soluciones, por lo que cabe pensar que no serán suficientemente tenidos en consideración.

Esto puede provocar que acaben presentándose graves conflictos sobre la infracción de los requerimientos jurídicos que rodean la implantación de estas soluciones. Y que acaben judicializándose, desembocando en la anulación de las decisiones adoptadas mediante estos sistemas, así como el sistema mismo, impidiendo su continuidad. Ya ha ocurrido así en otros países, como Holanda³⁹,

³⁹ El Tribunal de Distrito de La Haya pronunció una sentencia el 5 de febrero de 2020 (ECLI:NL:RBDHA:2020:865), mediante la que anuló una elaboración de perfiles para la lucha contra el fraude implantada por el Gobierno holandés que únicamente se utilizó en los calificados como “vecindarios problemáticos”, concluyendo el Tribunal que ese criterio puede provocar exclusión, discriminación y estigmatización injustificadas. Desde 2013, 26.000 familias fueron injustamente acusadas de fraude y obligadas a devolver las subvenciones que recibieron. El

Italia⁴⁰ o Austria⁴¹, así como en el TJUE⁴² –por mencionar algunos ejemplos– en los que los tribunales han anulado decisiones del sector público adoptadas mediante estos sistemas: en unos casos, por vulneración del derecho a la igualdad, al incurrir en sesgos de alcance discriminatorio; en otros, por vulnerar garantías esenciales del derecho a la defensa, como la necesidad de conocer las razones o motivos por los que estos sistemas de información adoptaron las decisiones.

Esto representa un grave riesgo reputacional para las Administraciones públicas que promuevan ámbitos de robotización o IA, así como acaba por arruinar la cuantiosa inversión económica afrontada para su puesta en marcha.

Centrándonos en la manera de articular la supervisión, auditoría y control, desde la perspectiva jurídica, casi huelga decir que los profesionales jurídicos implicados en la supervisión, auditoría y control de procesos de robotización o implantación de soluciones de IA han de tener conocimientos específicos respecto al régimen jurídico de las TICs. Pero no solo eso: deben mostrar un manejo solvente de términos y conocimientos tecnológicos, o para ser más precisos, de conocimientos especializados en relación con el uso de la IA en el sector

impacto de este asunto ha sido tan alto, que acabó forzando la convocatoria de elecciones en Holanda.

⁴⁰ El Consejo de Estado italiano (equivalente a la Sala de lo Contencioso-Administrativo de nuestro Tribunal Supremo), ya ha tenido ocasión de pronunciarse sobre la exigencia de motivación de actuaciones administrativas basadas en algoritmos en dos sentencias que anulan decisiones administrativas que incumplen este requisito. Se trata de las sentencias núm.2270, de 8 de abril de 2019; y la núm.8472 de 13 de diciembre de 2019, entre otras.

⁴¹ Un tribunal contencioso-administrativo austríaco, en un caso en el que la Administración justificaba la imposibilidad de ofrecer una motivación adecuada respecto de la decisión adoptada por un sistema informático (VwSlg 11.728/1985), afirmó que “La legalidad de una decisión no puede medirse en función de las capacidades del sistema informático existente. Más bien, el sistema informático debe adaptarse a los requisitos legales”; el asunto se expone por F. MERLI, en el informe “Automated Decision Making Systems in national Administrative Law: An Austrian Perspective”, presentado en el seminario internacional celebrado en la Universidad de Friburgo los días 29 y 30 de septiembre de 2022, p.3.

⁴² Volveremos sobre ello en el apartado relativo a la motivación de los actos basados en sistemas de IA.

público, pues solo comprendiendo cómo se diseñan y funcionan estas soluciones tecnológicas estarán en disposición de supervisar, auditar y controlar que satisfacen adecuadamente las exigencias impuestas por el Derecho público.

Además, tratándose de desarrollos para el sector público implicados en la toma de decisiones (lo que significa integrarse, de manera directa o indirecta, en un procedimiento administrativo), estos profesionales deben contar con conocimientos solventes específicamente jurídico-administrativos, debido al juego, en este ámbito, de reglas y principios específicos, como exponía en la Introducción. Así, deben garantizar que el sistema automatizado o de IA se ajusta a requerimientos del procedimiento administrativo como los que analizamos en el apartado siguiente de este trabajo.

La literatura jurídica ya citada insiste en que esta labor de supervisión, auditoría y control del sistema no es puntual o concreta, en un determinado momento del desarrollo de la solución. Por el contrario, debe iniciarse desde el momento mismo del diseño del sistema y proyectarse a todo lo largo de su ciclo de vida: antes de su puesta en servicio, mediante las correspondientes pruebas y comprobaciones; y, una vez iniciado su funcionamiento, revisándolo periódicamente para validar su correcto rendimiento.

En efecto, no basta con auditar y supervisar el sistema antes de que comience a funcionar: es asimismo necesario llevar a cabo revisiones periódicas de su funcionamiento, para verificar su adecuación a Derecho. Desde el punto de vista jurídico, estas revisiones se orientan especialmente a depurar o evitar fallos de funcionamiento derivados de los datos empleados para entrenar el sistema y de las consecuencias que pueden deparar: por una parte, hay que comprobar adecuadamente que no se van produciendo sesgos –discriminatorios– de algún tipo; por otra parte, los sistemas se entrenan con datos antiguos que se usan para predecir o aplicarse al futuro, lo cual puede producir un desajuste o desactualización susceptible de alcanzar, asimismo, efectos discriminatorios, por lo que no solo es necesario analizar cómo se está comportando el sistema, sino también actualizar periódicamente los datos de que se nutre, evitando la congelación o esclerosis de la actuación administrativa y su inadaptación a las nuevas circunstancias contextuales que puedan sobrevenir.

Por otra parte, en ciertos procesos de control, y especialmente en la auditoría, será conveniente que el supervisor sea externo, a fin de preservar su objetividad. Esto determina la conveniencia –o en su caso, necesidad– de segmentar y dotar de independencia los procesos de control y auditoría respecto a los de desarrollo del sistema; en el caso de contratación, ya sea estableciendo en los pliegos las condiciones de tal independencia, ya sea dividiendo el contrato en lotes y contemplando la incompatibilidad de las adjudicaciones entre sí. Si bien, en este último caso, debe quedar suficientemente claro el modo de preservar el secreto industrial sin que ello llegue a impedir el efectivo despliegue de las labores de supervisión, auditoría y control.

Cuando los profesionales que se ocupan de todas estas tareas carecen de formación o conocimientos específicos del ámbito en cuestión es posible que se vean incapaces de desarrollar su labor, y, si se contrata por la Administración el desarrollo de un sistema de IA se corre el riesgo de “captura del regulador”, es decir, que los responsables administrativos sean persuadidos por el contratista de los términos en el que éste ha de ejecutar el contrato.

Ello puede conducir a riesgos muy diversos: no solo al incumplimiento del marco jurídico-público al que debe ceñirse la solución informática, sino que además, puede suceder que un proceso de toma de decisiones sea parametrizado finalmente por una empresa privada –no por el poder público–, quien no solo selecciona y procesa los datos mediante los que se nutre y entrena el sistema, sino que materialmente acaba decidiendo también los criterios que el propio sistema aplica para procesar esos datos y ofrecer decisiones.

Estos riesgos pueden atemperarse mediante la formación de empleados que puedan desempeñar eficientemente el control de ejecución del contrato y, por derivada, la supervisión, auditoría y control del sistema de información y de su código fuente⁴³. Si atendemos al volumen de contratación que están adquiriendo estas contrataciones,

⁴³ El borrador de la Estrategia Andaluza de Inteligencia Artificial contempla, entre sus medidas aplicables al sector público, la formación específica de equipos multidisciplinarios que puedan dar soporte transversal a las necesidades que planteen los diversos servicios y órganos administrativos en relación con los sistemas de IA.

la inversión que representa esta formación interna podría estar justificada.

Ahora bien, dada la especificidad y complejidad de esta materia, puede ser preferible para la Administración la externalización de un servicio de asesoramiento técnico en la contratación pública de estas soluciones. El servicio puede contemplar, así, la asistencia a la Administración contratante durante todas las fases de diseño y desarrollo de la solución: desde la redacción de los pliegos, a la adjudicación del contrato (asistiendo a la mesa de contratación mediante informes técnicos de valoración de las ofertas) y su ejecución (verificando las exigencias, ya apuntadas, de revisión periódica y auditoría del sistema).

Sería un control externo a la empresa desarrolladora, compatible con la exigencia de juristas en los equipos de trabajo de esta última, y que podría ser el mejor modelo de desarrollo de estas soluciones, al fortalecer la posición de la Administración y diferenciar entre un control, auditoría o supervisión interno (de la propia empresa contratista sobre su sistema) y otro externo, independiente y objetivo (el de la Administración pública, realizado a su vez por una empresa asesora cualificada).

2.4.4.- Incumplimiento del requisito de mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema: invalidez de la actuación automatizada por vulneración del art.48.2 LPAC

Como hemos podido comprobar a lo largo de este apartado, las exigencias impuestas sobre los sistemas de actuación administrativa automatizada no son menores, sino que tienen que ver con aspectos cruciales para garantizar su adecuado funcionamiento. Debemos tomar en consideración la radical importancia de estos requisitos para valorar las consecuencias de su eventual incumplimiento.

A mi juicio, si se adopta un acto administrativo mediante actuación administrativa automatizada (o decisivamente basado en ella), sin implementarse previamente las garantías anteriormente descritas, tal acto administrativo adolecería de un relevante vicio de forma que sería determinante de su anulabilidad. El art.48.2 LPAC dispone que “el defecto de forma sólo determinará la anulabilidad cuando el acto

carezca de los requisitos formales indispensables para alcanzar su fin o dé lugar a la indefensión de los interesados”. En el apartado siguiente analizaremos supuestos en los que el vicio de forma da lugar a indefensión. En este apartado, los requisitos formales que hemos comentado se encuadran en el otro ámbito de anulabilidad, por cuanto que las exigencias examinadas son absolutamente esenciales para acreditar que la actuación administrativa automatizada ha respondido a unos estándares mínimos de calidad en su diseño, implementación y mantenimiento⁴⁴.

Se trata de una situación análoga –*mutatis mutandis*– a la imposición de multas de tráfico mediante la captación de la velocidad de los vehículos por radares: ha de tratarse de radares homologados, que han de someterse además a labores periódicas de mantenimiento y calibración, pues en otro caso no se podrá demostrar la fiabilidad del sistema metrológico y el acto sancionador carecerá de los requisitos formales indispensables para alcanzar su fin⁴⁵.

En el caso de la IA, una particularidad relevante es que todavía no existen instrumentos unívocos y estandarizados mediante los que llevar a cabo las labores de control, supervisión y auditoría. Ante esta eventual objeción hay que responder con dos argumentos: 1) Sí existen métodos suficientemente depurados en algunas áreas; por ejemplo, en relación con la auditoría de datos se ofrece, guías de auditoría de datos para IA por la AEPD, de manera que ignorar sus postulados debería conducir a la anulación de los actos administrativos en que se haya traducido la actuación administrativa automatizada; 2) Aunque

⁴⁴ Algunos exponentes jurisprudenciales de anulación de actos administrativos por vulneración de requisitos de forma que impiden al acto alcanzar su fin son la STS 28/10/1998, rec.12.170/1991, en el que la comisión evaluadora de unas oposiciones no siguió los requisitos formales impuestos por las bases de la convocatoria.

⁴⁵ Hay dos particularidades a reseñar entre el supuesto del ejemplo y la omisión de requisitos de la actuación administrativa automatizada: el primero, que el radar constituye un elemento probatorio de la infracción que debe acreditar la Administración conforme al principio acusatorio, por lo que el vicio supondría indefensión, aunque a nuestro juicio, incurra simultáneamente en el supuesto de carecer de los elementos formales indispensables para alcanzar su fin; y el segundo, que se trata del ejercicio de la potestad sancionadora, por lo que el alcance del vicio no sería en este caso de anulabilidad, sino de nulidad de pleno Derecho.

existan dudas acerca de los mecanismos *idóneos* para llevar a cabo el control y supervisión de la actuación administrativa automatizada en general, y de la IA en particular, lo cierto es que al menos habrá tenido que implantarse *alguno*, y que tenga, además, un mínimo de contenido: a la Administración actuante no le bastaría con decir que el estado de la ciencia no le permite implantar estos controles, y omitirlos todos, sino que, al menos, deberá probar que ha implantado el sistema de IA con un mínimo de supervisión y control, desde el diseño y durante todo su ciclo de vida.

Afirmar lo contrario supondría reconocer la total impunidad del sector público en actuaciones que afectan a derechos e intereses de la ciudadanía, y en las que reputadas voces vienen advirtiendo de los peligros que entraña la aplicación de esta tecnología. Por lo que, ante la omisión de instrumentos mínimamente creíbles y fundamentados de control, supervisión y auditoría, la actuación automatizada debería anularse por carecer de los requisitos formales indispensables para alcanzar su fin.

3.- INTELIGENCIA ARTIFICIAL Y PROCEDIMIENTO ADMINISTRATIVO: GARANTÍAS DEL PROCEDIMIENTO ADMINISTRATIVO COMÚN

3.1.- *Encuadre: la pérdida de derechos ante la administración digital*

En este apartado recordamos la necesidad de que los sistemas de IA respeten las exigencias que la Ley 39/2015, de 1 de octubre del Procedimiento Administrativo Común de las Administraciones Públicas (LPAC) establece como reglas generales, aplicables a cualesquiera procedimientos, ya utilicen o no la IA o la actuación administrativa automatizada y, por tanto, también en el caso de que lo hagan⁴⁶. Dado que se trata de requisitos y garantías ampliamente

⁴⁶ En el derecho comparado se vienen manifestando posiciones en el mismo sentido. Por lo que se refiere al Derecho europeo, véase lo que expone MIR en el Capítulo XVI. Por lo que respecta a los Estados Unidos (pero incluyendo un análisis

estudiados y asentados en nuestro Derecho, prescindimos de reproducir aquí su teoría o encuadre general, remitiendo a la bibliografía general para ampliar la cuestión⁴⁷, y centrándonos en los aspectos específicos que se presentan cuando se trata de procedimientos administrativos en los que se utilizan soluciones de IA como soporte a la toma de decisiones.

El art.33.1 de la Ley 11/2007 establecía un recordatorio general de que el uso de medios electrónicos debe respetar los requisitos formales y materiales establecidos en las normas que regulen la correspondiente actividad⁴⁸, lo que incluye, naturalmente, las normas básicas o comunes relativas al régimen jurídico del sector público y del proce-

comparado más amplio), KAMINSKI, M.E. y URBAN, J.M. (2021): “The right to contest AI”, *Columbia Law Review*, vol.121, n.º 7, págs.1957 ss., disponible en <https://ssrn.com/abstract=3965041> (última visita: 28/03/2023).

⁴⁷ Véanse los capítulos correspondientes a las correlativas cuestiones en GAME-RO CASADO, E. (Dir.); FERNÁNDEZ RAMOS, S. y VALERO TORRIJOS, J. (Coords.) (2017): *Tratado de procedimiento administrativo común y régimen jurídico básico del sector público*, 2 vols., Tirant lo Blanch, Valencia; BAÑO LEÓN, J.M. y LAVILLA RUBIRA, J.J. (Drs.) (2021): *Comentarios al procedimiento administrativo*, Tirant lo Blanch, Valencia; CAMPOS ACUÑA, M.C. (coord.) (2021): *Comentarios a la Ley 39/2015, de procedimiento administrativo común de las Administraciones públicas*, La Ley-El Consultor, Madrid; TOR-NOS MAS, J. (Dir.) (2017): *Estudios sobre las leyes 39/2015 del procedimiento administrativo común y 40/2015 del régimen jurídico del sector público*, Atelier, Barcelona; RECUERDA GIRELA, M.A. (Dir.) (2016): *Procedimiento Administrativo Común de las Administraciones Públicas*, Aranzadi; GALLARDO CASTILLO, M. J. (2016): *Los procedimientos administrativos en la ley 39/2015: análisis y valoración de la reforma*, Tecnos, Madrid; GOSÁLBEZ PEQUEÑO, H. (Dir.) (2016): *La nueva Ley del procedimiento administrativo común*, Wolters Kluwer, Madrid; MARTÍNEZ GUTIÉRREZ, R. (2016): *El régimen jurídico del nuevo procedimiento administrativo común*, Aranzadi, Cizur Menor (Navarra); PALOMAR OLMEDA, A. (2022): *Procedimiento administrativo*, 3ª ed., Aranzadi-Thomson-Reuters, Cizur Menor (Navarra); y RIVERO GONZÁLEZ, M. (Dir.) (2016): *Procedimiento Administrativo Común y Régimen Jurídico*, Comares, Granada.

⁴⁸ Según este precepto, “La gestión electrónica de la actividad administrativa respetará la titularidad y el ejercicio de la competencia por la Administración Pública, órgano o entidad que la tenga atribuida y el cumplimiento de los requisitos formales y materiales establecidos en las normas que regulen la correspondiente actividad. A estos efectos, y en todo caso bajo criterios de simplificación administrativa, se impulsará la aplicación de medios electrónicos a los procesos de trabajo y la gestión de los procedimientos y de la actuación administrativa”.

dimiento administrativo común. A pesar de la indudable utilidad de esta disposición, la legislación vigente no contempla una previsión similar, lo que, aun siendo una lástima, no obsta, evidentemente, a que se exija su cumplimiento efectivo, como exponíamos en la Introducción, precisamente por tratarse de normas básicas o comunes aplicables al conjunto de la actividad administrativa sin que se establezcan excepciones cuando se usan medios electrónicos o soluciones de IA.

También debe traerse aquí a colación un precepto de carácter general que abarca a los dos primeros epígrafes de este apartado –el derecho a subsanar y a la audiencia–: el art.22 RGPD, conforme al cual cualquier persona física “tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar”⁴⁹.

De otro lado, aunque se trate de *soft law*, no puede olvidarse aquí que el 14 de julio de 2021 el Gobierno español aprobó la Carta de Derechos Digitales⁵⁰. Nos interesa la sección XVIII de la Carta, titulada: “Derechos digitales de la ciudadanía en sus relaciones con las Administraciones Públicas”. En particular, los apartados 6 y 7, relativos a la IA, establecen:

“6. Se promoverán los derechos de la ciudadanía en relación con la inteligencia artificial reconocidos en esta Carta en el marco de la actuación administrativa reconociéndose en todo caso los derechos a:

a) Que las decisiones y actividades en el entorno digital respeten los principios de buen gobierno y el derecho a una buena Administración digital, así como los principios éticos que guían el diseño y los usos de la inteligencia artificial.

b) La transparencia sobre el uso de instrumentos de inteligencia artificial y sobre su funcionamiento y alcance en cada procedimiento concreto y, en particular, acerca de los datos utilizados, su margen de error, su ámbito de aplicación y su carácter decisorio o no decisorio. La ley podrá regular las condiciones de transparencia y el acceso al código fuente,

⁴⁹ Sobre este precepto, véase lo que expone VALERO en el Capítulo IX de esta obra, apartado 4.1.1.

⁵⁰ En su parte introductoria este documento establece: “La Carta no tiene carácter normativo, sino que su objetivo es reconocer los novísimos retos de aplicación e interpretación que la adaptación de los derechos al entorno digital plantea, así como sugerir principios y políticas referidas a ellos en el citado contexto”.

especialmente con objeto de verificar que no produce resultados discriminatorios.

c) Obtener una motivación comprensible en lenguaje natural de las decisiones que se adopten en el entorno digital, con justificación de las normas jurídicas relevantes, tecnología empleada, así como de los criterios de aplicación de las mismas al caso. El interesado tendrá derecho a que se motive o se explique la decisión administrativa cuando esta se separe del criterio propuesto por un sistema automatizado o inteligente.

d) Que la adopción de decisiones discrecionales quede reservada a personas, salvo que normativamente se prevea la adopción de decisiones automatizadas con garantías adecuadas.

7. Será necesaria una evaluación de impacto en los derechos digitales en el diseño de los algoritmos en el caso de adopción de decisiones automatizadas o semiautomatizadas”.

El documento ha tenido un gran impacto mediático. La versión final fue mejorada en algunos aspectos, teniendo en cuenta las propuestas de un grupo de académicos pertenecientes a la Red de Derecho Administrativo e Inteligencia Artificial⁵¹. Aunque su aprobación se ha considerado positiva, se critica que, en lugar de proclamarse nuevos derechos en una Carta no vinculante, lo que debería hacerse es garantizar la efectividad de los derechos que ya tienen reconocidos los ciudadanos en la legislación sobre procedimiento administrativo común: la administración digital está suponiendo una pérdida de derechos y garantías para la ciudadanía, y tanto las normas como la jurisprudencia deben velar especialmente por el respeto de estas exigencias, tradicionales en nuestro acervo jurídico y esenciales en un Estado de Derecho ante las actuaciones del poder público⁵².

⁵¹ Red Derecho Administrativo e Inteligencia Artificial (Red DAIA), a la que pertenezco. Se puede consultar su web en <http://reddaia.org/> (última visita 28 de marzo de 2023). Para ver el documento de la propuesta, siga este enlace: <https://www.dropbox.com/s/gsbirh4kcn6v1k0/DAIAcartadic20.pdf?dl=0>, última visita 28 de marzo de 2023.

⁵² Ciertamente, la Administración electrónica está provocando una grave pérdida de garantías, que debe abordarse con carácter prioritario; véase GAMERO CASADO, E. (2022): “Reflexiones introductorias: de la administración electrónica a la digital (o la historia interminable)”, en A. Cerrillo Martínez (dir.), *La Administración Digital*, Dykinson, Madrid, págs. 36 y ss.

Tras el encuadre de la cuestión, pasemos a analizar las tres garantías esenciales que pueden verse –que se están viendo– comprometidas con el empleo de la IA en el sector público.

3.2.- El derecho a subsanar la presentación de escritos

El art.68.1 LPAC reconoce a los interesados en el procedimiento el derecho a subsanar sus solicitudes de iniciación cuando no reúnan los requisitos establecidos en las disposiciones aplicables al caso. Abunda en ello el art.14 RAFESP. A esta previsión se le debe añadir el derecho del interesado –art.28.2 LPAC– a no aportar datos que ya se encuentren en poder de la Administración actuante o hayan sido elaborados por cualquier otra Administración.

La consecuencia inevitablemente derivada de estos preceptos es que, si tras la presentación de la solicitud, un sistema robotizado o que aplique soluciones de IA procede de manera automatizada al cruce de datos, y no obtiene los necesarios para resolver el procedimiento, la consecuencia no puede ser la emisión directa de una resolución desfavorable, ni tampoco un plazo de alegaciones –que debe tener lugar una vez admitida a trámite la solicitud–, sino el otorgamiento de un trámite de subsanación “para que, en un plazo de diez días, subsane la falta o acompañe los documentos preceptivos, con indicación de que, si así no lo hiciera, se le tendrá por desistido de su petición” –art.68.1 LPAC–.

La jurisprudencia del Tribunal Supremo considera que la omisión del trámite de subsanación cuando el interesado no aportó con la solicitud los datos y documentos necesarios es un vicio invalidante que conlleva la anulación del acto resolutorio; así, STS 1623/2017, de 26 de octubre, rec.399/2015. No hay razón para concluir de distinto modo cuando esta situación se produzca como consecuencia de la tramitación automatizada del procedimiento administrativo.

Una vez presentada la solicitud, el procedimiento puede automatizarse íntegramente y resolverse mediante herramientas de IA cuando la resolución definitiva sea favorable y conceda lo solicitado. Pero si el motivo para desestimar la solicitud reside en la omisión de datos o documentos necesarios para resolver, es preceptivo llevar a cabo este trámite de subsanación. Un trámite que, por otra parte, también puede

automatizarse: el sistema puede requerir la subsanación de manera automatizada una vez comprueba que carece de los datos necesarios para resolver. Pero no cabe dictar directamente una resolución desfavorable por la mera omisión de datos o documentos necesarios para resolver que no se hayan aportado con la solicitud o no hayan podido recabarse *prima facie* mediante cruces de datos.

3.3.- *El derecho de audiencia*

El trámite de audiencia constituye una garantía tradicional del procedimiento administrativo, al que la jurisprudencia ha dedicado las más enfáticas calificaciones: trámite esencial, capital, fundamental e, incluso, “sagrado”. El art. 105.c) CE, al ordenar al legislador la regulación del procedimiento administrativo, precisa que la ley deberá garantizar, “cuando proceda”, la audiencia del interesado, de donde se infiere la posibilidad de omitir este trámite. El art.82.4 LPAC permite prescindir del trámite de audiencia cuando no figuren en el procedimiento ni sean tenidos en cuenta en la resolución otros hechos ni otras alegaciones y pruebas que las aducidas por el interesado, pues en tales casos es evidente que el cumplimiento del trámite de audiencia carece de función alguna, de modo que no existe incompatibilidad entre los derechos de contradicción y defensa y los principios de celeridad y economía procedimental (SSTC 68/1985 y 181/1990; STS 7/10/2008, rec. 6828/2005). En caso de tramitación simplificada del procedimiento, el art.96.6 d) LPAC únicamente exige la audiencia cuando la resolución “vaya a ser desfavorable al interesado” (no ya cuando “pueda”).

La omisión del trámite de audiencia en los casos en que resulta preceptivo es determinante de un vicio de invalidez que conlleva la anulación del acto administrativo resolutorio del procedimiento. La jurisprudencia considera que la ausencia de este trámite genera indefensión, por lo que el grado de invalidez depende de la tipología del procedimiento: si la indefensión tiene lugar en un procedimiento sancionador, afecta al art.24 CE y supone un vicio de nulidad radical o absoluta; así, SSTC 18/1981, 262/1988, 291/2000 y concordantes; y art.47.1 a) LPAC. En el resto de casos, hay dos líneas jurisprudenciales: una, que considera que el trámite es de tal importancia que su omisión equivale a prescindir total y absolutamente del procedimien-

to establecido –art.47.1 e) LPAC–, por lo que el grado del vicio es de nulidad radical (p.ej., SSTS 30/9/04, rec.2262/2002; 253/2019, de 26 de febrero); en tanto que la otra considera infringido el art.48.2 LPAC y que en tal caso el acto es meramente anulable por haberse producido indefensión (STS 981/2016, de 4 de mayo).

Sea cual fuere el grado del vicio, lo que resulta incontestable es que, tanto el Derecho positivo como la jurisprudencia consideran que se debe sustanciar el trámite de audiencia cuando la resolución del procedimiento sea desfavorable al interesado, lo cual no solo sucede en procedimientos de suyo restrictivos de derechos, como los sancionadores, los expropiatorios, los reintegros de subvenciones, etc., sino también los casos de solicitudes denegadas, como las autorizaciones o licencias, las subvenciones o cualesquiera otras. Y si antes de dictar la resolución desfavorable no se concede trámite de audiencia, el acto resolutorio adolece de un vicio de invalidez.

En esas mismas coordenadas ha de enjuiciarse el uso de soluciones de IA: no cabe que el sistema adopte un acto administrativo desfavorable sin sustanciar el trámite de audiencia⁵³. De no hacerlo, tal acto administrativo sería inválido.

Debe notarse, además, que la audiencia –art.82 LPAC– no solo reconoce el derecho a presentar alegaciones, sino también, a acceder al expediente administrativo⁵⁴. En ese expediente debe constar la actuación desarrollada mediante herramientas de IA, de manera comprensible para el interesado, a fin de que pueda, en su caso, desvirtuar la decisión o conclusión alcanzada por el sistema, ya sea por basarse en datos erróneos o sesgados –discriminatorios–, por fallos de programación, etc. Esta exigencia nos remite a lo que expone MARTÍN DELGADO en el Capítulo IV, y a lo que se desarrolla a continuación acerca de la motivación de los actos administrativos.

⁵³ En la misma posición, véase especialmente MORAL SORIANO, L. (2022): “Decisiones automatizadas...”, *op.cit.*, pág.483.

⁵⁴ Idea en la que abundan HUERGO y MORAL, según hemos expuesto más atrás.

3.4.- *La motivación de los actos administrativos adoptados mediante sistemas de IA*

3.4.1.- La motivación como requisito de validez de los actos dictados mediante IA

El ar.35 LPAC establece toda una batería de actos administrativos que deben encontrarse motivados; entre otros, los que limiten derechos e intereses legítimos: en este caso, el Tribunal Constitucional tiene establecido que la motivación no es solo una elemental corte-sía, sino un riguroso requisito del acto restrictivo de derechos: SSTC 26/1981, 8/1992, 52/1995 y 46/2014. La amplitud de estos supuestos determina que prácticamente todos los actos administrativos deban motivarse: así lo defiende NAVARRO GONZÁLEZ en su monografía sobre la motivación de los actos administrativos⁵⁵, y yo mismo he apuntado (en el prólogo de esa obra), que en este momento la cuestión no es ya determinar qué actos deben motivarse, sino cómo y cuánto se deben motivar los actos administrativos. Y eso es lo que debemos hacer ahora: determinar cuándo la motivación es suficiente para explicar las decisiones adoptadas mediante sistemas de IA.

Recordemos, a los solos efectos de centrar la cuestión, que la motivación es una exigencia meramente formal: la necesidad de que en el propio acto administrativo consten las razones por las que se dicta, con una sucinta explicación de los hechos y de los fundamentos de Derecho de la decisión. No se trata, por tanto, de que existan motivos para dictar el acto, sino de que esos concretos motivos sean transparentes y visibles, que se hagan constar en el propio acto. No constituye una garantía material o de fondo, sino formal, aunque la exigencia de motivar se establezca para dar soporte a un derecho material como el de no padecer indefensión (más adelante volveremos sobre este punto).

El Derecho positivo no establece expresamente en ninguna parte que los actos dictados mediante algoritmos o IA deban motivarse. Porque no hace falta: ya establece el art. 35 LPAC todos los actos que deben motivarse, sin necesidad de especificar su soporte o el proce-

⁵⁵ Véase NAVARRO GONZÁLEZ, R. (2017): *La motivación de los actos administrativos*, Aranzadi-Thomson-Reuters, Cizur Menor (Navarra).

so mediante el que se dictan. No se establece en ninguna parte una excepción para actos basados en algoritmos o IA y, por tanto, tal excepción no existe. El problema reside entonces en determinar cómo deben motivarse estos actos, no en discutir si deben motivarse o no.

La importancia de la motivación ha sido especialmente puesta de manifiesto en relación con las actuaciones automatizadas por MARTÍN DELGADO⁵⁶, URÍOS Y ALAMILLO⁵⁷, NAVARRO GONZÁLEZ⁵⁸, BERNING⁵⁹ y CERRILLO⁶⁰, señalando la esencialidad de la programación, y de su público conocimiento, en orden a satisfacer este requisito⁶¹.

Esta exigencia ha sido expresamente destacada en relación con la aplicación de la IA en el sector público por la Agencia de los Derechos Fundamentales de la Unión Europea⁶², encuadrándola en el derecho

⁵⁶ En Isaac MARTÍN DELGADO (2009): “Naturaleza...”, *op.cit.*, pág.377

⁵⁷ Véase URÍOS APARISI, X, y ALAMILLO DOMINGO, I. (2011): *La actuación...*, *op.cit.*, págs.21-31, 88, 133...

⁵⁸ Véase NAVARRO GONZÁLEZ, R. (2017): *La motivación...*, *op.cit.*, págs.474 ss.

⁵⁹ Véase Antonio David BERNING PRIETO (2019): *Validez e invalidez de los actos administrativos en soporte electrónico*, Aranzadi-Thomson-Reuters, Cizur Menor, págs.267 ss.

⁶⁰ En Agustí CERRILLO MARTÍNEZ (2020): “¿Son fiables...”, *op.cit.*, pp.31 ss.

⁶¹ En la misma línea se posicionan los trabajos de Juli PONCE SOLÉ en varias aportaciones; por todas, véase (2018): “La prevención de riesgos de mala administración y corrupción, la inteligencia artificial y el derecho a una buena Administración”, *Revista Internacional Transparencia e Integridad*, n.º 6. Entre los muchos autores que han insistido en la necesidad de motivar las decisiones sustentadas en algoritmos o en IA también pueden destacarse CRIADO, I; VALERO, J. y VILLODRE, J. (2020): “Algorithmic transparency and bureaucratic discretion: The case of SALER early warning system”, *Information Polity*, vol.25, n.º4, pág.452. Si bien, la mayoría de las argumentaciones ofrecidas por los autores citados en esta nota aluden a la transparencia, mientras que, como indicamos en el texto, nosotros encuadramos la exigencia no ya solo en ese contexto, sino más concretamente en la motivación del acto administrativo, como exigencia formal inexcusable que se debe incorporar al propio acto en cuanto que requisito de su contenido. Véase, asimismo, REICHEL, J. (2023): “Regulating Automation of Swedish Public Administration”, *Rivista Interdisciplinare sul Diritto delle Amministrazioni Pubbliche*, n.º 1/2023, págs.83 ss.

⁶² En su informe *Getting the future right: Artificial intelligence and fundamental rights*, European Union Agency for Fundamental Rights, Viena, 2020, pág.81; accesible en <https://fra.europa.eu/en/publication/2021/getting-future-right-artificial-intelligence-and-fundamental-rights-summary> (última visita: 28/03/2023).

a la buena administración reconocido en la Carta Europea de los Derechos Fundamentales, que menciona expresamente la necesidad de motivar los actos administrativos [artículo 41.2 c)]

El problema tiene mucho que ver con el binomio explicabilidad-opacidad de los algoritmos de IA, cuestión a la que muchos autores se han aproximado desde la perspectiva de la transparencia⁶³, o de su integración en el expediente administrativo⁶⁴. Nosotros lo abordamos aquí bajo el prisma de la motivación del acto administrativo, como exigencia formal inexcusable que se debe incorporar al propio acto, como requisito de su contenido, al igual que han hecho otros autores: así, MARTÍN DELGADO, URÍOS y ALAMILLO, NAVARRO GONZÁLEZ, CERRILLO, HUERGO y MORAL⁶⁵. Esta aproximación es más contundente desde el punto de vista de sus efectos, puesto que incumplir los deberes de transparencia suele resultar bastante inocuo para el sector público, mientras que la omisión de motivación es causa de invalidez de los actos administrativos, como veremos enseguida.

⁶³ Véanse, especialmente, el Capítulo IV de esta obra, así como COTINO HUESO, L. y CASTELLANOS CLARAMUNT, J. (Eds.) (2020): *Transparencia y explicabilidad de la inteligencia artificial*, Tirant lo Blanch, Valencia.

⁶⁴ Para HUERGO LORA, A. (2020): “Una aproximación...”, *op.cit.*, pág.85, “en la medida en que un algoritmo sirva para tomar decisiones que afecten a un ciudadano, la Administración debe hacerlo público del mismo modo que cualquier otro documento que forme parte del expediente administrativo, en el que el algoritmo se integraría en virtud del artículo 70 LPAC, puesto que sirve de antecedente y fundamento a la decisión administrativa”. En la misma posición, MORAL SORIANO, L. (2022): “Decisiones automatizadas, Derecho administrativo y argumentación jurídica”, en F. LLANO ALONSO, (Dir.), *Inteligencia artificial y Filosofía del Derecho*, Laborum Ediciones, Murcia, págs. 475-499

⁶⁵ Véanse MARTÍN DELGADO, I. (2009): “Naturaleza...”, *op.cit.*, pág.377; URÍOS APARISI, X. y ALAMILLO DOMINGO, I. (2011): *La actuación...*, *op.cit.*, págs. 21 ss.; NAVARRO GONZÁLEZ, R.: *La motivación...*, *op.cit.*, pág.474; y CERRILLO MARTÍNEZ, A. (2020): “Automatización...”, *op.cit.*, pág.405; HUERGO LORA, A. (2021): “Administraciones públicas e inteligencia artificial: ¿más o menos discrecionalidad”, *El Cronista del Estado Social y Democrático de Derecho*, n.º 96-97, pág.12 (hemos manejado la versión accesible en la sección “Estudios y comentarios” del INAP, <https://laadministracionaldia.inap.es/noticia.asp?id=1512494>); y MORAL SORIANO, L. (2022): Decisiones automatizadas...”, *op.cit.*, págs. 483 ss.; entre otros.

3.4.2.- Alcance y medio de la motivación: reglas especiales en relación con la IA

El problema central en toda esta cuestión es cómo articular la motivación cuando se trata de decisiones sustentadas en herramientas de IA: ¿en qué debe consistir en este caso la sucinta referencia a hechos y fundamentos de Derecho a que alude el art.35.1 LPAC?

Como venimos diciendo, la cuestión, en definitiva, tiene que ver con la explicabilidad de los algoritmos y de la IA, en la que se depositan actualmente grandes esfuerzos, entre los que debemos destacar el Centro para la Transparencia Algorítmica que el *Joint Research Centre* de la Comisión Europea va a implantar en Sevilla. Sus estudios ayudarán sin duda a ofrecer soluciones a este problema. En el estado actual de la tecnología y de nuestros conocimientos, podemos esbozar apenas algunos apuntes.

Estamos haciendo referencia a decisiones que directa o indirectamente se fundamenten en procesos algorítmicos o inteligencia artificial, y el requisito de la motivación, como advierte VALERO, debe adaptarse a las particularidades que entraña el uso de esta tecnología⁶⁶.

Comenzando por la sucinta referencia a *hechos*, desde el momento en que el aprendizaje (y la decisión) de los sistemas de IA se realiza mediante procesamiento de *big data*, resulta crucial conocer los bancos de datos seleccionados y ofrecidos al sistema, pues constituyen en este caso los *hechos* sobre los que se sustenta la decisión. Solo de esta manera será posible analizar la calidad de los datos –volumen, ausencia de sesgos discriminatorios, etc.– y la licitud de su uso.

En cuanto a los *fundamentos de Derecho*, consistirá en la explicación de por qué se ha dictado el acto. En el caso de la IA, esta información reside en el instrumento mediante el que se emite la decisión o en el que la misma se sustenta: es el proceso lógico que conduce a adoptar el acto. La motivación, en este tipo de actos, es necesariamente la explicación transparente (aunque pueda ser sucinta), en el propio acto o en algún otro medio a disposición del interesado, del proceso lógico que conduce a su adopción.

⁶⁶ Véase VALERO TORRIJOS, J. (2019): “Las garantías...”, *op.cit.*, pág.88.

En particular, URÍOS y ALAMILLO⁶⁷ consideran que la exigencia de motivación supone en este caso “la necesidad de codificar y poder reconstruir, para cada caso singular, las reglas lógicas – ya hemos visto que en una aproximación híbrida, con la aplicación de un método integrado por la lógica de predicados de primer orden, por las lógicas modales, deóntica y refutable aplicables, y por la lógica descriptiva en cuanto a la representación del dominio de conocimiento jurídico – que han sido aplicadas en el acto administrativo automático singular”. Como se ve, no se trata tanto de ofrecer una explicación en lenguaje natural, sino de disponer de los medios e instrumentos necesarios para poder justificar adecuadamente los criterios y procesos mediante los que el sistema toma la decisión. Tampoco se alude necesariamente al código fuente, cuyo conocimiento puede ser a menudo decisivo, pero no puede olvidar otros elementos de necesaria consideración.

El Consejo de Estado italiano (equivalente a la Sala de lo Contencioso-Administrativo de nuestro Tribunal Supremo), ya ha tenido ocasión de pronunciarse sobre la exigencia de motivación de actuaciones administrativas basadas en algoritmos en dos sentencias que anulan decisiones administrativas que incumplen este requisito. Se trata de las sentencias núm.2270, de 8 de abril de 2019; y la núm.8472 de 13 de diciembre de 2019. En esta última se exige, como requisito para la aplicación de algoritmos⁶⁸, el “pleno conocimiento del módulo utilizado y de los criterios aplicados”, así como que “este conocimiento del algoritmo debe garantizarse en todos los aspectos: desde sus autores hasta el procedimiento utilizado para su elaboración, el mecanismo de decisión, las prioridades asignadas en el procedimiento de evaluación y toma de decisiones y los datos seleccionados como relevantes (...) a fin de poder verificar que los criterios, condiciones y resultados del procedimiento robótico cumplen con las prescripciones y las finalidades establecidas por la ley o por la propia administración sobre dicho procedimiento y para que queden claras —y, por lo tanto,

⁶⁷ *Op.ult.cit.*, pág.134.

⁶⁸ Tomamos la traducción de Agustí CERRILLO MARTÍNEZ (2020): “¿Son fiables...”, *op.cit.*, págs. 31 ss. La recomendable lectura de este trabajo ofrece otros posicionamientos alineados con la necesidad de motivar las decisiones sustentadas en algoritmos.

sean cuestionables— las modalidades y reglas a partir de las que se haya programado”.

En efecto, conocer los criterios con los que se ha programado la toma de decisión (o uno de los elementos en los que se apoya) es crucial para determinar si se ha incurrido en un sesgo de alcance discriminatorio⁶⁹, o más sencillamente, si los criterios en cuestión son objetivos y no incurrir en la arbitrariedad proscrita por los artículos 9.1 y 103.2 CE.

Uno de los aspectos concernidos por esa necesidad de información es determinar si resulta o no preciso el acceso al código fuente. Esta cuestión se expone ampliamente por MARTÍN DELGADO en el Capítulo IV, al que remito para no ser reiterativo⁷⁰. En el mismo lugar, pero ya con una perspectiva más amplia de las exigencias de motivación en casos de uso de IA, razona el autor que: “la información relativa al diseño del sistema, su finalidad de uso, su lógica de funcionamiento, los datos de los que se nutre, el margen de error testado y el nivel de precisión verificado o los resultados de las auditorías realizadas contribuye a la motivación de la decisión adoptada sobre la base del mismo por la Administración; por el contrario, aun cuando se realicen ingentes esfuerzos por motivar en el caso concreto la decisión, difícilmente en ausencia de la información anteriormente mencionada podrá aceptarse la misma cuando se trate de sistemas autónomos. No bastan las razones; resulta imprescindible saber que el sistema funciona de forma correcta”.

Por su parte, la Carta de Derechos Digitales (que no tiene fuerza vinculante), reconoce el derecho de la ciudadanía en sus relaciones con las Administraciones públicas a:

“c) Obtener una motivación comprensible en lenguaje natural de las decisiones que se adopten en el entorno digital, con justificación de las normas jurídicas relevantes, tecnología empleada, así como de los criterios de aplicación de las mismas al caso. El interesado tendrá derecho

⁶⁹ En esa línea, Agustí CERRILLO MARTÍNEZ (2019): “El impacto de la inteligencia artificial en el Derecho administrativo: ¿nuevos conceptos para nuevas realidades técnicas?”, *Revista General de Derecho Administrativo*, n.º 50, pág.14.

⁷⁰ La profusa explicación, con abundante aparato bibliográfico, la realiza el autor en la parte final del Capítulo.

a que se motive o se explique la decisión administrativa cuando esta se separe del criterio propuesto por un sistema automatizado o inteligente”.

A mi juicio, esto no significa que, en todo caso, se deba suministrar la información en un lenguaje comprensible para el común de la ciudadanía. Al igual que sucede con decisiones administrativas adoptadas en ámbitos sectoriales específicos, puede ser necesario que un especialista se pronuncie sobre ciertos aspectos, que solo son verdaderamente comprensibles para los expertos del ramo. Como advierten KEARNS y ROTH⁷¹, “la primera pregunta que acude a la mente cuando se intenta formular una teoría de la interpretabilidad algorítmica es ¿interpretable para quién?”. En relación con la IA sería el caso, por ejemplo, de la programación del sistema: sería absurdo pretender que se facilite al interesado en lenguaje natural. Ahora bien, en relación con este problema es necesario proclamar dos reglas claras: *a)* Se debe ofrecer en lenguaje natural y comprensible toda la información que sea posible; *b)* Se debe ofrecer toda la información necesaria para motivar adecuadamente la decisión, sin que la Administración pueda escudarse en su tecnicismo para denegar el acceso.

Acerca de esto último, no hacemos más que predicar el mismo esquema formal que se manifiesta en cualquier ámbito de actuación administrativa, por ejemplo, cuando un acto se sustenta en un infor-

⁷¹ KEARNS, M. y ROTH, A. (2020): *El algoritmo ético. La ciencia del diseño de algoritmos socialmente responsables*, LaLey-Wolters Kluwer, Madrid, pág.235. Los autores desarrollan esta idea del siguiente modo pág.236: “A alguien con una formación matemática escasa o nula, el concepto de correspondencia de objetos matemáticos exactos, como, por ejemplo, las características extraídas de una aplicación crediticia para predecir la probabilidad de devolución del préstamo, le resultará totalmente ajeno. Simplemente, carecerá de la formación necesaria para reflexionar sobre dicha función y, para él, el modelo no será interpretable. En cambio, para un observador familiarizado con los fundamentos de la modelización estadística clásica, aunque quizá no tanto con el aprendizaje automático moderno, un modelo lineal que efectúe una suma ponderada de las características de entrada sería el ejemplo canónico de la interpretabilidad; por ejemplo, mirará los signos de los datos y verá que permanecer muchos años en un trabajo guarda una correlación positiva con la devolución de un préstamo, y que tener antecedentes penales se asocia a una correlación negativa. Sin embargo, este mismo observador podría considerar que una red neuronal multicapa no es interpretable, porque no permite extraer con facilidad correlaciones tan simples”.

me técnico desfavorable: el ciudadano no podrá por sí mismo comprender y rebatir el contenido del informe, pero sí podrá acudir a un experto para que manifieste su posición, desvirtuándolo y logrando, de ese modo, destruir la motivación de la decisión, lo que conlleva su anulación, como destacaremos más adelante.

Ahora bien, debe quedar claro que una cosa es la explicabilidad del algoritmo desde el punto de vista de su operativa tecnológica, y otra bien distinta que tal explicabilidad o interpretabilidad sea suficiente para cumplir con la exigencia de motivar los actos administrativos. No basta con una descripción –inteligible– del proceso, sino que es preciso demostrar que ese proceso se ajusta a Derecho⁷².

Por otra parte, el requisito de la motivación no exige estrictamente que ésta se consigne en el propio acto administrativo. El art. 88.6 LPAC dispone que “La aceptación de informes o dictámenes servirá de motivación a la resolución cuando se incorporen al texto de la misma”; es lo que se conoce como motivación *in allunde*, o por remisión. Ciertamente el precepto restringe ese modo de articular la motivación a la existencia de informes y dictámenes, que además se incorporen al texto del propio acto administrativo; pero la jurisprudencia es más flexible, admitiendo otras formas de motivar, siempre que se permita de alguna manera al interesado conocer suficientemente las razones en las que se ha basado la decisión⁷³. Esto permite que no sea estrictamente al acto adoptado mediante IA el que contenga la información necesaria para motivarlo, es decir, que el propio acto administrativo dictado en aplicación de algoritmos o IA no tiene que incorporar necesariamente los pormenores de la programación del sistema ni la acreditación de su verificación de adecuado funcionamiento, pero sí debe suministrar información suficiente para que el interesado pueda localizar y obtener esos datos, o bien haber dispuesto de ellos durante la tramitación del procedimiento administrativo, constando en

⁷² Como sostiene HUERGO LORA, A. (2020): “Una aproximación...”, *op.cit.*, pág. 74, “es preciso conocer el algoritmo para saber si puede haber introducido algún contenido no previsto en la norma, distorsionando su aplicación o llevándolo, en definitiva, a que se dicte una decisión contraria a Derecho.

⁷³ Para profundizar en la cuestión, véase la extensa exposición de la cuestión que lleva a cabo al exponer la motivación *in allunde* NAVARRO GONZÁLEZ, R. (2017): *La motivación...*, *op.cit.*, Capítulo V, § 2.2.

el expediente o en otras fuentes de información a su alcance⁷⁴. Por ejemplo, CANALS propone la creación de un repositorio de acceso restringido (protección de secretos comerciales) que permita el acceso al código fuente para realizar el control.⁷⁵

No puede descartarse que la motivación de este tipo de actos deba incluir también la acreditación de que el sistema de información mediante el que se ha adoptado la decisión haya superado exitosamente las exigencias de auditoría, control y supervisión que analizamos más atrás, con ocasión del estudio de la actuación administrativa automatizada. Por ejemplo, facilitando las pruebas de funcionamiento de que ha sido objeto el sistema, así como el resto de instrumentos de auditoría, control y supervisión que se hayan desplegado. Se trata aquí de acreditar que el sistema funciona correctamente, como se hace, por ejemplo, con los radares que se utilizan para imponer multas de tráfico: el acuerdo de incoación del procedimiento no solo incluye la fotografía del vehículo, sino también la acreditación de que el radar se encuentra homologado y de que ha superado las inspecciones metrológicas periódicas correspondientes⁷⁶.

⁷⁴ En el mismo sentido, se ha defendido que el algoritmo mediante el que se toma la decisión debe constar en el expediente, así, HUERGO LORA, A. (2021): “Administraciones...”, *op.cit.*, pág.12.

⁷⁵ Véase CANALS AMETLLER, D. (2021): “El ejercicio de potestades administrativas por operadores privados en régimen de mercado”, en E. Gamero Casado (Dir.): *La potestad administrativa. Concepto y alcance práctico de un criterio clave para la aplicación del Derecho administrativo*, Tirant lo Blanch, Valencia, pág.375.

⁷⁶ Cuando no se acredita que el cinemómetro (radar) con el que se obtuvo la medición de velocidad del vehículo ha pasado las revisiones periódicas que verifican su fiabilidad (actualmente impuestas en la Orden ICT/155/2020, de 7 de febrero, por la que se regula el control metrológico del Estado de determinados instrumentos de medida), no existe prueba de cargo que permita sancionar (así, Sentencia de la Sala de lo Contencioso-Administrativo del Tribunal superior de Justicia de Castilla y León 1465/2001, de 29 de septiembre; y otras muchas concordantes, especialmente de los Juzgados de lo Contencioso-Administrativo). Ante el alud de recursos que esta jurisprudencia supuso, la práctica actual de la DGT es acompañar el correspondiente certificado al acuerdo de incoación del expediente sancionador, lo cual sirve de fundamento probatorio (y, por consiguiente, motivación) de la sanción.

Todas estas exigencias deben tenerse en cuenta en el diseño de los sistemas de IA, que han de tener prevista la forma de satisfacerlas⁷⁷.

Como reflexión final en relación con el alcance de la motivación en los sistemas de IA del sector público podemos llegar a la *conclusión de que todavía no está completamente claro cómo motivar las decisiones sustentadas en IA*, pues se sigue madurando en sus instrumentos de auditoría, supervisión y control. Pero en el apartado siguiente comprobaremos que, cuando no es posible ofrecer una explicación clara y convincente de cómo funciona el sistema, los tribunales están anulando las decisiones adoptadas mediante IA. *En suma: no podemos determinar con precisión cómo lograr que la motivación de actos administrativos sustentados en IA sea suficiente; pero sí sabemos que, de ser insuficiente (incapaz de explicar cómo y por qué se ha tomado la decisión), el acto administrativo se anulará*. Comprobémoslo.

3.4.3.- La omisión de motivación genera indefensión y determina la invalidez de los actos dictados con sistemas de IA. Jurisprudencia nacional, comparada y del TJUE

El fin de la motivación (o, al menos, el más importante de ellos), es habilitar el derecho a la tutela judicial efectiva de los sujetos cuyos derechos e intereses legítimos puedan verse afectados por la decisión. Solo podemos rebatir un acto administrativo si sabemos por qué se ha dictado. La falta de motivación suprime un elemento clave de control e impide verificar la adecuación a Derecho del acto administrativo.

La jurisprudencia está plenamente habilitada para efectuar este control, por complejo que sea, puesto que no hay diferencia sustantiva entre controlar el uso de la tecnología y cualquier otro escenario de control asignado a los Tribunales, en materias sectoriales de extraordinario tecnicismo, sobre las que inevitablemente se tienen que pronunciar. Como recuerda HUERGO⁷⁸, los jueces y magistrados han de pronunciarse frecuentemente sobre cuestiones extremadamente complejas, que sobrepasan el ámbito de su formación –jurídica– específica, lo cual

⁷⁷ Como reclama VALERO TORRIJOS, J. (2019): “Las garantías...”, *op.cit.*, pág.91, con cita de Steels y López

⁷⁸ En HUERGO LORA, A (2021): “Administraciones...”, *op.cit.*, pág.12.

no obsta a que puedan, y deban, llevar a cabo ese control. Está en juego el derecho a la tutela judicial efectiva sin que pueda producirse indefensión –art.24 CE–, y, como afirma DE LA SIERRA⁷⁹, la jurisdicción contencioso-administrativa está decisivamente llamada a desplegar su control sobre la motivación de los actos administrativos sustentados en IA para impedir la vulneración de este derecho⁸⁰.

La jurisprudencia contencioso-administrativa ha establecido que los actos administrativos carentes de motivación (cuando ésta es preceptiva) generan indefensión (SSTS 1198/2019, de 19 de septiembre; 198/2019, de 19 de febrero; y 10/5/2000, rec.5.760/1995, entre otras muchas). Pero, además, cuando se encuentra en entredicho un derecho fundamental (derecho de reunión, procedimientos sancionadores), la exigencia de motivación tiene alcance constitucional (SSTC 26/1981, f.j.14; 236/2007, f.j.12; 17/2009, f.j.2; y 46/2014, f.j.4, entre otras muchas), por lo que determina un vicio de nulidad radical o de pleno Derecho [artículo 47.1 a) LPAC].

Por tanto, debo insistir en la posición de que, aun cuando actualmente no exista un precepto que exija la motivación de los actos administrativos sustentados en algoritmos o IA, y a pesar también de que dudemos cómo lograr una motivación suficiente de dichos actos, tal requisito es inequívocamente predicable en los supuestos establecidos con carácter general por el artículo 35.1 LPAC; y que, en tales supuestos, la omisión del requisito determina un vicio de invalidez del acto, susceptible de impugnación en vía administrativa y contencioso-administrativa.

También URÍOS y ALAMILLO consideran que la omisión de motivación (en los términos expuestos) representa un vicio de invalidez

⁷⁹ En DE LA SIERRA MORÓN, S. (2021): “Control judicial de los algoritmos: robots, administración y estado de derecho”, Tribuna publicada en *ElDerecho.com*, el 11/06/2021, apartado de conclusiones; accesible en <https://elderecho.com/control-judicial-de-los-algoritmos-robots-administracion-y-estado-de-derecho> (última visita: 28/03/2023). Véase también, de la misma autora, “Inteligencia artificial y justicia administrativa: una aproximación desde la teoría del control de la Administración pública”, *Revista General de Derecho Administrativo*, n.º 53, 2020.

⁸⁰ En la misma línea, véanse TERRÓN SANTOS, D. y DOMÍNGUEZ ÁLVAREZ, J.L. (2022): *i-Administración pública...*, *op.cit.*, págs.161 ss.

de las decisiones administrativas automatizadas⁸¹. Debemos extrapolar esa conclusión a las decisiones adoptadas mediante algoritmos e IA, como ha hecho VALERO⁸².

No nos consta la existencia de jurisprudencia española que haya acogido aún esta causa de invalidez para anular decisiones basadas en IA. Un precedente que presenta cierta proximidad con esta cuestión es la Sentencia del Juzgado de lo contencioso-administrativo n.º 8 de Madrid, de 30/12/2021, sobre BOSCO, el algoritmo mediante el que se decide reconocer o no el bono social eléctrico. Tras comprobarse que el algoritmo discriminaba a ciertas personas (como las viudas), la Fundación CIVIO presentó una solicitud de acceso a su código fuente, basada en la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno. El derecho de acceso se denegó por el Juzgado acogiendo la alegación de la Abogacía del Estado, a cuyo decir afectaba a la seguridad nacional, al incluirse en el código fuente la manera en que el sistema se interconecta con el resto de sistemas de información del Estado, inclusive los de Defensa e Interior. La sentencia está recurrida ante la Audiencia Nacional, y con ocasión de esta apelación el Consejo de Transparencia ha modificado su posición y apoya el recurso, basado en la idea de segmentar el código fuente, suprimiendo los elementos que afecten a la seguridad nacional, y manteniendo los que explican las reglas o criterios aplicados por el algoritmo. A mi juicio, la cuestión estaría ya resuelta de otro modo si, en lugar de canalizarse este asunto mediante el ejercicio del derecho de acceso, se hubiera articulado mediante la impugnación de una denegación del bono social invocando su falta de motivación.

En el Derecho comparado sí existen importantes precedentes que anulan decisiones adoptadas mediante herramientas de IA que no permitían ofrecer una motivación suficiente⁸³.

⁸¹ Véase URÍOS APARISI, X. y ALAMILLO DOMINGO, I. (2011): *La actuación...*, *op.cit.*, pág.94.

⁸² En VALERO TORRIJOS, J. (2019): “Las garantías...”, *op.cit.*, págs.87 y 90. También SORIANO ARNÁNZ, A. (2021): “Decisiones automatizadas: problemas...”, *op.cit.*, p. 94, considera que la falta de transparencia de los sistemas de IA dificulta su control judicial.

⁸³ Además de los trabajos citados en el resto del apartado, véase DE LA SIERRA MORÓN, S. (2021): “Control judicial de los algoritmos...”, *op.cit.*, §§ 4 y 5.

Así, la falta de motivación es una de las causas en las que se sustentó la sentencia del Tribunal de Distrito de La Haya de 5 de febrero de 2020 (ECLI:NL:RBDHA:2020:865), mediante la que anuló una elaboración de perfiles para la lucha contra el fraude implantada por el Gobierno holandés que únicamente se utilizó en los calificados como “vecindarios problemáticos”, concluyendo el Tribunal que ese criterio puede provocar exclusión, discriminación y estigmatización injustificadas⁸⁴. Desde 2013, 26.000 familias fueron injustamente acusadas de fraude.

La falta de motivación fue, asimismo, el fundamento de sendas sentencias, ya citadas, del Consejo de Estado italiano, la núm. 2270, de 8 de abril de 2019; y la núm.8472 de 13 de diciembre de 2019, basan la anulación de decisiones adoptadas por la Administración mediante sistemas de IA por no ser capaces de aportar motivación suficiente; estos y otros pronunciamientos han sido comentados por GALETTA, quien sostiene que la jurisprudencia italiana exige que la Administración ofrezca un modo de motivar las decisiones sustentadas en algoritmos⁸⁵. También en Italia se ha utilizado el mismo argumento en una sentencia relativa a los *riders* (ciclistas) de Deliveroo: Sentencia del Tribunal Ordinario de Bolonia de 31/12/2020⁸⁶; aunque se trata de un pronunciamiento relativo a empresas privadas y no a la Administración pública, es evidente su clara identidad con el problema que analizamos.

⁸⁴ Véase un comentario a este asunto en COTINO HUESO, L. (2020): “‘SyRI, ¿a quién sanciono?’ Garantías frente al uso de inteligencia artificial y decisiones automatizadas en el sector público y la sentencia holandesa de febrero de 2020”, *La Ley Privacidad*, n.º 4; accesible en https://www.academia.edu/42979949/_SyRI_a_qui%C3%A9n_sanciono_Garant%C3%ADas_frente_al_uso_de_inteligencia_artificial_y_decisiones_automatizadas_en_el_sector_p%C3%BAblico_y_la_sentencia_holandesa_de_febrero_de_2020 (última visita: 28/03/2023).

⁸⁵ Véase GALETTA, D.U. (2020): “Algoritmi, procedimento amministrativo e garanzie: brevi riflessioni, anche alla luce degli ultimi arresti giurisprudenziali in materia”, *Rivista Italiana di Diritto Pubblico Comunitario*, n.º 4/2020, p. 501 ss.

⁸⁶ Un comentario a esta sentencia en CARRILLO PARRILLA, J.A. (2021): “La discriminación a través del algoritmo”, *Observatorio Pagoda*, 11/03/2021, accesible en <https://www.uv.es/fatwireed/catedra-pagoda/es/novedades-1286053802801/Novetat.html?id=1286182093538> (última visita: 28/03/2023).

También en Austria los tribunales rechazan los sistemas de IA que no pueden explicar su funcionamiento, como explica MERLI⁸⁷. Y según explica HUERGO, en Francia, el Consejo Constitucional ha establecido que, cuando se interpone un recurso contencioso-administrativo contra una decisión basada en algoritmos el juez puede pedir a la Administración que le comunique sus características⁸⁸.

Hemos de destacar especialmente la jurisprudencia del TJUE en la cuestión. En particular, la Sentencia del Tribunal de Justicia -Gran Sala- de 21 de junio de 2022, *Ligue des droits humains*, asunto C-817/19 (ECLI:EU:C:2022:491), que se extiende ampliamente sobre la necesidad de que los sistemas de IA ofrezcan a los interesados una explicación –motivación– suficiente de la decisión. El caso versa sobre la cuestión prejudicial que el Tribunal Constitucional belga elevó al TJUE, acerca del tratamiento automatizado de los datos del registro de nombres de los pasajeros (PNR)⁸⁹, dentro de la UE. Tales datos pueden ser utilizados para la prevención, detección, investigación y enjuiciamiento de delitos terroristas y delitos graves, u pueden utilizarse, por ejemplo, para la evaluación previa a la llegada o a la salida de los pasajeros en comparación con criterios de riesgo predeterminados, sirviendo de base a las autorizaciones de entrada o salida del país. Conviene resaltar los siguientes pasajes de la sentencia:

⁸⁷ En MERLI, F. (2023): “Automated Decision-Making Systems in Austrian Administrative Law”, *Rivista Interdisciplinare sul Diritto delle Amministrazioni Pubbliche*, n.º 1/2023, pág.45, el autor sostiene que “unexplainable AI results are therefore no option under Austrian law. (...) the holdings on the reasoning requirements are motivated to a large degree by the necessity to secure effective legal protection”.

⁸⁸ HUERGO LORA, A. (2020): “Una aproximación...”, *op.cit.*, pág.82. La sentencia del Consejo Constitucional versó sobre un precepto legal que autoriza a la Administración a utilizar algoritmos siempre que informen al interesado: véase, en el mismo trabajo, pág.80 y nota 83.

⁸⁹ Este tratamiento automatizado de los datos del registro de nombre de pasajeros (PNR) se regula en la Directiva (UE) 2016/681 del PE y del Consejo, de 27 de abril de 2016 relativa a la utilización de datos del registro de nombre de los pasajeros (PNR) para la prevención, detección, investigación y enjuiciamiento de los delitos de terrorismo y de la delincuencia grave; fue traspuesta al Derecho español por la Ley Orgánica 1/2020, de 16 de septiembre, sobre la utilización de los datos del Registro de Nombres de Pasajeros para la prevención, detección, investigación y enjuiciamiento de delitos de terrorismo y delitos graves.

“(195) Es preciso añadir que el recurso a estas tecnologías entrañaría el riesgo de privar de efecto útil a la revisión individualizada de los resultados positivos y al control de licitud exigido por las disposiciones de la Directiva PNR. En efecto, como hace constar, en esencia, el Abogado General en el punto 228 de sus conclusiones, habida cuenta de la opacidad que caracteriza el funcionamiento de las tecnologías de inteligencia artificial, puede resultar imposible comprender la razón por la cual un determinado programa ha alcanzado una concordancia positiva. En estas circunstancias, el uso de esas tecnologías también podría privar a las personas afectadas de su derecho a la tutela judicial efectiva, que consagra el artículo 47 de la Carta y que la Directiva PNR pretende, a tenor de su considerando 28, garantizar en un nivel elevado, en particular para cuestionar el carácter no discriminatorio de los resultados obtenidos. (...)”

(210) En particular, las autoridades competentes deben asegurarse de que el interesado, al que no necesariamente se permite, en el procedimiento administrativo, tomar conocimiento de los criterios de evaluación predeterminados y de los programas que aplican tales criterios, pueda comprender el funcionamiento de esos criterios y de esos programas, de forma que pueda decidir, con pleno conocimiento de causa, si ejerce o no su derecho a un recurso judicial, garantizado en el artículo 13, apartado 1, de la Directiva PNR, con objeto de cuestionar, en su caso, el carácter ilícito y, en particular, discriminatorio de los mencionados criterios (véase, por analogía, la sentencia de 24 de noviembre de 2020, *Minister van Buitenlandse Zaken*, C-225/19 y C-226/19, EU:C:2020:951, apartado 43 y jurisprudencia citada). Así debe ocurrir también con los criterios de revisión mencionados en el apartado 206 de la presente sentencia.

(211) Por último, al sustanciar un recurso interpuesto con arreglo al artículo 13, apartado 1, de la Directiva PNR, el juez que ejerce el control de la legalidad de la resolución adoptada por las autoridades competentes y, salvo en los casos de amenazas para la seguridad del Estado, el propio interesado, deben poder acceder al conjunto de motivos y pruebas sobre cuya base se ha adoptado esta resolución (véase, por analogía, la sentencia de 4 de junio de 2013, *ZZ*, C-300/11, EU:C:2013:363, apartados 54 a 59), incluidos los criterios de evaluación predeterminados y el funcionamiento de los programas que aplican estos criterios.”

El sistema implantado por la Directiva en examen permite satisfacer estas exigencias, por lo que el TJUE considera que no se ha puesto de manifiesto ningún elemento que pueda afectar a su validez (§ 228). Por lo que *sensu contrario*, cuando un sistema de IA no permita satisfacer estos requisitos será contrario a la Carta Europea de Derechos Humanos, y consiguientemente, inválido.

3.4.4.- Consecuencia: prohibición de algoritmos de caja negra en la toma de decisiones por el sector público. ¿Posibilidad de articular en el futuro una manera distinta de motivar, cuando se trata de decisiones sustentadas en IA?

De lo antedicho debemos extraer otra consecuencia adicional: no solo serán inválidos los sistemas de IA que no ofrezcan motivación suficiente, sino que, además, no podrán implantarse sistemas de IA basados en tecnologías que no permitan explicar su funcionamiento. Como afirman KEARNS y ROTH⁹⁰, en relación con la interpretabilidad de los algoritmos –correlativamente añadiríamos nosotros: y su motivación– hay que situarse “en el intervalo entre el «no se permiten algoritmos» y el «todo vale»”.

Existen sistemas de IA programados de forma que evolucionan por sí mismos, de manera que los humanos somos incapaces de comprender o determinar por qué llegaron a una determinada conclusión o tomaron una decisión concreta. Son algoritmos *inteligentes*, que evoluciona en sus decisiones mediante procesos que ignoramos y no somos capaces de explicar. Siendo así, parece claro que no nos resulta cabalmente posible ofrecer una motivación suficiente de los actos administrativos que se basen en estos sistemas.

El caso se puede producir en diferentes tecnologías de inteligencia artificial conocidas como de caja negra (*black box*), que no serían admisibles en cuanto que no resultan explicables⁹¹: en particular, las

⁹⁰ Cfr. KEARNS, M. y ROTH, A. (2020): *El algoritmo ético, op.cit.*, pág.237. En línea con lo que expresamos en el texto, los autores consideran que no parece haber una metodología prometedora sobre la interpretabilidad de las redes neuronales profundas (pág.240).

⁹¹ Véase el Capítulo I de esta obra; especialmente, § 5.2. Y las específicas aportaciones en esta materia de CERRILLO MARTÍNEZ, A. (2019): “Com obrir les caixes negres de les Administracions públiques? Transparència i rendició de comptes en l’ús dels algorismes”, *Revista Catalana de Derecho Público*, n.º 58, págs. 13 ss.; y RUDIN, C. (2019): “Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead”, *Nature Machine Intelligence*, vol.1, págs.206 ss. Por su parte, el Tribunal de Cuentas de Países Bajos también se opone a los algoritmos de caja negra, llegando a afirmar que debería ser más fácil verificar los resultados de un algoritmo que los de una persona humana; véase (2021), *Understanding algorithms*, Netherlands Court of Audit, pág. 36.

redes neuronales profundas (*deep learning*), es decir, sistemas que aprenden por sí solos y cuyo proceso lógico realmente ignoramos, sin que lleguemos a saber por qué optaron por una determinada decisión. Los algoritmos que nos plantean estos problemas de interpretabilidad son conocidos como algoritmos de caja negra. Sobre ellos sostiene KOENE que no resulta posible determinar si la decisión que tomó el algoritmo es justa o no⁹².

En el caso *Ligue des droits humains*, citado (sentencia de 21 de junio de 2022), el TJUE destaca la imposibilidad de utilizar de sistemas de aprendizaje automático impuesta sobre el PNR como aval de su adecuación a Derecho, es decir, que no aceptaría el empleo de esta tecnología, pero al descartarse su uso en este concreto caso, el sistema resulta conforme a Derecho:

“(194): En cuanto a los criterios a los que la UIP [Unidad de Información sobre los Pasajeros] puede recurrir a estos efectos, es preciso señalar, en primer lugar, que, según los propios términos del artículo 6, apartado 3, letra b), de la Directiva PNR, debe tratarse de criterios «predeterminados». Como observa el Abogado General en el punto 228 de sus conclusiones, este requisito se opone a la utilización de las tecnologías de inteligencia artificial en el marco de sistemas de autoaprendizaje (*machine learning*), que puedan alterar, sin intervención y sin control humanos, el proceso de evaluación y, en particular, los criterios de evaluación en los que se basa el resultado de la aplicación del proceso, así como la ponderación de dichos criterios”.

Para defender la aplicación de estos algoritmos podría afirmarse que siempre sería posible conocer y difundir pública y conscientemente la programación del proceso, parametrizada (esta sí) con criterios objetivos y transparentes, y aceptando que la decisión (finalmente impredecible y por ello no completamente explicable) la tome la máquina. Es decir, que la motivación en este caso consistiría en explicar detalladamente el proceso y reconocer abiertamente su impredecibi-

⁹² Ansgar KOENE, de la Universidad de Nottingham, es ingeniero y doctor en física y neurociencia computacional, especializado en inteligencia artificial. La afirmación se encuentra en la entrevista que le realizó *Digital Future Society* el 07/02/2020, accesible aquí: <https://digitalfuturesociety.com/es/qanda/ansgar-koene-y-los-peligros-de-los-algoritmos-de-caja-negra-en-el-mercado-laboral/> (última visita: 28/03/2023).

lidad, pero aceptando como premisa de partida que la programación del algoritmo es acertada. Ciertamente, a la vista de la particularidad que representa la IA, podría articularse otra metodología a efectos de motivación de las decisiones que se adopten aplicando esta tecnología, en el sentido anteriormente expuesto. Por ejemplo, asegurando que el proceso cumple, al menos, 4 condiciones:

1. Preservando la adecuada selección y depuración de los datos con los que se nutre el sistema.
2. Explicando con transparencia qué tipo de tecnología se va a aplicar: redes neuronales profundas, etc.
3. Acreditando la realización de los medios de control, supervisión y auditoría de que deben disponer estos sistemas, lo que permitiría crear cierta confianza *-iuris tantum-* en que las decisiones adoptadas por el sistema son conforme a Derecho.
4. Aceptando el resultado por asentarse en esas premisas predefinidas, aun cuando no sepamos exactamente el porqué de la decisión.

Esta modulación o excepción a la metodología propia de la motivación en lenguaje natural habría de ser expresamente regulada para resultar admisible. Mientras que no se lleve a cabo esta regulación específica, los requisitos generales de la motivación son plenamente aplicables a las decisiones adoptadas mediante IA, como también sostiene VALERO⁹³.

Por otra parte, esta regulación específica podría superar el listón de la motivación, pero no rebasaría otros diferentes, como el de la objetividad de la decisión: ¿cómo podemos afirmar que es objetiva una decisión que no sabemos *por qué* se toma, aunque sepamos *cómo* se toma? Desde el plano técnico se puede objetar que cabe evitar los sesgos discriminatorios mediante una adecuada selección de los datos utilizados para la programación del sistema: si esos bancos de datos no incurrían en sesgos, tampoco lo haría el sistema que los utilice. El problema, como ya indiqué más atrás, es que las máquinas no razonan como los humanos, y criterios de programación que no tendrían

⁹³ En VALERO TORRIJOS, J. (2019): “Las garantías...”, *op.cit.*, pág.87.

que parecemos sesgados son interpretados por el sistema induciendo a tales sesgos. Si la programación es transparente, podremos corregir el sesgo; pero si se articula mediante algoritmos de caja negra, no parece posible hacerlo, o cuando menos, sólo se podría corregir la desviación tras un período suficiente de aplicación del sistema y posterior auditoría y supervisión de sus resultados.

3.5.- *Inteligencia artificial y discrecionalidad administrativa*

¿Pueden adoptarse actos administrativos discrecionales mediante IA? ¿O esta posibilidad está vetada por el Derecho administrativo? Esta misma pregunta también se ha planteado en relación con la actuación administrativa automatizada, y las reflexiones son extrapolables a uno y otro caso de uso.

Hay autores que consideran que la discrecionalidad administrativa es incompatible con la actuación automatizada –y, por consiguiente, también con la IA–, en la medida que exige juicios de valor, incompatibles con sistemas cuyo funcionamiento resulta ajeno a la voluntad humana, por lo que solo cabría implantarlos para adoptar actos reglados⁹⁴. En cambio, otros autores no encuentran objeción a que mediante actuación administrativa automatizada se adopten actos discrecionales: por ejemplo, MARTÍN DELGADO⁹⁵ considera que cabe aceptar estos sistemas cuando se trata de discrecionalidad administrativa técnica, o “de baja intensidad”; en tanto que CERRILLO entiende que la cuestión no estriba en proscribir radicalmente la adopción de actos administrativos de manera automatizada en ámbitos de discrecionalidad administrativa, sino en determinar si la

⁹⁴ Así, VALERO TORRIJOS, J. (2007): *El régimen jurídico de la e-Administración*, Comares, Granada, pág.75; URIOS APARISI, X. y ALAMILLO DOMINGO, I. (2011): *La actuación administrativa automatizada...*, *op.cit.*, págs. 94 ss.; BAUZÁ MARTORELL, FJ (2017): “Identificación, autenticación...”, *op.cit.*, pág 740; PONCE SOLÉ (2019): “Inteligencia artificial, Derecho administrativo y reserva de Humanidad: algoritmos y procedimiento administrativo debido tecnológico”, *Revista General de Derecho Administrativo*, n.º 50, págs.28 ss.; véase también lo que expone el mismo autor en el Capítulo V de esta obra; y TERRÓN SANTOS, D. y DOMÍNGUEZ ÁLVAREZ, J.L. (2022): *i-Administración pública...*, *op.cit.*, pág.109.

⁹⁵ Véase MARTÍN DELGADO, I. (2009): “Naturaleza...”, *op.cit.*, págs. 369 ss.,

decisión –discrecional– será o no de mayor calidad empleando la actuación automatizada en lugar de la decisión humana⁹⁶. Por su parte, HUERGO considera que no toda solución de IA supone actuación automatizada, y así como considera que esta última no puede implantarse en ámbitos de discrecionalidad, su posición es más matizada por lo que se refiere a la IA⁹⁷.

Algunas disposiciones se han hecho los posicionamientos restrictivos, acogiendo limitaciones al uso de sistemas automatizados y limitándolos, por lo general, a actuaciones de comprobación y de carácter más mecánico que intelectual. Así, el art.44 de la Ley 26/2010, de 3 de agosto, de régimen jurídico y de procedimiento de las administraciones públicas de Cataluña⁹⁸; el art.40 del Decreto 622/2019, de 27 de diciembre, de administración electrónica, simplificación de procedimientos y racionalización organizativa de la Junta de Andalucía⁹⁹; y

⁹⁶ CERRILLO MARTÍNEZ, A. (2022): “Actividad administrativa automatizada...”, *op.cit.*, págs. 266 ss., y del mismo autor, “¿Son fiables las decisiones de las Administraciones públicas tomadas por algoritmos?”, *European Review of digital administration & law*, vol.1, n.º1-2, págs. 18 ss.

⁹⁷ Véase HUERGO LORA, A. (2021): “Administraciones...”, *op.cit.*, págs. 14 ss.

⁹⁸ Que, bajo la rúbrica *Actuación administrativa automatizada*, dispone: “1. Las administraciones públicas catalanas pueden realizar actuaciones automatizadas para constatar la concurrencia de los requisitos que establece el ordenamiento jurídico, declarar las consecuencias previstas, adoptar las resoluciones y comunicar o certificar los datos, actos, resoluciones o acuerdos que consten en sus sistemas de información, mediante la utilización del sistema de firma electrónica que determinen. 2. Sólo son susceptibles de actuación administrativa automatizada los actos que puedan adoptarse con una programación basada en criterios y parámetros objetivos”.

⁹⁹ Con la misma rúbrica, este es el precepto más minucioso de cuantos se han dictado en lo relativo a la delimitación precisa de las actuaciones que pueden llevarse a cabo por medios automatizados. Dispone lo siguiente:

“1. Se promoverá activamente la actuación administrativa automatizada en actividades que puedan producirse mediante un sistema de información adecuadamente programado sin necesidad de intervención directa de una persona empleada pública en cada caso singular, y especialmente las que consistan en:

a) La adopción de un acuerdo o decisión administrativa mediante la aplicación de fórmulas matemáticas y otros procesos puramente mecánicos en los que se utilicen valores cuantificables y susceptibles de ser expresados en cifras y porcentajes.

b) La certificación de hechos o datos preexistentes en registros o en sistemas de información, incluso del silencio administrativo.

el art.43 de la Ley 5/2021, de 29 de junio, de Organización y Régimen Jurídico del Sector Público Autonómico de Aragón¹⁰⁰.

En mi opinión, fuera de los ámbitos expresamente vetados por el Derecho positivo no existe una objeción de carácter dogmático que impida adoptar decisiones automatizadas en ejercicio de potestades administrativas discrecionales. Con arreglo a los postulados que regentan la cuestión –no solo clásicos, sino unívocos, por lo que resulta innecesaria su cita–, la discrecionalidad concede a la Administración un margen de apreciación para adaptarse a las circunstancias del caso concreto. Es perfectamente defendible que esa adaptación a las circunstancias tenga lugar mediante el desarrollo de un sistema algorítmico, de decisión automatizada y, finalmente, de IA, si las circunstancias concurrentes y la parametrización del sistema determinan que ese sea el mejor modo de ejercer la discrecionalidad en un concreto ámbito. Ahora bien, es igualmente pacífico que discrecionalidad no equivale a arbitrariedad, así como que los actos discrecionales han de motivarse –art.35.1 i) LPAC –, por lo que la Administración debe explicar razonadamente por qué considera que en ese ámbito es preferible un sistema automatizado o de IA, y ser capaz además de motivar

c) La constatación puramente mecánica de requisitos previstos en la normativa aplicable y la posterior declaración, en su caso, de la consecuencia jurídica prevista en la misma.

d) La comunicación o declaración de un hecho, acto o acuerdo preexistente a través de su transcripción total o parcial.

e) La práctica de las notificaciones electrónicas.

2. No cabrá realizar mediante actuación administrativa automatizada actividades que supongan juicios de valor”.

¹⁰⁰ También bajo la rúbrica *Actuación administrativa automatizada*, y tras dos apartados que reproducen o concretan lo establecido en el Derecho estatal básico, este precepto dispone: “3. No se considerará actuación administrativa automatizada cuando sea necesario aplicar criterios subjetivos de decisión, bien de forma individualizada o colectiva, ni la ejecución de procesos fragmentados dentro de un mismo acto o actuación administrativa. 4. El departamento con competencias en materia de administración electrónica definirá el marco al que deben ajustarse los órganos de la Administración de la Comunidad Autónoma de Aragón en el desarrollo de actuaciones administrativas automatizadas. Reglamentariamente, identificará los criterios de modelización de algoritmos, de publicación de los árboles de decisión y los códigos fuente, de almacenaje, de identificación de los responsables, así como los criterios éticos de actuación en el tratamiento de la información”.

cada decisión que se adopte en su aplicación. En relación con la motivación, remito a lo expuesto más arriba. Y en cuanto a la idoneidad del sistema, suscribo la opinión de CERRILLO¹⁰¹, a saber:

“Podemos observar que los medios electrónicos y, en particular, la inteligencia artificial, pueden ser mucho más eficientes que las personas para tomar decisiones de calidad en entornos de baja complejidad incluso donde haya cierta incertidumbre. En cambio, cuando los entornos en los que deben adoptarse las decisiones sean complejos, el nivel de incertidumbre sea alto o los datos disponibles sean escasos o de baja calidad, los medios electrónicos pueden tener un menor papel en la toma de decisiones, debiendo atribuirse a personas. Asimismo, si la decisión exige desarrollar proceso analíticos puede ser eficiente el uso de la tecnología sustituyendo a las personas, mientras que si comporta la realización de un proceso racional complejo, como realizar juicios o ponderaciones, no lo será. En última instancia, hoy en día, en la medida en que los algoritmos carecen de empatía o sentido común, que pueden ser fundamentales en la toma de decisiones públicas, la toma de decisiones deberá reservarse a las personas si ello es fundamental para garantizar su calidad.

De este modo, la respuesta a la pregunta de si las Administraciones públicas pueden automatizar la toma de decisiones discrecionales, que debe estar necesariamente presidida por el principio de precaución, pasa por analizar el tipo de decisión que debe adoptarse, el entorno en el que deba adoptarse, la disponibilidad de datos y su calidad y la necesidad de estar dotado de determinadas capacidades personales para poder adoptar una decisión de calidad. Junto a este análisis deberá conocerse el nivel de desarrollo de la tecnología para poder contrastar su adecuación a lo anterior”.

Como colofón de lo antedicho, yo recordaría que, cuando el resultado deparado por el sistema de IA en cualquier entorno de decisión –discrecional o no– sea desfavorable, deben preservarse garantías como el derecho de audiencia; y que la reserva de humanidad, destacada por PONCE en su capítulo de esta obra, y por CERRILLO en el pasaje transcrito, invitan a que, en tales casos de decisiones desfavorables,

¹⁰¹ Cfr. CERRILLO MARTÍNEZ, A. (2022): “Actividad administrativa automatizada...”, *op.cit.*, pág.267.

estos sistemas sean un apoyo a la decisión, pero que no lleguen a adoptar la decisión definitiva del procedimiento, reservándose ese ámbito para personas humanas que puedan desplegar elementos de valoración que, al menos hasta la fecha, escapan a las posibilidades de la tecnología.

4.- FINAL: LA INACEPTABLE TENSION ENTRE PROGRESO TECNOLÓGICO Y GARANTÍAS DE LA CIUDADANÍA; LUCES Y SOMBRAS DE LA REGULACIÓN ESPECÍFICA DE LA IA EN EL SECTOR PÚBLICO

El indudable desarrollo de la IA está suponiendo una extraordinaria tensión entre intereses enfrentados. Por una parte, la industria del sector y los científicos *tecnólogos* vienen destacando las bondades de la tecnología y su importancia para el desarrollo económico y social. Cuando se les exponen reparos jurídicos, alegan, por una parte, que el argumento no debería detener el avance científico, como si en las sociedades avanzadas cualquier nueva tecnología signifique necesariamente un avance (no es preciso recordar el armamento atómico: también las industrias contaminantes o las de excesivo consumo energético); y por otra, que si los europeos somos tan puristas con la protección de los derechos, otros actores mundiales, como EE.UU., Rusia o China, nos sobrepasarán en el desarrollo de estas tecnologías generando una frustración de desarrollo económico para Europa¹⁰².

Me parece inaudito que, a estas alturas de la civilización, sigamos planteándonos debates como éste. La protección de los derechos fundamentales debe estar por encima de toda duda, no solo en Europa, sino en todas las sociedades avanzadas. El hecho de que las advertencias sobre la degradación de derechos que puede conllevar la aplicación descontrolada de la IA se hayan oído con especial rotundidad en Europa –por parte de la doctrina jurídica y de la actuación judicial–, y con menos intensidad, en EE.UU., es sintomático de las prioridades

¹⁰² Un interesante análisis comparado de los modelos europeo y norteamericano en MÖKANDER, J. y otros (2022): “The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: what can they learn from each other?”, *Mind and Machines* (Springer), publicación en línea 18/08/2022, <https://doi.org/10.1007/s11023-022-09612-y>

que los grandes modelos de la civilización contemporánea vienen auspiciando, y que se encuentra, indudablemente, en contraposición de valores.

Nadie se plantea en la actualidad que una nueva tecnología de producción alimentaria permita contaminar los acuíferos o emitir un volumen excesivo de gases contaminantes a la atmósfera. Ese mismo planteamiento de partida debe ser el que acompañe a los desarrollos de IA en relación con los derechos fundamentales, especialmente cuando se trata de sistemas implantados para la toma de decisiones en el sector público.

Como se puede comprobar a lo largo de esta obra, hay al menos tres derechos fundamentales que se pueden ver afectados en tales casos: *a)* El derecho a la igualdad, sin que pueda prevalecer discriminación –art.14 CE–, que se puede poner en entredicho cuando los modelos de IA no están bien diseñados e incurrir en sesgos de alcance discriminatorio; *b)* El derecho a la protección de datos –art.18 CE–, susceptible de vulnerarse cuando el entrenamiento de los modelos de IA se sustenta en tratamientos ilícitos de datos; *c)* El derecho a la tutela judicial efectiva, sin que pueda producirse indefensión –art.24 CE–, que tiene también una vertiente en el procedimiento administrativo determinante de un vicio de anulabilidad –art.48.2 LPAC– que llega a ser de nulidad –art.47.1 a) LPAC– cuando el procedimiento afecta a derechos fundamentales. Son, por tanto, gravísimos los daños que se pueden infligir a los derechos fundamentales si los modelos de IA no son adecuadamente diseñados, auditados, controlados y revisados, siguiendo las metodologías que hemos expuesto en este trabajo.

A la situación actual, en la que, con un mínimo esfuerzo –como el que hemos llevado a cabo en esta obra– se le puede concretar un determinado marco jurídico de actuación, con pautas y determinaciones a seguir para lograr la adecuación a Derecho de los modelos de IA, se le deben añadir las propuestas regulatorias en camino, y en especial, la que lidera la UE, que se expone en los Capítulos XVII y XVIII de esta obra, a los que remito para conocer los pormenores de esa iniciativa normativa¹⁰³.

¹⁰³ Adicionalmente puede verse HERNÁNDEZ PEÑA, J.C. (2022): *El marco jurídico de la inteligencia artificial. Principios, procedimientos y estructuras de gober-*

Como he dicho en otro lugar¹⁰⁴, el enfoque europeo sobre la IA supone indudablemente un hándicap al desarrollo de la tecnología y un coste añadido para su puesta en mercado. Imponer limitaciones reduce las potencialidades de los modelos, y exigir requisitos determina costes de cumplimiento. Los desarrolladores de IA asiáticos y norteamericanos, al no encontrarse limitados por estas circunstancias, pueden implementar modelos de IA que aprendan más rápidamente y a menor coste, lo cual representa una ventaja competitiva. Pero es una carga que Europa ha estado siempre dispuesta a pagar, no solo en materia de IA, sino en otros muchos ámbitos. Representa una línea de pensamiento profundamente imbricada en los valores europeos y que, desde la Escuela de Salamanca, entronca indudablemente con el humanismo cristiano, que sitúa al individuo en el centro de la escena y valora su dignidad como valor supremo del Ordenamiento jurídico: de hecho, todo el sistema constitucional en los países europeos está centrado en el respeto y protección de los derechos fundamentales de la persona frente a cualesquiera injerencias o potenciales lesiones. La intención –expresada de la iniciativa regulatoria europea es que, al tratarse frecuentemente de modelos de comercialización mundial, los requisitos impuestos para la implantación de desarrollos de IA en Europa conduzcan a un cumplimiento universal de tales requisitos, al menos cuando se trate de esos sistemas tan potentes –ChatGPT, sistemas de conducción automática de vehículos, etc.– y distribuidos por todo el orbe.

Como he apuntado más atrás, de igual forma que Europa no se plantea modificar sus políticas en materia de medio ambiente o de derechos laborales – aunque ello represente una desventaja competitiva para las industrias europeas respecto a las del resto del mundo y haya conducido a la deslocalización de muchas fábricas–, sino, antes al contrario, profundizar en ellas, no cabe dar un paso atrás en la protección de los derechos fundamentales y las garantías del procedimiento administrativo ante la implantación de sistemas de IA que sustenten la toma de decisiones.

nanza, Aranzadi-Thomson-Reuters, Cizur Menor (Navarra), págs.76-207.

¹⁰⁴ GAMERO CASADO, E. (2021): “El enfoque europeo de inteligencia artificial”, *Revista de Derecho Administrativo*, n.º 20, págs.286 ss.

Para concluir, debo advertir un nuevo riesgo: el de pensar que, al aprobarse una norma específica sobre IA, para que los modelos desarrollados sean lícitos les basta con cumplir con lo establecido en esa disposición. Si analizamos con detalle la propuesta europea, veremos que no siempre rebasa los estándares que hemos predicado en este trabajo, o dicho de otro modo, que no impone expresamente el cumplimiento de ciertas garantías que hemos exigido aquí, cuando se trata de sistemas que sustentan la toma de decisiones en un procedimiento administrativo. Pues bien, recordemos lo obvio: la ley especial –reglamento europeo– se aplica preferentemente a la general, desplazándola cuando haya contradicción o antinomia; pero mientras que no exista colisión, los preceptos de la ley general se aplican también a los supuestos de hecho regulados por la ley especial. De tal manera que, los requisitos como el derecho a subsanar y a la audiencia, o el deber de motivar los actos administrativos, son predicables de actuaciones en las que se aplique la IA como soporte a la toma de decisiones en procedimientos administrativos, aunque el reglamento europeo no lo establezca expresamente.

BIBLIOGRAFÍA CITADA

- ALAMILLO DOMINGO, I. (2017): Comentarios a los artículos 41 y 42, en CAMPOS ACUÑA, C. (Dir.), *Comentarios a la Ley 40/2015, de Régimen Jurídico del Sector Público*, Wolters Kluwer, Madrid, págs. 222-231.
- BAÑO LEÓN, J.M. y LAVILLA RUBIRA, J.J. (Dirs.) (2021): *Comentarios al procedimiento administrativo*, Tirant lo Blanch, Valencia.
- BAUZÁ MARTORELL, F.J (2017): “Identificación, autenticación y actuación automatizada de las Administraciones públicas”, en GAMERO CASADO, E. (Dir.), *Tratado de procedimiento administrativo común y régimen jurídico básico del sector público*, vol.1, Tirant lo Blanch, Valencia, págs.787 ss.
- BERGHOFF, C. y otros (2021): *Towards Autitable AI Systems. Current status and future directions*, Whitepaper, Federal Office for Information Security-Fraunhofer Forum.
- BERNING PRIETO, A.D. (2019): *Validez e invalidez de los actos administrativos en soporte electrónico*, Aranzadi-Thomson-Reuters, Cizur Menor (Navarra).
- BONE, J. (2020): “Auditoría de Inteligencia Artificial”, disponible en <https://www.intedya.com>

- CAMPOS ACUÑA, M.C. (coord.) (2021): *Comentarios a la Ley 39/2015, de procedimiento administrativo común de las Administraciones públicas*, La Ley-El Consultor, Madrid.
- CANALS AMETLLER, D. (2021): “El ejercicio de potestades administrativas por operadores privados en régimen de mercado”, en E. Gamero Casado (Dir.): *La potestad administrativa. Concepto y alcance práctico de un criterio clave para la aplicación del Derecho administrativo*, Tirant lo Blanch, Valencia, págs. 320-387.
- CAPDEFERRO VILLAGRASA, O. (2023): “Las obligaciones sujetas a control administrativo automatizado”, *Revista de Internet, Derecho y Política*, n.º 37.
- CARRILLO PARRILLA, J.A. (2021): “La discriminación a través del algoritmo”, Observatorio Pagoda, 11/03/2021, accesible en <https://www.uv.es/fatwireed/catedra-pagoda/es/novedades-1286053802801/Novetat.html?id=1286182093538> (última visita: 28/03/2023).
- CERRILLO MARTÍNEZ (2019): “El impacto de la inteligencia artificial en el Derecho administrativo: ¿nuevos conceptos para nuevas realidades técnicas?”, *Revista General de Derecho Administrativo*, n.º 50.
- (2019): “Com obrir les caixes negres de les Administracions públiques? Transparència i rendició de comptes en l'ús dels algorismes”, *Revista Catalana de Derecho Público*, n.º 58, 13-28.
- (2020): “Automatización e inteligencia artificial”, en MARTÍN DELGADO, I. (Dir.), *El procedimiento administrativo y el régimen jurídico de la Administración pública desde la perspectiva de la innovación tecnológica*, Iustel-IVAP, Madrid, págs.389-415.
- (2020): “¿Son fiables las decisiones de las Administraciones públicas tomadas por algoritmos?”, *European Review of digital administration & law*, vol.1, n.º1-2, págs.18-36
- (2021): “Robots, asistentes virtuales y automatización de las administraciones públicas”, *Revista Galega de Administración Pública*, n.º 61, págs. 271 sigs.
- (2022): “Actividad administrativa automatizada y utilización de algoritmos”, en VV.AA., *Las políticas de buen gobierno en Andalucía (I): Digitalización y transparencia*, Instituto Andaluz de Administración Pública, Sevilla, págs.259-287.
- COTINO HUESO, L. y CASTELLANOS CLARAMUNT, J. (Eds.) (2022). *Transparencia y explicabilidad de la inteligencia artificial*, Tirant lo Blanch, Valencia.
- CRIADO, I; VALERO, J. y VILLODRE, J. (2020) “Algorithmic transparency and bureaucratic discretion: The case of SALER early warning system”, *Information Polity*, vol.25, n.º4, PÁGS 449-470.

- DE LAS HERAS BALLELL, T. (2022): *Guiding Principles for Automated Decision-Making in the EU*, European Law Institute, Vienna.
- DE LA SIERRA MORÓN, S. (2020): “Inteligencia artificial y justicia administrativa: una aproximación desde la teoría del control de la Administración pública”, *Revista General de Derecho Administrativo*, n.º 53, 2020.
- (2021): “Control judicial de los algoritmos: robots, Administración y Estado de derecho”, Tribuna publicada en *ElDerecho.com*, el 11/06/2021, apartado de conclusiones; accesible en <https://elderecho.com/control-judicial-de-los-algoritmos-robots-administracion-y-estado-de-derecho> (última visita: 28/03/2023).
- EZEANI, G.; KOENE, A.; KUMAR, R.; SANTIAGO, N.; y WRIGHT, D. (2021): *A survey of artificial intelligence risk assessment methodologies. The global state of play and leading practices identified*, Ernst&Young-Trilateral Research.
- FLORIDI, L. y otros (2022): *capAI. A procedure for conducting conformity assessment of AI systems in line with the EU Artificial Intelligence Act*, University of Oxford, Oxford.
- GALÁN, C. (2019): “The certification as a mechanism for control of artificial intelligence in Europe”, accesible en <https://ec.europa.eu/futurium/en/european-ai-alliance/certification-mechanism-control-artificial-intelligence-europe-0.html>
- GALETTA, D.U. (2020): “Algoritmi, procedimento amministrativo e garanzie: brevi riflessioni, anche alla luce degli ultimi arresti giurisprudenziali in materia”, *Rivista Italiana di Diritto Pubblico Comunitario*, n.º 4/2020, 501-516.
- GALLARDO CASTILLO, M. J. (2016): *Los procedimientos administrativos en la ley 39/2015: análisis y valoración de la reforma*, Tecnos, Madrid.
- GAMERO CASADO, E. (2021): “Necesidad de motivación e invalidez de los actos administrativos sustentados en inteligencia artificial o en algoritmos”, Almacén de Derecho, <https://almacenederecho.org>
- (2021): “Compliance (o cumplimiento normativo) de desarrollos de inteligencia artificial para la toma de decisiones administrativas”, *Diario La Ley*, n.º 50, Sección Ciberderecho.
- (2021): “El enfoque europeo de inteligencia artificial”, *Revista de Derecho Administrativo*, n.º 20, 2021, 268-288.
- (2022): “Reflexiones introductorias: de la administración electrónica a la digital (o la historia interminable)”, en A. Cerrillo Martínez (dir.), *La Administración Digital*, Dykinson, Madrid, págs. 33-50.
- GAMERO CASADO, E. (Dir.); FERNÁNDEZ RAMOS, S. y VALERO TORRIJOS, J. (Coords.) (2017): *Tratado de procedimiento administrativo*

- común y régimen jurídico básico del sector público*, 2 vols., Tirant lo Blanch, Valencia.
- GEORGIEVA, I.; TIMAN, T. y HOEKSTRA, M. (2022): *Regulatory divergences in the draft AI Act. Differences in public and private sector obligations*, European Parliamentary Research Service, Brussels.
- GÓMEZ PADILLA, R. (2016): “Comentario al art.41 LRJSP”, en RECUE-DA GIRELA, M.A. (Dir.), *Régimen jurídico del sector público y procedimiento administrativo común*, Aranzadi-Thomson-Reuters, Cizur Menor (Navarra), págs.1.222 ss.
- GOSÁLBEZ PEQUEÑO, H. (Dir.) (2016): *La nueva Ley del procedimiento administrativo común*, Wolters Kluwer, Madrid.
- HERNÁNDEZ PEÑA, J.C. (2022): *El marco jurídico de la inteligencia artificial. Principios, procedimientos y estructuras de gobernanza*, Aranzadi-Thomson-Reuters, Cizur Menor (Navarra).
- HUERGO LORA, A. (2020): “Una aproximación a los algoritmos desde el Derecho administrativo”, en la obra dirigida por él mismo, *La regulación de los algoritmos*, Aranzadi-Thomson-Reuters, Cizur Menor (Navarra), págs.23-87.
- HUERGO LORA, A. (2021): “Administraciones públicas e inteligencia artificial: ¿más o menos discrecionalidad?”, *El Cronista del Estado Social y Democrático de Derecho*, n.º 96-97 (hemos manejado la versión accesible en la sección “Estudios y comentarios” del INAP, <https://laadministracionaldia.inap.es/noticia.asp?id=1512494>)
- JOSHI, D. y otros (2021): *Algorithmic accountability for the public sector. Learning from the first wave of policy implementation*, Ada Lovelace Institute, AI Now Institute and Open Government Partnership, accessible en <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/> (última visita: 28/03/2023).
- KAMINSKI, M.E. y URBAN, J.M. (2021): “The right to contest AI”, *Columbia Law Review*, vol.121, n.º 7, 1.957-2.048, disponible en <https://ssrn.com/abstract=3965041> (última visita: 28/03/2023)
- KAZIM, E., MENDES, D., DENNY, T. y KOSHIYAMA, A. (2021): “AI auditing and impact assessment: according to the UK information commissioner’s office”, *AI Ethics*, <https://doi.org/10.1007/s43681-021-00039-2>, última visita 28 de marzo de 2023.
- KEARNS, M. y ROTH, A. (2020): *El algoritmo ético. La ciencia del diseño de algoritmos socialmente responsables*, LaLey-Wolters Kluwer, Madrid.
- LESLIE, D. (2019): *Understanding artificial intelligence ethics and safety. A guide for the responsible design and implementation of AI*, The Alan Turing Institute (<https://doi.org/10.5281/zenodo.3240529>).
- MALGIERI, G. (2019): “Automated decision-making in the EU Member States: The right to explanation and other *suitable safeguards* in the national

- legislation”, *Computer Law & Security Review*, n.º 35 (<https://dx.doi.org/10.2139/ssrn.3233611>).
- MANTELERO, A. (2022): *Beyond Data. Human Rights, Ethical and Social Impact Assessment in Artificial Intelligence*, Asser Press–Springer, The Hague.
- MANZONI M., MEDAGLIA R., TANGI L., VAN NOORDT C., VACCARI L., y GATTWINKEL D. (2022): *AI Watch. Road to the Adoption of Artificial Intelligence by the Public Sector*, JRC-European Commission, Luxembourg.
- MARTÍN DELGADO, I. (2009): “Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada”, *Revista de Administración Pública*, n.º 180, págs.353-386.
- MARTÍNEZ MARTÍNEZ, R. (2019): “inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo”, *Revista Catalana de Derecho Público*, n.º 58, 64-81.
- MARTÍNEZ GUTIÉRREZ, R. (2016): *El régimen jurídico del nuevo procedimiento administrativo común*, Aranzadi, Cizur Menor (Navarra).
- MERCHÁN ARRIBAS, M. (2020): *Guía de uso de la Inteligencia Artificial en el Sector Público*, accesible en <https://www.clubdeinnovacion.es/guia-uso-la-ia-sector-publico/> (última visita: 28/03/2023).
- MERLI, F. (2022): “Automated Decision Making Systems in national Administrative Law: An Austrian Perspective”, Informe presentado en el seminario internacional celebrado en la Universidad de Friburgo Proyecto INDIGO) los días 29 y 30 de septiembre de 2022.
- (2023): “Automated Decision-Making Systems in Austrian Administrative Law”, *Rivista Interdisciplinare sul Diritto delle Amministrazioni Pubbliche*, n.º 1/2023, págs.41-50.
- MÖKANDER, J. y otros (2022): “The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: what can they learn from each other?”, *Mind and Machines* (Springer), publicación en línea 18/08/2022, <https://doi.org/10.1007/s11023-022-09612-y>
- MORAL SORIANO, L. (2022): “Decisiones automatizadas, Derecho administrativo y argumentación jurídica”, en F. LLANO ALONSO, (Dir.), *Inteligencia artificial y Filosofía del Derecho*, Laborum Ediciones, Murcia, 475-499.
- MOSS, E.; WATKINS E.A.; METCALF, J.; ELISH, M.C. (2020): “Governing with Algorithmic Impact Assessments: Six Observations”, *Conference on Artificial Intelligence, Ethics, and Society (AIES)*, Disponible en <https://papers.ssrn.com>
- NAVARRO GONZÁLEZ, R. (2017): *La motivación de los actos administrativos*, Aranzadi-Thomson-Reuters, Cizur Menor (Navarra).

- PALOMAR OLMEDA, A. (2022): *Procedimiento administrativo*, 3ª ed., Aranzadi-Thomson-Reuters, Cizur Menor (Navarra).
- RECUERDA GIRELA, M.A. (Dir.) (2016): *Procedimiento Administrativo Común de las Administraciones Públicas*, Aranzadi.
- REICHEL, J. (2023): “Regulating Automation of Swedish Public Administration”, *Rivista Interdisciplinare sul Diritto delle Amministrazioni Pubbliche*, n.º 1/2023, págs.75-94.
- REISMAN, D.; SCHULTZ, J.; CRAWFORD, K.; y WHITAKER, M. (2018): *Algorithmic Impact Assessments: a practical framework for public agency accountability*, AINow Institute (University of New York).
- RIVERO GONZÁLEZ, M. (Dir.) (2016): *Procedimiento Administrativo Común y Régimen Jurídico*, Comares, Granada.
- RUDIN, C. (2019): “Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead”, *Nature Machine Intelligence*, vol.1, 206-215.
- SORIANO ARNANZ, A. (2021): “Decisiones automatizadas: problemas y soluciones jurídicas. Más allá de la protección de datos”, *Revista de Derecho Público: Teoría y Método*, vol. 3, págs. 85-127.
- (2021b): “Decisiones automatizadas y discriminación: aproximación y propuestas generales”, *Revista General de Derecho Administrativo*, n.º 56.
- SCHWARTZ, R. y otros (2022): *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence*, National Institute of Standards and Technology (United States Department of Commerce), accessible en <https://doi.org/10.6028/NIST.SP.1270> (última visita: 28/03/2023).
- TARRES VIVES, M. (2012): “La actuación administrativa automatizada”, en J. TORNOS MAS, *La Ley 26/2010, del 3 de agosto, de régimen jurídico y procedimiento de las Administraciones públicas de Cataluña*, Iustel, Madrid.
- TERRÓN SANTOS, D. (2022): *Administración inteligente y automática. Una visión más allá del algoritmo*, Colex, Madrid.
- TERRÓN SANTOS, D. y DOMÍNGUEZ ÁLVAREZ, J.L. (2022): *i-Administración pública, sistemas algorítmicos y protección de datos*, Iustel, Madrid.
- TORNOS MAS, J. (Dir.) (2017): *Estudios sobre las leyes 39/2015 del procedimiento administrativo común y 40/2015 del régimen jurídico del sector público*, Atelier, Barcelona.
- URÍOS APARISI, X. y ALAMILLO DOMINGO, I. (2021): *La actuación administrativa automatizada en el ámbito de las Administraciones públicas. Análisis jurídico y metodológico para la construcción y explotación de trámites automáticos*, Escola d’Administració Pública de Catalunya, Barcelona.

- VALERO TORRIJOS, J. (2007): *El régimen jurídico de la e-Administración*, Comares, Granada.
- (2019): “Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración”, *Revista catalana de dret públic*, n.º 58, págs.82-96 (<https://doi.org/10.2436/rcdp.i58.2019.3307> , última visita 28/3/2023).
- WIERZBOWSKI, M. (Coord.) (2022): *Model Rules on Impact Assessment of Algorithmic Decision-Making System Used by Public Administration*, European Law Institute, Viena.
- ZLOTNIK, A. (2019): “Inteligencia Artificial en las Administraciones Públicas: definiciones, evaluación de viabilidad de proyectos y áreas de aplicación”, *Boletic*, n.º 84, págs. 24 sigs.

TERCERA PARTE:
MEDIOS PARA LA
IMPLEMENTACIÓN DE LA
INTELIGENCIA ARTIFICIAL EN EL
SECTOR PÚBLICO

Capítulo X

Reutilización de la información del sector público e inteligencia artificial

AGUSTÍ CERRILLO I MARTÍNEZ¹

Catedrático de Derecho administrativo

Universitat Oberta de Catalunya

Resumen: Las Administraciones públicas producen, recopilan y utilizan grandes volúmenes de datos que pueden ser de gran utilidad para el avance de la inteligencia artificial. En este trabajo se analizan los distintos mecanismos que prevé el ordenamiento jurídico para facilitar la reutilización de datos públicos así como el intercambio de otros datos con el fin de que puedan ser utilizados por las Administraciones públicas en el desarrollo de la inteligencia artificial.

Abstract: Public Administrations produce, collect and use large volumes of data that can be very useful for the progress of artificial intelligence. This paper analyzes the different mechanisms provided by the legal system to facilitate the reuse of public data as well as the exchange of other data so that they can be used by public administrations in the development of artificial intelligence.

Palabras-clave: datos; información pública; reutilización; inteligencia artificial

Key-words: data; public sector information; reuse; artificial intelligence

SUMARIO.- 1. DATOS E INTELIGENCIA ARTIFICIAL.- 2. EL IMPULSO EUROPEO DE LA REUTILIZACIÓN DE DATOS PÚBLICOS Y SU CONTRIBUCIÓN AL DESARROLLO DE LA INTELIGENCIA ARTIFICIAL.- 3. LA REUTILIZACIÓN DE DATOS PARA EL DESARROLLO DE LA INTELIGENCIA ARTIFICIAL.- 3.1. La reutilización de datos públicos.- 3.1.1. La apertura de datos públicos.- 3.1.2. La reutilización de determinadas categorías de datos públicos protegidos por la confidencialidad.- 3.2. El intercambio de datos.- 3.2.1. Los servicios de intermediación de datos.- 3.1.2. La cesión altruista de datos.- 3.3. La obtención de datos necesarios para el desempeño de las tareas realizadas en interés público.- 4. REFLEXIONES FINALES: DISPONIBILIDAD Y CALIDAD DE LOS DATOS COMO FUNDAMENTO DEL DESARROLLO DE LA INTELIGENCIA ARTIFICIAL.- BIBLIOGRAFÍA CITADA

¹ ORCID: 0000-0002-1577-4194

1.- DATOS E INTELIGENCIA ARTIFICIAL

En los últimos años, los datos han ido adquiriendo una mayor importancia en nuestra sociedad, en nuestra economía y en nuestras instituciones públicas a medida que ha ido aumentando la capacidad de generación, almacenamiento y análisis. En esta dirección, hoy en día, es común referirse a los datos como *el oro* (MAYER-SCHÖNBERGER y CUKIER, 2013) o *el petróleo* (THE ECONOMIST, 2017) del siglo XXI.

Esta importancia ha sido ampliamente reconocida. En esta dirección, podemos traer a colación la exposición de motivos de la Propuesta de Reglamento del Parlamento Europeo y del Consejo sobre normas armonizadas para un acceso justo a los datos y su utilización (en adelante, propuesta de Ley de Datos), en la que la Comisión Europea reconoce que “Los datos son un componente fundamental de la economía digital y un recurso esencial para garantizar las transiciones ecológica y digital”². Asimismo, constata que “A lo largo de la última década, las tecnologías digitales han transformado la economía y la sociedad, lo que ha tenido efectos en todos los sectores de actividad y la vida cotidiana. Los datos son elementos centrales de esa transformación: la innovación basada en los datos reportará enormes beneficios tanto a los ciudadanos de la Unión como a la economía”³.

Las Administraciones públicas no son ajenas a ello puesto que producen, recopilan y también utilizan grandes volúmenes de datos⁴. Estos datos se obtienen como resultado de la actividad y funcionamiento de las propias Administraciones públicas. Además, las entidades públicas consiguen datos de sensores situados en los servicios públicos y las infraestructuras en el territorio o de dispositivos conectados a través del Internet de las cosas. Los datos en poder de las Administraciones públicas también pueden ser aportados por las

² COM (2022) 68 final.

³ Considerando 2.

⁴ Ya a inicios de siglo, la Directiva 2003/98/CE del Parlamento Europeo y del Consejo, de 17 de noviembre de 2003, relativa a la reutilización de la información del sector público reconocía que “El sector público recoge, produce, reproduce y difunde una amplia gama de información relativa a numerosos ámbitos, por ejemplo información social, económica, geográfica, meteorológica o turística y sobre empresas, patentes y educación” (considerando 4).

personas y empresas en cumplimiento de las obligaciones previstas en el ordenamiento jurídico o los pueden obtener las Administraciones públicas de los dispositivos o aplicaciones móviles que porten las personas usuarias de los servicios públicos o vecinas de un determinado territorio o del seguimiento de la actividad realizada por ellas en las redes sociales (HINTZ, DENCİK y WAHL-JORGENSEN, 2019).

Estos datos pueden ser de gran utilidad para la producción de nuevos productos y servicios tanto por las Administraciones públicas como por las empresas y la sociedad civil (CERRILLO I MARTÍNEZ, 2012, 790)⁵. Así fue enfáticamente reconocido ya a inicios de este siglo en la Directiva 2003/98/CE del Parlamento Europeo y del Consejo, de 17 de noviembre de 2003, relativa a la reutilización de la información del sector público donde se afirmaba que “La información del sector público constituye una materia prima importante para diversos productos y servicios de contenidos digitales y se convertirá en un recurso cada vez más importante con el desarrollo de los servicios inalámbricos de contenidos”⁶.

No obstante, a medida que ha ido avanzando el desarrollo tecnológico, el valor de los datos en poder de las Administraciones públicas se ha ido incrementando significativamente. Así lo ha advertido, en general, la Comisión Europea en su Comunicación al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones “Configurar el futuro digital de Europa” de 19 de febrero de 2020 donde se reconoce que “Los datos se han convertido en un factor clave de la producción, y el valor que crean tiene que ser compartido con toda la sociedad que contribuye a facilitarlos. Esta es la razón por la que necesitamos construir un verdadero mercado único europeo de datos: un espacio europeo de datos basado en normas y valores europeos”⁷.

En particular, la mayor disponibilidad de datos tiene una aplicación inmediata en el desarrollo de la inteligencia artificial. Tal y como se

⁵ Como ha puesto de manifiesto ONTSI (2020), el 69,9% de los reutilizadores, utiliza datos provenientes del sector público y del sector privado, el 12% solo utiliza datos del sector público y el 18,1% únicamente datos provenientes del sector privado.

⁶ Considerando 5.

⁷ COM (2020) 67 final.

recogió en el Libro blanco sobre la inteligencia artificial—un enfoque europeo orientado a la excelencia y la confianza “Sin datos, el desarrollo de la IA y otras aplicaciones digitales resulta imposible. El enorme volumen de datos nuevos que está por generarse es una oportunidad para que Europa se posicione en la primera línea de la transformación en materia de datos e inteligencia artificial”⁸. Por ello, destaca la exposición de motivos de la Directiva de datos abiertos, “El empleo inteligente de los datos, incluido su tratamiento a través de aplicaciones de inteligencia artificial, puede tener un efecto transformador en todos los sectores de la economía”⁹.

En efecto, a medida que se ha incrementado la disponibilidad de datos y se ha impulsado su uso, la inteligencia artificial ha experimentado un importante progreso. En particular, los nuevos conjuntos de datos disponibles facilitan que los algoritmos de aprendizaje automático (*machine learning*) vayan avanzando en la identificación de patrones y correlaciones entre los datos y en la generación de modelos—cuya calidad es evaluada continuamente y, en su caso, mejorada— que posteriormente son utilizados en la toma de decisiones automatizadas. En esta dirección, Mayer-Schonberger y Cukier (2013, 13) se refieren a *dejar hablar a los datos*.

La Comisión Europea ha reconocido el valor de los datos en el desarrollo de la inteligencia artificial en Europa. En esta dirección, la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (propuesta de Ley de inteligencia artificial) observa que “Muchos sistemas de IA necesitan datos de alta calidad para funcionar correctamente, en especial cuando se emplean técnicas que implican el entrenamiento de modelos, con vistas a garantizar que el sistema de IA de alto riesgo funciona del modo previsto y en condiciones de seguridad y no se convierte en la fuente de alguno de los tipos de discriminación prohibidos por el Derecho de la Unión”¹⁰.

Los datos pueden ser utilizados por aquellos que los producen o por quienes los tienen en su poder. Pero también pueden ser reutiliza-

⁸ COM (2020) 65 final

⁹ Considerando 9.

¹⁰ Considerando 44.

dos por otras personas, empresas y centros de investigación o, incluso, Administraciones públicas en el desarrollo de la inteligencia artificial.

En los últimos años se han ido impulsando diferentes medidas para facilitar que los datos puedan ser utilizados no solo por aquellos que los producen sino también por otras personas o entidades que puedan tener interés en reutilizar estos datos para producir nuevos bienes o servicios. De este modo es importante y necesario no solo garantizar que los datos estén disponibles entre todos aquellos que los han generado sino también que puedan ser fácilmente reutilizados por aquellos otros que participan activamente en el desarrollo de la inteligencia artificial.

2.- EL IMPULSO EUROPEO DE LA REUTILIZACIÓN DE DATOS PÚBLICOS Y SU CONTRIBUCIÓN AL DESARROLLO DE LA INTELIGENCIA ARTIFICIAL

La reutilización de datos públicos consiste en el uso de datos con un propósito distinto para el que estos datos fueron generados.

En las últimas décadas, la reutilización de datos públicos ha experimentado una evolución impulsada por distintas fuerzas.

En primer lugar, la reutilización de los datos públicos se ha visto impulsada por el desarrollo tecnológico. En efecto, la extensión de las redes de comunicación, el incremento de la capacidad de computación de los ordenadores o la evolución en las aplicaciones informáticas están generando un mayor volumen de datos y facilitando su transmisión. Al mismo tiempo, se está incrementando el interés y la utilidad de estos datos al ser más diversos los productos y servicios que pueden crearse a partir de la reutilización de dichos datos. En esta dirección, la Comisión Europea ha observado cómo “el uso inteligente de los datos tiene un efecto transformador en todos los sectores de la economía y el sector público” (COMISIÓN EUROPEA, 2018, 2). En particular, la rápida evolución que está experimentando la inteligencia artificial está dando un nuevo valor a los datos que son utilizados para entrenar o evaluar el desempeño de los algoritmos.

En segundo lugar, la reutilización se ha visto avivada por la apertura de las Administraciones públicas a la ciudadanía con el fin de que

las personas puedan conocer la actividad y el funcionamiento de las Administraciones públicas y estas puedan escuchar las preferencias de la ciudadanía y contar con su participación en la toma de decisiones públicas y su colaboración en la implementación de las políticas públicas y en la prestación de los servicios públicos. Como afirma Noveck "el gobierno abierto es una estrategia innovadora para cambiar la forma como trabaja el gobierno. Usando las tecnologías de las redes para conectar al público con el gobierno y estar mutuamente informados gracias al open data, un gobierno abierto busca apoyar en la solución problemas. El resultado final son instituciones más eficaces y una democracia más robusta" (NOVECK, 2011).

En tercer lugar, la reutilización se ha visto propiciada por el impulso internacional. En 2015, se adoptó la Carta Internacional de Datos Abiertos con la participación de gobiernos, sociedad civil y expertos de todo el mundo en la que se recogen seis principios (abiertos por defecto, oportunos y exhaustivos, accesibles y utilizables, comparables e interoperables, para mejorar la gobernanza y la participación ciudadana y para el desarrollo inclusivo y la innovación)¹¹.

Todas estas fuerzas han propiciado que desde la Unión Europea se hayan ido adoptando en las últimas décadas diferentes políticas y aprobando distintas normas que han ido evolucionando en paralelo.

El interés de la Unión Europea por la reutilización de la información pública surgió a finales del siglo pasado (JANSSEN, K. y DUMORTIER, 2003). En aquel momento desde la Comisión Europea se observó que "una buena disponibilidad de la información pública es un requisito previo indispensable para la competitividad de la industria europea"¹². Asimismo, se constató que existían considerables diferencias entre los estados miembros en relación con las normas y prácticas de explotación de la información del sector público.

¹¹ Entre los impulsos realizados a nivel internacional, no pueden desconocerse los trabajos realizados por instituciones como la OECD (2008): "Recommendation of the Council for enhanced access and more effective use of public sector information [C(2008)36]".

¹² Véase al respecto, el libro verde *La información del sector público: un recurso clave para Europa* [COM(1998)585] y, previamente, las Directrices para mejorar la sinergia entre los sectores público y privado en el mercado de la información de 1989.

Para dar respuesta a todo ello, se impulsó la aprobación de la Directiva 2003/98/CE del Parlamento Europeo y del Consejo de 17 de noviembre, de reutilización de la información del sector público¹³. Esta norma perseguía establecer un conjunto mínimo de normas que facilitasen la reutilización de los documentos existentes conservados por organismos del sector público de los Estados miembros con la finalidad de avanzar en el establecimiento de un mercado interior y dar respuesta a las diferencias entre las normas y prácticas de explotación de los recursos de información del sector público existentes entre los Estados miembros que constituían barreras para el uso pleno de la información pública (CERRILLO I MARTÍNEZ y GALÁN GALÁN, 2006).

Posteriormente, diez años más tarde, se adoptó la Directiva 2013/37/UE del Parlamento Europeo y del Consejo, de 26 de junio de 2013, por la que se modifica la Directiva 2003/98/CE relativa a la reutilización de la información del sector público. La nueva directiva se dictó con el fin de “establecer la obligación inequívoca para los Estados miembros de autorizar la reutilización de todos los documentos, salvo si el acceso está restringido o excluido en virtud de normas nacionales sobre acceso a los documentos”. Asimismo, la Directiva 2013/37/UE amplió el ámbito de aplicación de la normativa europea (por ejemplo, a las bibliotecas, los museos y los archivos) y fomentó el uso de formatos abiertos y legibles por máquina (CERRILLO I MARTÍNEZ, 2020).

A pesar del avance que significó la aprobación de la modificación de la Directiva 2003/98/CE en 2013, pronto se advirtió que la normativa europea quedaba desfasada respecto al desarrollo de la tecnología lo que podía determinar una pérdida de oportunidad para las empresas y la sociedad civil europeas a la hora de crear servicios basados en los datos generados por el sector público. Además, se constató de nuevo que existían profundas diferencias en las regulaciones y en las

¹³ Directiva 2003/98/CE del Parlamento Europeo y del Consejo de 17 de noviembre, de reutilización de la información del sector público. Véase al respecto, CERRILLO I MARTÍNEZ, A. (2012): "The re-use of public sector information in Europe and its impact on transparency", *European Law Journal*, nº 18(6), 770-792.

prácticas de los Estados miembros con respecto a la reutilización de la información del sector público.

Por todo ello, en 2019 se aprobó la Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativa a los datos abiertos y la reutilización de la información del sector público (Directiva de datos abiertos). Esta norma persigue fomentar el uso de datos abiertos y estimular la innovación de los productos y servicios. Entre otras novedades, como veremos en detalle a continuación, la Directiva de datos abiertos amplía los ámbitos subjetivo y objetivo de aplicación de la normativa europea incluyendo, por ejemplo, a las empresas públicas en relación con la información generada durante la prestación de servicios de interés general o a las universidades en relación con los datos producidos en las investigaciones financiadas con fondos públicos. Asimismo, anima a los Estados miembros a promover la creación de datos basados en el *principio abiertos desde el diseño y por defecto*. También promueve la reutilización de datos dinámicos de alto valor económico¹⁴.

El último estadio en la evolución del impulso europeo de la reutilización de la información pública lo encontramos con la adopción de la Estrategia Europea de Datos en febrero de 2020¹⁵. Este documento parte de la constatación de que “La UE tiene el potencial para tener éxito en una economía ágil en el manejo de los datos”. Para ello, la Estrategia Europea de Datos persigue convertir a la Unión Europea en *un modelo de referencia de una sociedad empoderada por los datos para tomar mejores decisiones* a través de la creación de un espacio único europeo de datos, un verdadero mercado único en el que los datos puedan circular por toda la Unión y entre sectores; se respeten

¹⁴ Como pone de manifiesto Valero, “no sólo se ha establecido una nueva regulación de carácter general para impulsar el acceso a los datos en poder de las entidades públicas en toda la Unión Europea sino que, además, se han incorporado unas medidas concretas tendentes a que ciertos conjuntos de datos sean fácilmente reutilizables en unas condiciones favorables debido a la importancia del papel que están llamados a jugar en el actual ecosistema digital” VALERO TORRIJOS, J. (2022): “El régimen jurídico de los datos de alto valor: dificultades y retos para su utilización”, en VALERO TORRIJOS, J. y MARTÍNEZ GUTIÉRREZ, R. (eds.), *Datos abiertos y reutilización de la información del sector público* (pp. 81-102), Granada, Comares.

¹⁵ COM(2020) 66 final.

las normas europeas (por ejemplo, en materia de privacidad, protección de datos o competencia) y existan unas normas justas, prácticas y claras para el acceso a los datos y su utilización. Según la Estrategia, el espacio común europeo de datos que persigue crear ha de ser “como un mercado interior de datos en el que estos puedan utilizarse independientemente de su ubicación física de almacenamiento en la Unión de conformidad con el Derecho aplicable, lo que, entre otras cosas, podría resultar fundamental para el rápido desarrollo de las tecnologías de inteligencia artificial”. Todo ello con el fin de que, en 2030, “la cuota de la UE en la economía de los datos — datos almacenados, tratados y valiosos para su uso en Europa — al menos se corresponda con su peso económico”.

La Estrategia Europea de Datos parte de la constatación de la existencia de distintas cuestiones que impiden o frenan la consecución de esta finalidad como la baja disponibilidad de datos para impulsar el desarrollo de la inteligencia artificial (por ejemplo, se constata que en ocasiones los mismos datos públicos no están disponibles para su reutilización en todos los Estados miembros o que no existen incentivos económicos suficientes o de la confianza entre operadores necesaria para avanzar en el intercambio y el uso de datos entre empresas); la existencia de desequilibrios en el poder de mercado (por ejemplo, la prestación de servicios en la nube o infraestructuras de datos se halla concentrada en pocas empresas o dificultad por parte de las PYMES para acceder a los datos); la baja interoperabilidad y calidad de los datos; la necesidad de reforzar la gobernanza de los datos; la infraestructuras y tecnologías de datos insuficientes (por ejemplo, son pocos los proveedores de servicios en la nube establecidos en la Unión Europea o están sujetos también a la legislación de terceros estados lo que puede contradecir el marco relativo a la protección de datos o generar inseguridad jurídica o, desde la perspectiva de la demanda, hay una baja utilización de los servicios en la nube o existen problemas de interoperabilidad en relación con la portabilidad); la ausencia de normas y herramientas para el ejercicio de los derechos que reconoce la normativa sobre protección de datos personales y privacidad (por ejemplo, práctica inexistencia de herramientas de gestión del consentimiento o de la información personal o de cooperativas de datos); la baja alfabetización en materia de datos; o los nuevos retos que van surgiendo a diario en materia de ciberseguridad.

Para hacerles frente, la Estrategia Europea de Datos prevé la adopción de distintas medidas que se articulan alrededor de cuatro ejes. En primer lugar, una gobernanza intersectorial para el acceso a los datos y su utilización con la idea de desarrollar ecosistemas de datos animados, dinámicos y vigorosos a través de un marco normativo adecuado. En segundo lugar, inversiones en datos y refuerzo de las capacidades e infraestructuras de Europa para albergar, tratar y utilizar los datos, interoperabilidad que puedan servir como catalizadores de la innovación basada en los datos y de la demanda de productos y servicios basados en ellos. En tercer lugar, reforzar las competencias, empoderando a las personas con respecto a los datos e invirtiendo en cualificaciones para poder reducir la brecha en las capacidades existentes en materia de macrodatos y ciencia de los datos y, específicamente, en las pymes y las empresas emergentes. Finalmente, la creación de espacios comunes europeos de datos en sectores estratégicos y en ámbitos de interés público (por ejemplo, en relación a la industria, el Pacto Verde Europeo, la movilidad, la salud, en materia financiera, la energía, el sector agrario, las Administraciones públicas o las cualificaciones) y, complementariamente, el impulso de políticas que estimulen el uso de datos y la demanda de servicios enriquecidos con datos.

La Estrategia va más allá de los datos públicos, pero no descuida su especial relevancia ni la necesidad de apuntar la adopción de medidas específicamente orientadas a facilitar su reutilización.

La Estrategia Europea de Datos se está desplegando a través de distintos instrumentos como la Directiva sobre datos abiertos; el Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos; o la futura Ley de Datos¹⁶. A ellos nos referimos a continuación.

¹⁶ Propuesta adoptada el 23 de febrero de 2022 [COM (2022) 68 final].

Además de estos instrumentos, otras normas adoptadas por la Unión Europea también contribuyen significativamente a los objetivos de la Estrategia Europea de Datos. Por ejemplo, el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE; el Reglamento (UE) 2018/1807 del Parlamento Europeo y del Consejo, de 14 de noviembre de 2018, relativo a un marco para la libre circulación de datos no personales en la Unión Europea; o el Reglamento (UE) 2019/881 del Parlamento Europeo y

3.- LA REUTILIZACIÓN DE DATOS PARA EL DESARROLLO DE LA INTELIGENCIA ARTIFICIAL

Las Administraciones públicas pueden facilitar la reutilización de datos que están en su poder a través de distintos mecanismos que han sido regulados a nivel europeo. Estos datos que pueden ser reutilizados por las propias Administraciones públicas o bien, y especialmente, por personas, empresas o centros de investigación. Pero al mismo tiempo, las Administraciones Públicas pueden utilizar datos que han sido generados por las personas o las empresas con el fin de desarrollar o utilizar la inteligencia artificial.

A continuación, examinamos los distintos mecanismos que pueden ser utilizados para facilitar el intercambio y la reutilización de datos previstos en el ordenamiento comunitario que pueden tener un impacto significativo en el progreso de la inteligencia artificial en Europa.

3.1.- *La reutilización de datos públicos*

En la actualidad, la reutilización de datos públicos está regulada en dos normas que ordenan, por un lado, la apertura de datos públicos para facilitar su reutilización y, por otro, la reutilización de determinadas categorías de datos públicos que están protegidos por algunas normas (por ejemplo, las relativas a la protección de datos personales o la propiedad intelectual).

3.1.1.- La apertura de datos públicos

Uno de los instrumentos que integran la Estrategia Europea de Datos es la apertura de datos públicos. La apertura de datos es el estadio actual de la evolución de los mecanismos de reutilización del sector

del Consejo, de 17 de abril de 2019, relativo a ENISA y a la certificación de la ciberseguridad de las tecnologías de la información y la comunicación y por el que se deroga el Reglamento (UE) 526/2013. Asimismo, hay otras normas sectoriales que también regulan el acceso de los datos en ámbitos como automoción; proveedores de servicios de pago; medición inteligente; red eléctrica; o sistemas de transporte inteligente.

público y se basa en el reconocimiento de la necesidad de remover los obstáculos técnicos y jurídicos con el fin de facilitar la reutilización¹⁷.

La apertura de datos públicos y, en general, la reutilización de la información del sector público se encuentra regulada en la Directiva de datos abiertos y en su transposición a nuestro ordenamiento jurídico a través del Real Decreto-ley 24/2021, de 2 de noviembre que modificó la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público (LRISP).

La LRISP es la regulación básica del régimen jurídico aplicable a la reutilización de los documentos elaborados o custodiados por los sujetos incluidos en su ámbito de aplicación¹⁸. La LRISP regula la reutilización sin perjuicio del régimen aplicable al derecho de acceso a los documentos¹⁹.

De acuerdo con la LRISP, la reutilización es el uso por personas físicas o jurídicas de documentos elaborados o custodiados por las entidades del sector público, con fines comerciales o no comerciales distintos del propósito inicial que tenían esos documentos en la actividad de servicio público para la que se produjeron, excepto para el

¹⁷ Como afirma Cotino “La apertura de los datos es toda una ética, filosofía y práctica para que los datos queden libres sin restricciones de copyright, patentes u otros mecanismos de control, bajo las finalidades compatibles de generar riqueza y plusvalías, así como en su caso, reutilizarlos para la transparencia y el control de las administraciones” COTINO HUESO, L. (2020): “Ética, valores y principios del «open data» y los retos futuros de la apertura de datos públicos”, *Consultor de los ayuntamientos y de los juzgados: Revista técnica especializada en administración local y justicia municipal*, nº 3, 147-166.

¹⁸ Disposición final primera LRISP.

¹⁹ Artículo 1 LRISP. Véase a la remisión a la normativa sobre acceso a la información pública, véase, FERNÁNDEZ SALMERÓN, M. (2022): “Acceso a la información pública y datos abiertos: A vueltas con el ámbito objetivo de la disciplina sobre reutilización y su armonización con la normativa sobre transparencia”, en VALERO TORRIJOS, J. y MARTÍNEZ GUTIÉRREZ, R. (eds.), *Datos abiertos y reutilización de la información del sector público* (pp. 1-20), Granada, Comares.

En relación con esta cuestión, cabe recordar que la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno dispone que “La información sujeta a las obligaciones de transparencia será publicada en las correspondientes sedes electrónicas o páginas web y de una manera clara, estructurada y entendible para los interesados y, preferiblemente, en formatos reutilizables” (artículo 5.4).

intercambio de documentos entre dichos sujetos en el marco de sus actividades de servicio público²⁰.

Ni la Directiva de datos abiertos ni la LRISP imponen una obligación de las Administraciones públicas de facilitar la reutilización de la información del sector público. En efecto, la Directiva (UE) 2019/1024 que dispone que los documentos se facilitarán “siempre que sea posible y apropiado, por medios electrónicos, en formas o formatos que sean abiertos, legibles por máquina, accesibles, fáciles de localizar y reutilizables, conjuntamente con sus metadatos”²¹. Por su parte, la LRISP se refiere a que, si bien los documentos serán reutilizables, los sujetos incluidos en el ámbito de aplicación velarán porque los documentos a los que se aplica dicha norma puedan ser reutilizados para fines comerciales o no comerciales²².

Estas normas tampoco prevén que las entidades del sector público deban difundir los datos en formato abierto. En relación con esta cuestión, la LRISP contempla que “La elaboración y la puesta a disposición de los documentos incluidos en el ámbito de aplicación de la presente Ley se efectuará, en la medida de lo posible, conforme al principio de documentos abiertos desde el diseño y por defecto”²³ y que “Los sujetos previstos en el artículo 2 facilitarán sus documentos en cualquier formato o lengua preexistente, pero también procurarán, siempre que ello sea posible y apropiado, proporcionarlos en formato abierto, accesible, legible por máquina” y que “Tanto el formato como los metadatos, en la medida de lo posible, deben cumplir estándares y normas formales abiertas”. Pero, también se advierte que “Esto no implicará que estén obligados a crear documentos, adaptarlos o facilitar extractos de documentos, cuando ello suponga

²⁰ Artículo 3.1 LRISP. El propio artículo también prevé que la reutilización consisten en el uso de los documentos elaborados o custodiados por “Las sociedades mercantiles públicas a que se refiere el párrafo c) del artículo 2 con fines comerciales o no comerciales distintos del propósito inicial que tenían esos documentos de prestar servicios de interés general para el que se produjeron, excepto para el intercambio de documentos entre estas sociedades mercantiles públicas y el resto de sujetos previstos en el artículo 2 que se realice exclusivamente en el desarrollo de las actividades de servicio público de estos últimos”.

²¹ Artículo 5.1 Directiva (UE) 2019/1024.

²² Artículo 4.1 LRISP.

²³ Artículo 5.1 LRISP.

un esfuerzo desproporcionado que conlleve algo más que una simple manipulación”²⁴.

La LRISP, a la senda de lo dispuesto en la Directiva de datos abiertos, expande el ámbito de sujetos que deben fomentar la reutilización de los datos en su poder. En esta dirección, cabe señalar que todas las administraciones, entidades y organismos que conforman el sector público tienen la obligación de facilitar la reutilización de la información. En particular, según la LRISP, se entiende por Administraciones y organismos del sector público la Administración General del Estado, las administraciones de las Comunidades Autónomas, las entidades que integran la Administración Local, el sector público institucional (organismos públicos y entidades de derecho público vinculados o dependientes de las Administraciones públicas y entidades de derecho privado vinculadas o dependientes de las Administraciones públicas cuando ejerzan potestades administrativas), las universidades públicas así como las empresas públicas por lo que respecta a la prestación de servicios de interés general²⁵.

Asimismo, la LRISP amplía el conjunto de datos y documentos que pueden ser reutilizados. Al respecto, la LRISP se refiere a la reutilización de los documentos que obran en poder de las Administraciones y organismos del sector público bien hayan sido elaborados bien estén custodiados por ellas²⁶. Según el anexo de dicha norma, el documento es “toda información o parte de ella, cualquiera que sea su soporte o forma de expresión, sea esta textual, gráfica, sonora visual o audiovisual, incluyendo los metadatos asociados y los datos contenidos con los niveles más elevados de precisión y desagregación”. A pesar de la amplitud con la que se define el objeto de la reutilización, cabe advertir que existen distintas categorías de documento que quedan al margen de la aplicación de la LRISP, lo que no excluye que pueda facilitarse su reutilización. En esta dirección, no será aplicable a aquellos documentos sobre los que existan prohibiciones o limitaciones en el derecho de acceso a la información pública; los documentos para cuyo acceso se requiera ser titular de un derecho o interés legítimo;

²⁴ Artículo 5.3 LRISP.

²⁵ Artículo 2 LRISP en relación con los artículos 1 y 2 Directiva de datos abiertos.

²⁶ Artículos 3 y 4 LRISP.

los documentos que obran en poder de las Administraciones públicas para finalidades ajenas a las funciones de servicio público; los documentos sobre los que existan derechos de propiedad intelectual o industrial por parte de terceros; los documentos conservados por las entidades que gestionen los servicios esenciales de radiodifusión sonora y televisiva; los documentos producidos o conservados por instituciones educativas y de investigación así como los museos y archivos y otras instituciones culturales; las partes de documentos que solo incluyan logotipos, divisas e insignias; los documentos a los que no pueda accederse o cuyo acceso esté limitado por motivos de protección de los datos personales a menos que se produzca la disociación de los datos; los documentos elaborados por entidades del sector público empresarial y fundacional en el ejercicio de las funciones atribuidas legalmente y los de carácter comercial, industrial o mercantil elaborado en ejecución del objeto social previsto en sus estatutos y los estudios realizados por entidades del sector público en colaboración con el sector privado²⁷.

También se amplía el ámbito objetivo de la reutilización al preverse que las Administraciones públicas pondrán a disposición datos en formatos que permitan la reutilización rápida y sencilla. En esta dirección, la LRISP prevé que los organismos del sector público pondrán a disposición datos dinámicos para su reutilización inmediatamente después de su recopilación, por ejemplo, los datos generados por sensores²⁸. Es decir, se deberá facilitar la rápida reutilización de aquellos datos “en formato digital, sujetos a actualizaciones frecuentes o en tiempo real, debido en particular, a su volatilidad o rápida obsolescencia”²⁹. Para la rápida difusión de los datos dinámicos las Administraciones públicas deben utilizar interfaces de programación de aplicaciones (API) que, cuando sea posible, permitan la descarga masiva de los datos³⁰.

²⁷ Artículo 4.3 LRISP.

²⁸ Artículo 5.5 Directiva (UE) 2019/1024.

²⁹ Artículo 2.8 Directiva (UE) 2019/1024.

³⁰ Artículo 5.4 LRISP. Según el considerando 32 Directiva (UE) 2019/1024, las API deben basarse en los principios de disponibilidad, estabilidad, mantenimiento durante la vida útil, uniformidad de uso y normas, facilidad de uso y seguridad.

Algunos de estos datos tienen alto valor al generar importantes beneficios sociales o económicos porque son idóneos para crear servicios de valor añadido³¹. Los datos de alto valor deben estar disponibles gratuitamente; ser legibles por máquina; ser suministrados a través de APIs, y, cuando sea posible, poder ser descargados en forma masiva³².

Para facilitar la reutilización, los documentos deben estar disponibles con los metadatos pertinentes. Asimismo, estos documentos deberán ser accesibles para personas con discapacidad, siempre que sea posible, apropiado y razonable³³. Además, cuando sea posible, los datos y documentos deben estar disponibles a través de medios electrónicos y mediante plataformas multicanal³⁴. Los documentos deben ser facilitados en cualquier formato o lengua preexistente. En última instancia, la LRISP dispone que los documentos también se difundan en formatos abiertos, o sea, en formatos independientes de plataformas y puestos a disposición del público sin restricciones que impidan la reutilización de los documentos³⁵.

Según la LRISP, las Administraciones públicas pueden optar por facilitar la reutilización a través de distintas modalidades.

En primer lugar, la reutilización puede realizarse sin sujeción a condiciones.

³¹ El anexo I Directiva de datos abiertos recoge una lista de datos de alto valor (datos en el ámbito geoespacial, la observación de la Tierra y medio ambiente, la meteorología, la estadística, las sociedades y propiedad de sociedades y la movilidad. Asimismo, el artículo 3.ter.1 LRISP prevé la posibilidad de fijar otros conjuntos de datos a nivel estatal. Véase al respecto VALERO TORRIJOS, J. (2022): "El régimen jurídico de los datos de alto valor: dificultades y retos para su utilización", en VALERO TORRIJOS, J. y MARTÍNEZ GUTIÉRREZ, R. (eds.), *Datos abiertos y reutilización de la información del sector público* (pp. 81-102), Granada, Comares..

³² Artículo 3.ter.2 LRISP.

³³ Artículo 5.3 LRISP.

³⁴ Artículos 5.2 y 5.3 LRISP.

³⁵ Artículo 5.2 y anexo LRISP. Véase asimismo la Norma Técnica de Interoperabilidad sobre reutilización de recursos de la información aprobada por la Resolución de 19 de febrero de 2013, de la Secretaría de Estado de Administraciones Públicas que tiene por objeto establecer el conjunto de pautas básicas para la reutilización de documentos y recursos de información elaborados o custodiados por el sector público.

En segundo lugar, la reutilización puede sujetarse a condiciones establecidas en licencias-tipo. Estas licencias deben estar disponibles en formato digital y ser procesables electrónicamente³⁶. Las condiciones a las que se someta la reutilización deben ser claras, justas y transparentes; no pueden restringir las posibilidades de reutilización ni limitar la competencia ni ser discriminatorias para categorías comparables de reutilización³⁷. En esta dirección, la propia norma prevé algunas condiciones que se pueden establecer como que el contenido de la información, incluyendo los metadatos, no sea alterado; que no se desnaturalice el sentido de la información; que se cite la fuente; que se mencione la fecha de la última actualización; la finalidad o finalidades concretas de la reutilización cuando contenga datos de carácter personal o la prohibición de revertir el procedimiento de disociación mediante la adición de nuevos datos obtenidos de otras fuentes³⁸.

En tercer lugar, la reutilización también puede someterse a la solicitud previa de una licencia. En este caso, la licencia debe reflejar la información relativa a la finalidad concreta, comercial o no comercial, para la que se concede la reutilización, la duración de la licencia, las obligaciones del beneficiario y del organismo concedente, las responsabilidades de uso y modalidades financieras, indicándose el carácter gratuito o, en su caso, la tasa o precio público aplicable³⁹. Cuando sea posible, las Administraciones públicas usarán licencias abiertas con las mínimas restricciones posibles sobre la reutilización de la información⁴⁰.

Finalmente, en determinadas circunstancias se prevé la posibilidad de que las Administraciones públicas faciliten la reutilización bajo un acuerdo exclusivo cuando sea necesario para la prestación de un servicio de interés público, la Administración pública competente podrá hacerlo, de manera transparente y pública, bajo la

³⁶ Artículo 4.4 LRISP.

³⁷ Artículo 4.3 LRISP.

³⁸ Artículo 8 LRISP.

³⁹ Artículo 9 LRISP.

⁴⁰ Artículo 9.2 LRISP. Véase un análisis en detalle en LÓPEZ-TARRUELLA MARTÍNEZ, A. (2022): "La necesidad de una licencia gubernamental abierta en el ordenamiento español para favorecer el acceso a los datos", en VALERO TORRIJOS, J. y MARTÍNEZ GUTIÉRREZ, R. (eds.), *Datos abiertos y reutilización de la información del sector público* (pp. 159-186), Granada, Comares.

obligación de revisar periódicamente la permanencia del motivo que justificó la concesión del derecho de reutilización en exclusividad⁴¹. No obstante, debe tenerse presente que, en principio, la LRISP prohíbe que se puedan otorgar derechos exclusivos a un reutilizador⁴².

En cualquier caso, la reutilización de los datos se realizará bajo la responsabilidad y riesgo del reutilizador siendo este -y no las Administraciones públicas- quien deberá responder de los daños que pueda generar el uso o la reutilización de la información⁴³.

En principio, para facilitar la reutilización, las Administraciones públicas no deben cobrar ninguna cantidad económica debiendo poner los documentos a disposición de los reutilizadores de manera gratuita. Sin embargo, la LRISP contempla la posibilidad de que las Administraciones públicas puedan establecer una tasa o precio público con el que recuperar los costes marginales que genere la reproducción, puesta a disposición y difusión de los documentos para su reutilización⁴⁴. A la hora de fijar esta tasa o precio público las Administraciones públicas pueden tener en cuenta si la reutilización

⁴¹ Artículo 6.2 LRISP. En particular, por lo que se refiere a los acuerdos exclusivos, véase, GÓMEZ MANRESA, M. F.: "Acuerdos exclusivos y reutilización de la información del sector público", en VALERO TORRIJOS, J. y MARTÍNEZ GUTIÉRREZ, R. (eds.), *Datos abiertos y reutilización de la información del sector público* (pp. 187-213), Granada, Comares.

⁴² Artículo 6.1 LRISP.

⁴³ Artículo 4.7 LRISP. Sobre esta cuestión, véase, BELTRÁN CASTELLANOS, J. M. (2022): "Alcance de la obligación, seguridad jurídica y responsabilidad en la reutilización de la información del sector público", en VALERO TORRIJOS, J. y MARTÍNEZ GUTIÉRREZ, R. (eds.), *Datos abiertos y reutilización de la información del sector público* (pp. 127-158), Granada, Comares. y, previamente, MARTÍNEZ GUTIÉRREZ, R. (2014): "La responsabilidad patrimonial en el acceso, la difusión y la reutilización de la información", en VALERO TORRIJOS, J. y FERNÁNDEZ SALMERÓN, M. (eds.), *Régimen jurídico de la transparencia del sector público: del Derecho de acceso a la reutilización de la información* (pp. 427-458), Cizur Menor, Thomson-Aranzadi.

⁴⁴ Véase un análisis detallado de la regulación de las tasas o precio público por la reutilización en SÁNCHEZ GARCÍA, A. (2022): "El valor económico de la información pública y la regulación del precio de su reutilización", en VALERO TORRIJOS, J. y MARTÍNEZ GUTIÉRREZ, R. (eds.), *Datos abiertos y reutilización de la información del sector público* (pp. 215-252), Granada, Comares.

tendrá fines comerciales o no⁴⁵. Más allá de este principio general, en determinadas circunstancias (por ejemplo, reutilización de documentos facilitados entidades del sector público que no tengan carácter industrial o mercantil, bibliotecas, museos y archivos o sociedades mercantiles), al calcular el precio también pueden tenerse en cuenta los costes de recogida y producción de la información e incrementar el coste total con un margen de beneficio razonable de la inversión⁴⁶. En el caso de que se prevea cobrar alguna tarifa, las Administraciones públicas deben informar de ello por medios electrónicos⁴⁷.

3.1.2.- La reutilización de determinadas categorías de datos públicos protegidos por la confidencialidad

El segundo de los instrumentos que integran la Estrategia Europea de Datos es la reutilización de determinadas categorías de datos que están protegidos por la confidencialidad, por ejemplo, por contener datos personales o secretos comerciales.

Actualmente, este instrumento está previsto en el Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos y por el que se modifica el Reglamento (UE) 2018/1724 (Reglamento de Gobernanza de Datos)⁴⁸.

Con el fin de extender el mercado interior de datos, esta norma persigue crear un marco armonizado entre los Estados miembros para

⁴⁵ Artículo 7 LRISP. La Directiva de datos abiertos reconoce la posibilidad de que las administraciones públicas puedan cobrar por otros servicios que proporcionen respecto a datos de alto valor como, por ejemplo, certificando la autenticidad o veracidad de los datos (considerando 69).

⁴⁶ La Directiva de datos abiertos concreta que el margen de beneficio razonable de la inversión no debería ser superior al 5 % del tipo de interés fijo del Banco Central Europeo (artículo 2.16).

⁴⁷ Artículo 7.3 LRISP.

⁴⁸ Posteriormente, examinaremos el marco de notificación y supervisión para la prestación de servicios de intermediación de datos y el marco para la inscripción voluntaria en un registro de las entidades que recojan y traten datos cedidos con fines altruistas. Además de los aspectos destacados en el texto, el Reglamento de Gobernanza de Datos también regula la creación de un Comité Europeo de Innovación en materia de Datos (artículos 29 y ss.).

el intercambio de determinadas categorías de datos con el fin de generar confianza entre la ciudadanía y las empresas en relación con el acceso, el control, el intercambio, la utilización y la reutilización a los mismos⁴⁹. En particular, se prevén unas normas para el intercambio y la reutilización de aquellos datos públicos y privados que quedan excluidos de la regulación vigente de reutilización por diversos motivos (por ejemplo, la protección datos personales, la existencia de propiedad intelectual o de secretos comerciales).

Al respecto, debemos tener presente que la Directiva de datos abiertos excluye de su ámbito de aplicación determinados datos. En efecto, como destaca la exposición de motivos de la Directiva de datos abiertos en relación con los datos de la investigación financiada públicamente “en este contexto, deben tenerse debidamente en cuenta las inquietudes relacionadas con la privacidad, la protección de datos personales, la confidencialidad, la seguridad nacional, los intereses comerciales legítimos, como los secretos comerciales, y los derechos de propiedad intelectual de terceros, conforme al principio «tan abiertos como sea posible, tan cerrados como sea necesario»”⁵⁰.

De este modo, de manera complementaria a la Directiva de datos públicos, el Reglamento de Gobernanza de Datos amplía el ámbito objetivo de la reutilización de datos públicos al incluir en él buena parte de los documentos excluidos de la aplicación de la Directiva. En particular, en el caso del Reglamento de Gobernanza de Datos, el objeto de la reutilización son aquellos datos que obren en poder de los organismos del sector público que estén protegidos por la confidencialidad comercial, incluidos los secretos comerciales, profesionales o empresariales; la confidencialidad estadística; la protección de los derechos de propiedad intelectual de terceros, o la protección de los datos personales⁵¹.

⁴⁹ El Reglamento de Gobernanza de Datos ya está en vigor aunque prevé que será obligatorio en todos sus elementos y directamente aplicable en cada Estado miembro a partir del 24 de septiembre de 2023. Contiene una disposición transitoria respecto a las entidades que ya presten determinados servicios de intermediación de datos a fecha 23 de junio de 2022, que deberán cumplir las obligaciones establecidas en el capítulo III de este Reglamento a más tardar el 24 de septiembre de 2025.

⁵⁰ Considerando 28.

⁵¹ Artículo 3.1 Reglamento de Gobernanza de Datos.

Así, con la aprobación del Reglamento de Gobernanza de Datos se amplía el conjunto de datos reutilizables. De todos modos, aún quedan documentos fuera del ámbito objetivo de la reutilización por las dificultades inherentes que puede comportar por la coexistencia de otros bienes o derechos como, por ejemplo, los documentos que obren en poder de empresas públicas excepto cuando puedan derivar del ejercicio de funciones del sector público o de la prestación de servicios de interés general; los datos que obren en poder de organismos públicos de radiodifusión; de centros culturales y de enseñanza; de organismos del sector público que estén protegidos por motivos de seguridad pública, defensa o seguridad nacional: o los datos cuya facilitación constituya una actividad ajena al ámbito de la misión de servicio público de los organismos del sector público de que se trate⁵².

El Reglamento de Gobernanza de Datos define la reutilización en el mismo sentido que la Directiva de datos públicos como “la utilización, por personas físicas o jurídicas, de los datos que obren en poder de organismos del sector público, con fines comerciales o no comerciales distintos del propósito inicial englobado en la misión de servicio público para el que se hayan producido tales datos, excepto en el caso del intercambio de datos entre organismos del sector público con la única finalidad de desempeñar sus actividades de servicio público”⁵³.

Del mismo modo que la Directiva de datos abiertos, el Reglamento de Gobernanza de Datos no obliga a los organismos del sector público a permitir la reutilización de los datos en su poder. Así lo advierte enfáticamente el artículo 1 al afirmar que “El presente Reglamento no obliga a los organismos del sector público a permitir la reutilización de datos”. No obstante, nuevamente en este caso, es evidente que la finalidad que persigue el legislador comunitario es promover la reutilización de estos datos exhortando a los organismos públicos a que los pongan a disposición de los reutilizadores.

El Reglamento de Gobernanza de Datos parte de una prohibición de acuerdos exclusivos o acuerdos cuyo objetivo o efecto sea conceder derechos exclusivos o restringir la disponibilidad de los datos para su

⁵² Artículo 3.2 Reglamento de Gobernanza de Datos.

⁵³ Artículo 2.2) Reglamento de Gobernanza de Datos.

reutilización por entidades distintas de las partes en tales acuerdos o prácticas⁵⁴. No obstante, como la Directiva de datos abiertos, el Reglamento admite excepciones a esta prohibición cuando el acuerdo exclusivo sea la única manera de garantizar la prestación de un servicio o el suministro de un producto de interés general. En este caso, el acuerdo deberá concederse de manera transparente y siendo compatible con la normativa vigente en materia de contratación pública y no puede tener una duración superior a los doce meses⁵⁵.

Respecto a la reutilización, el Reglamento de Gobernanza de Datos prevé que los organismos del sector público deben establecer unas condiciones para proteger de manera efectiva los derechos e intereses de terceros en relación con los datos susceptibles de ser reutilizados. Así, con carácter general, debe advertirse que el Reglamento de Gobernanza de Datos no exime de la aplicación de las obligaciones en materia de confidencialidad (por ejemplo, el Reglamento General de Protección de Datos, el derecho de la competencia o la relativa a la seguridad pública, la defensa y la seguridad nacional), ni tampoco de las que puedan derivarse de la normativa la relativa al acceso o la reutilización de la información pública⁵⁶.

De la lectura de esta norma se desprende la posibilidad de que los organismos públicos impongan condiciones a los reutilizadores como, por ejemplo, supeditar la reutilización a la observación de una obligación de confidencialidad que prohíba la divulgación de cualquier información que ponga en peligro los derechos e intereses de terceros, que el reutilizador pueda haber adquirido a pesar de las garantías establecidas exigiendo la anonimización de los datos personales; prever que el acceso y reutilización a distancia de los datos se lleven a cabo en un entorno de tratamiento seguro; exigir la preservación de la integridad del funcionamiento de los sistemas técnicos del entorno de tratamiento seguro utilizado o el cumplimiento de los derechos de propiedad intelectual). Estas condiciones deben perseguir servir a los intereses de los reutilizadores sin que al tiempo supongan una carga para los organismos del sector público. Además, las condiciones que

⁵⁴ Artículo 4.1 Reglamento de Gobernanza de Datos.

⁵⁵ Artículo 4 apartados 3, 4 y 5 Reglamento de Gobernanza de Datos.

⁵⁶ Artículo 1.2 Reglamento de Gobernanza de Datos.

establezcan los organismos públicos no deben ser discriminatorias ni pueden restringir la competencia promoviendo el acceso a los datos de las pymes y las empresas emergentes. Además, las condiciones deben ser transparentes y proporcionadas y estar justificadas objetivamente⁵⁷. Para facilitar el cumplimiento de las condiciones, el Reglamento dispone que los organismos del sector público deben proporcionar cuando sea posible orientaciones y asistencia a los reutilizadores.

El Reglamento de Gobernanza de Datos prevé que los organismos del sector público deben resolver las solicitudes de reutilización en el plazo máximo de dos meses salvo que el Derecho nacional prevea un plazo más breve. Este plazo se podrá ampliar en 30 días en el caso de solicitudes de reutilización excepcionalmente extensas y complejas⁵⁸.

Los Estados miembros deben designar los organismos competentes para asistir a los organismos del sector público encargados de gestionar las solicitudes de reutilización de datos, que podrán existir o ser de nueva creación⁵⁹. En esta dirección, estos organismos pueden encargarse, por ejemplo, de dar apoyo técnico mediante la habilitación de un entorno de tratamiento seguro; orientar y ofrecer apoyo técnico sobre la mejor manera de estructurar y almacenar los datos para que sean fácilmente accesibles; prestar apoyo técnico para la seudonimización y para garantizar que el tratamiento de los datos pueda llevarse a cabo de manera que se preserve de manera efectiva la privacidad, la confidencialidad, la integridad y la accesibilidad de la información; prestar asistencia a los reutilizadores para solicitar el consentimiento de los interesados para la reutilización o asistir en la evaluación de la idoneidad de las obligaciones contractuales contraídas por un reutilizador). A estos organismos también se les puede atribuir la facultad de conceder el acceso para la reutilización de los datos. Para ello deben otorgárseles los recursos jurídicos, financieros, técnicos y humanos adecuados⁶⁰.

Para facilitar la reutilización de los datos, el Reglamento de Gobernanza de Datos prevé que los Estados miembros creen un punto

⁵⁷ Artículo 5 Reglamento de Gobernanza de Datos.

⁵⁸ Artículo 8 Reglamento de Gobernanza de Datos.

⁵⁹ Los Estados miembros deben notificar a la Comisión la identidad de dichos organismos a más tardar el 24 de septiembre de 2023.

⁶⁰ Artículo 7.1 Reglamento de Gobernanza de Datos.

único de información donde toda la información esté disponible y sea fácilmente accesible. Este punto puede estar vinculado a los puntos de información de ámbito sectorial o regional o local existentes y puede desarrollar las funciones de manera automatizada. El punto podrá recibir las consultas y también las solicitudes de reutilización que serán transmitidas a los órganos u organismos competentes para su tramitación⁶¹.

El Reglamento de Gobernanza de Datos prevé que los organismos del sector público puedan cobrar una tasa por permitir la reutilización. Estas tasas deben ser transparentes, no discriminatorias y proporcionadas, estarán justificadas objetivamente y no podrán restringir la competencia. Para calcular las tasas deben tenerse en cuenta los costes relacionados con la tramitación de las solicitudes de reutilización así como los necesarios para la reproducción, la entrega y la difusión de los datos; la adquisición de derechos; la anonimización o preparación de los datos personales y de los datos comerciales confidenciales; el mantenimiento del entorno de tratamiento seguro; la adquisición, por parte de terceros ajenos al sector público, del derecho de terceros de permitir la reutilización, y la asistencia a los reutilizadores en la obtención del consentimiento de los interesados y del permiso de los titulares de datos cuyos derechos e intereses puedan verse afectados por la reutilización. Los criterios y la metodología para el cálculo de la tasa, los sujetos a los que se les aplique un descuento y los costes que integran deben publicarse. Por último, también se prevé que para incentivar la reutilización de los datos con fines no comerciales o por parte de pymes, empresas emergentes, sociedad civil o centros educativos, los organismos del sector público pueden prever descuentos en las tasas o incluso ofrecer los datos de manera gratuita⁶².

Finalmente, el Reglamento de Gobernanza de Datos prevé la existencia de mecanismos de garantía para que toda persona física o jurídica directamente afectada por una decisión adoptada en el marco del mismo disponga de un derecho efectivo de recurso en el Estado miembro y, asimismo, tenga la posibilidad de que su decisión pueda ser revisada por un órgano imparcial con los conocimientos especia-

⁶¹ Artículo 8 Reglamento de Gobernanza de Datos.

⁶² Artículo 6 Reglamento de Gobernanza de Datos.

lizados adecuados (por ejemplo, la autoridad nacional de defensa de la competencia, la autoridad reguladora del acceso a los documentos o la autoridad de control en materia de protección de datos) o un órgano jurisdiccional nacional⁶³.

3.2.- El intercambio de datos

Las Administraciones públicas obtienen sus datos principalmente en el marco del desarrollo de sus funciones. También porque las personas y las empresas se los facilitan en cumplimiento de las obligaciones previstas en el ordenamiento jurídico o en el marco de la prestación de servicios públicos.

No obstante, las Administraciones públicas también pueden obtener datos a través de su adquisición mediante servicios de intermediación de datos o gracias a la cesión altruista de datos.

Estos mecanismos para la obtención de datos están también regulados en el Reglamento de Gobernanza de Datos que, además de la reutilización de determinadas categorías de datos públicos protegidos por la confidencialidad que hemos analizado en el epígrafe anterior, regula otros mecanismos de intercambio de datos con los que desarrollar el mercado interior e impulsar una sociedad y una economía basada en los datos centradas en el ser humano y que sea fiable y segura⁶⁴.

3.2.1.- Los servicios de intermediación de datos

Los servicios de intermediación de datos son aquellos servicios que persiguen establecer relaciones comerciales para el intercambio de datos entre un número indeterminado de interesados y titulares de datos, por una parte, y usuarios de datos, por otra, a través de medios técnicos, jurídicos o de otro tipo⁶⁵.

El Reglamento de Gobernanza de Datos dispone que para prestar determinados servicios de intermediación de datos debe seguirse un

⁶³ Artículo 9.2 Reglamento de Gobernanza de Datos.

⁶⁴ Considerando 3.

⁶⁵ Artículo 2.11) Reglamento de Gobernanza de Datos.

procedimiento de notificación con posterioridad a la cual el proveedor de servicios de intermediación podrá iniciar su actividad⁶⁶. La notificación dará derecho a prestar los servicios en todos los Estados miembros si bien le será de aplicación del ordenamiento jurídico del Estado miembro donde tenga su establecimiento principal⁶⁷. Entre otros, el proveedor puede solicitar que la autoridad competente le confirme que cumple con lo que dispone el Reglamento de Gobernanza de Datos y recibida la confirmación de tal extremo, podrá usar la denominación “proveedor de servicios de intermediación de datos reconocido en la Unión” junto con el logotipo común que se establezca⁶⁸.

De acuerdo con el Reglamento de Gobernanza de datos, la prestación de servicios de intermediación de datos estará sujeta a distintas condiciones como que no se puedan utilizar los datos para fines diferentes de su puesta a disposición de los usuarios de datos; que las condiciones contractuales comerciales no dependan de que el titular de datos o el usuario de datos utilice otros servicios prestados por el mismo proveedor; que los proveedores de servicios intercambien los datos en el mismo formato en el que los reciban de parte del interesado o del titular de datos pudiendo únicamente convertirlos a otros formatos en los casos previstos (por ejemplo, para mejorar la interoperabilidad o lo solicita el usuario de datos); que los servicios de intermediación de datos únicamente incluyan la oferta de herramientas y servicios específicos adicionales con el objetivo específico de facilitar el intercambio de los datos (por ejemplo, almacenamiento temporal, organización, conversión, anonimización o seudonimización de datos); o que velen por que el procedimiento de acceso a sus servicios sea equitativo, transparente y no discriminatorio⁶⁹.

Las autoridades competentes en materia de servicios de intermediación de datos, que deben ser designadas en cada Estado miembro,

⁶⁶ Artículo 10 Reglamento de Gobernanza de Datos. El propio Reglamento también reconoce que la regulación prevista no se aplica a las organizaciones reconocidas de gestión de datos con fines altruistas ni a otras entidades sin ánimo de lucro (artículo 15).

⁶⁷ Artículo 11 apartados 2, 4 y 5 Reglamento de Gobernanza de Datos.

⁶⁸ Artículo 11.9 Reglamento de Gobernanza de Datos.

⁶⁹ Artículo 12 Reglamento de Gobernanza de Datos.

además de realizar las funciones relacionadas con el procedimiento de notificación, deben controlar y supervisar el cumplimiento de los requisitos previstos en el Reglamento de Gobernanza de Datos por los proveedores de servicios de intermediación de datos⁷⁰. Estas autoridades deben desarrollar sus funciones sin perjuicio de las que estén atribuidas a las autoridades de protección de datos, de defensa de la competencia o responsables de ciberseguridad. Asimismo deben articular los mecanismos necesarios para garantizar una sólida cooperación e intercambio de información para el desempeño de sus funciones.

En esta dirección, el Reglamento dispone que la autoridad competente puede solicitar a los proveedores de servicios de intermediación la información necesaria para comprobar que cumplen con los requisitos previstos. En el caso de que constaten que no cumplen con dichos requisitos, notificarán al proveedor las observaciones dándole la oportunidad de formular sus alegaciones en el plazo de diez días. Asimismo, la autoridad competente puede exigir el cese de la infracción en un plazo razonable, o de manera inmediata en el caso de las infracciones graves, pudiendo también adoptar las medidas adecuadas y proporcionadas necesarias para garantizar el cumplimiento. En esta dirección, el Reglamento de Gobernanza de Datos prevé que la autoridad competente pueda imponer sanciones; exigir un aplazamiento del inicio o una suspensión de la prestación del servicio de intermediación de datos hasta que las condiciones se hayan modificado; o exigir el cese de la prestación del servicio de intermediación de datos en caso de infracciones graves o reiteradas que no hayan sido corregidas⁷¹.

Por último, el Reglamento de Gobernanza de Datos prevé que toda persona tendrá derecho a presentar una reclamación en relación con cualquier cuestión comprendida en el ámbito de aplicación del mismo ante la autoridad competente⁷². Asimismo, reconoce que toda persona física o jurídica afectada tendrá derecho a una tutela judicial efectiva en lo que respecta a las decisiones jurídicamente vinculantes adoptadas por las autoridades competentes en materia de servicios de intermediación de datos⁷³.

⁷⁰ Artículos 12 y 13 Reglamento de Gobernanza de Datos.

⁷¹ Artículo 14 Reglamento de Gobernanza de Datos.

⁷² Artículo 27 Reglamento de Gobernanza de Datos.

⁷³ Artículo 28 Reglamento de Gobernanza de Datos.

3.1.2.- La cesión altruista de datos

La cesión altruista de datos consiste en el intercambio voluntario de datos basado en el consentimiento de los interesados para que se traten sus datos personales, o en el permiso de los titulares de datos para que se usen sus datos no personales, sin ánimo de obtener o recibir una gratificación con objetivos de interés general (por ejemplo, la asistencia sanitaria, la lucha contra el cambio climático o la mejora de la prestación de servicios públicos o la elaboración de políticas públicas)⁷⁴.

En relación con la cesión altruista de datos, el Reglamento de Gobernanza de Datos emplaza a los Estados miembros a que adopten las políticas y establezcan las medidas organizativas y técnicas necesarias para facilitar la cesión altruista de datos.

Asimismo, dispone que las organizaciones de gestión de datos con fines altruistas deben inscribirse en un registro nacional de organizaciones reconocidas de gestión de datos con fines altruistas gestionado por la autoridad competente⁷⁵. Cuando estén registradas las entidades podrán utilizar la denominación de “organización de gestión de datos con fines altruistas reconocida en la Unión”, así como un logotipo común diseñado por la Comisión⁷⁶.

Para poderse inscribir en el registro, entre otros requisitos, se prevé que las entidades deben cumplir un código normativo que será adoptado por la Comisión en el que se fijen los requisitos de información adecuados que deberá facilitarse a los interesados y a los titulares de datos, antes de conceder su consentimiento o permiso para la cesión altruista de datos; los requisitos técnicos y de seguridad adecuados para garantizar un nivel de seguridad adecuado; las hojas de ruta sobre comunicación; y las recomendaciones sobre las normas de interoperabilidad pertinentes⁷⁷.

Las entidades deben llevar un registro de las personas físicas o jurídicas a las que se haya permitido tratar datos que obren en su poder; la fecha o la duración del tratamiento de los datos; la finalidad

⁷⁴ Artículo 2.16) Reglamento de Gobernanza de Datos.

⁷⁵ Artículo 19 Reglamento de Gobernanza de Datos.

⁷⁶ Artículo 17 Reglamento de Gobernanza de Datos.

⁷⁷ Artículo 22 Reglamento de Gobernanza de Datos.

del tratamiento de datos declarada; y, en su caso, las tasas abonadas. Asimismo, deben elaborar anualmente un informe de actividades que enviarán a la autoridad competente⁷⁸.

Por otro lado, las entidades no pueden utilizar los datos para objetivos distintos de los de interés general para los que el interesado o el titular de datos permita el tratamiento a quienes deben informar de ello, así como, en el caso de que el tratamiento se efectúe en un tercer país, sobre la ubicación de cualquier actividad de tratamiento⁷⁹.

Las autoridades competentes para la inscripción en el registro serán las encargadas de controlar y supervisar el cumplimiento de los requisitos previstos. Estas autoridades podrán solicitar a las organizaciones reconocidas de gestión de datos con fines altruistas toda la información necesaria a fin de comprobar el cumplimiento de los requisitos y en el caso de que considere que no se cumplen con los requisitos establecidos notificará a la organización sus observaciones concediéndole un plazo de treinta días para presentar sus alegaciones. Asimismo, podrán exigir el cese de la infracción y adoptar las medidas necesarias para garantizarlo. Cuando la organización no cumpla con los requisitos, perderá el derecho a utilizar la denominación “organización de gestión de datos con fines altruistas reconocida en la Unión” y se cancelará su inscripción en el registro⁸⁰.

Como en el caso de los servicios de intermediación de datos, el reglamento también contempla que toda persona tendrá derecho a presentar una reclamación en relación con cualquier cuestión comprendida en el ámbito de aplicación del mismo ante la autoridad competente y afectada tendrá derecho a una tutela judicial efectiva en lo que respecta a las decisiones jurídicamente vinculantes adoptadas por las autoridades competentes en materia de cesión altruista de datos⁸¹.

⁷⁸ Artículo 20 Reglamento de Gobernanza de Datos.

⁷⁹ Artículo 21 Reglamento de Gobernanza de Datos.

⁸⁰ Artículo 24 Reglamento de Gobernanza de Datos.

⁸¹ Artículos 27 y 28 Reglamento de Gobernanza de Datos.

3.3.- *La obtención de datos necesarios para el desempeño de las tareas realizadas en interés público*

El último instrumento al que debemos referirnos es la obtención de datos para la realización de determinadas tareas de interés público que está previsto en la propuesta de Reglamento del Parlamento Europeo y del Consejo sobre normas armonizadas para un acceso justo a los datos y su utilización (propuesta de Ley de Datos) que será adoptado en el marco de la Estrategia Europea de Datos⁸².

Desde un punto de vista general, la propuesta de Ley de Datos persigue fomentar el acceso a los datos y su utilización así como garantizar la equidad en el proceso de asignación del valor de los datos. Con esta norma se quiere facilitar que las empresas de la Unión Europea tengan la oportunidad de innovar y competir a partir del uso de los datos a través del establecimiento de unas normas armonizadas sobre la puesta a disposición de datos generados por el uso de un producto o servicio, de datos por parte de los titulares de datos para los destinatarios de datos y de datos por parte de los titulares de datos para organismos públicos para el desempeño de una misión realizada en interés público, cuando exista una necesidad excepcional⁸³.

Para ello, entre otras medidas, la propuesta de Ley de Datos impone al titular de datos -es decir, a cualquier persona física o jurídica que tenga el derecho, la obligación o la capacidad, en el caso de los datos no personales, de poner a disposición determinados datos-⁸⁴ la obligación de poner los datos a disposición de los usuarios y de terceros designados por los usuarios en determinadas circunstancias⁸⁵. Todo ello en condiciones justas, razonables y no discriminatorias y de manera transparente⁸⁶. A tal efecto, regula aspectos como el ámbito de aplicación de esta obligación⁸⁷; la información sobre los productos o servicios que, en un formato claro y comprensible, debe ponerse a

⁸² COM(2022) 68 final.

⁸³ Artículo 1 propuesta Ley de datos.

⁸⁴ Artículo 2.6) propuesta Ley de datos.

⁸⁵ Artículo 3 propuesta Ley de datos.

⁸⁶ Artículo 8 propuesta Ley de datos.

⁸⁷ Artículo 12 propuesta Ley de datos.

disposición de las personas usuarias⁸⁸; el derecho de los usuarios a acceder y utilizar datos generados por el uso de productos o servicios relacionados⁸⁹; el derecho a compartir datos con terceros⁹⁰; las obligaciones de terceros que reciben datos a petición del usuario⁹¹; la compensación por la puesta a disposición de los datos⁹²; la resolución de litigios⁹³; las medidas técnicas de protección y disposiciones sobre la utilización o divulgación no autorizadas de datos⁹⁴; o la no vinculación de las cláusulas contractuales sobre el acceso a los datos y su uso cuando sean abusivas⁹⁵.

A pesar de la relevancia de esta obligación y su regulación, en este trabajo nos interesa especialmente destacar que la propuesta de Ley de Datos incluye específicamente la obligación del titular de datos de facilitar a los organismos públicos determinados datos debido a la concurrencia de una necesidad excepcional de utilizar dichos datos⁹⁶.

En particular, se prevé que existirá una necesidad excepcional de utilizar los datos cuando concurra alguna de las siguientes circunstancias⁹⁷:

- que los datos sean necesarios para responder a una emergencia pública. Según la propuesta, la emergencia pública es “una situación excepcional que afecta negativamente a la población de la Unión, de un Estado miembro o de una parte de él, que entraña un riesgo de repercusiones graves y duraderas en las condiciones de vida o la estabilidad económica, o la degrada-

⁸⁸ Artículo 3.2 propuesta Ley de datos.

⁸⁹ Artículo 4 propuesta Ley de datos.

⁹⁰ Artículo 5 propuesta Ley de datos.

⁹¹ Artículo 6 propuesta Ley de datos.

⁹² Artículo 9 propuesta Ley de datos.

⁹³ Artículo 10 propuesta Ley de datos.

⁹⁴ Artículo 11 propuesta Ley de datos.

⁹⁵ Artículo 13 propuesta Ley de datos.

⁹⁶ Artículo 14 propuesta Ley de datos. A pesar de la referencia en el texto a las Administraciones públicas, debemos señalar que en la propuesta se hace referencia a los organismos del sector público y a las instituciones, órganos u organismos de la Unión.

⁹⁷ Artículo 15 propuesta Ley de datos.

ción sustancial de los activos económicos de la Unión o del Estado miembro en cuestión”⁹⁸

- que los datos sean necesarios para prevenir una emergencia pública o para ayudar a la recuperación tras una emergencia pública y la solicitud de datos esté limitada en el tiempo y en su ámbito de aplicación;
- que la falta de datos disponibles impida a las Administraciones públicas desempeñar una tarea específica de interés público que esté explícitamente prevista por ley siempre y cuando no haya podido obtener esos datos por medios alternativos (por ejemplo, mediante la compra de datos en el mercado a precios de mercado o en cumplimiento de obligaciones de facilitar datos), y la adopción de nuevas medidas legislativas no pueda garantizar la disponibilidad oportuna de los datos; o la obtención de los datos de acuerdo con lo previsto en la nueva norma reduzca de manera sustancial la carga administrativa para los titulares de datos.

Las solicitudes que cursen las Administraciones públicas a los titulares de datos deben concretar los datos necesarios; acreditar la existencia de una necesidad de excepcional; explicar el uso que darán a los datos y la duración del mismo; indicar la base jurídica en qué se fundamenta la solicitud; concretar el plazo para facilitar los datos o para que el titular de los mismos pueda solicitar la modificación o retirada de la solicitud⁹⁹. Estas solicitudes deben ser formuladas en un lenguaje claro, conciso y sencillo comprensible para el titular de datos. Asimismo, las solicitudes deben guardar proporción con la necesidad excepcional, en cuanto a la granularidad y volumen de los datos solicitados y la frecuencia de acceso a los datos solicitados y respetar los objetivos legítimos del titular de datos. Por último, también debe informar de las sanciones que se pueden imponer en caso de que no se atienda la solicitud¹⁰⁰.

⁹⁸ Artículo 2.10) propuesta Ley de datos.

⁹⁹ Artículo 17.1 propuesta Ley de datos.

¹⁰⁰ Artículo 17.2 propuesta Ley de datos.

El titular de los datos debe poner a disposición de la administración solicitante los datos sin demora indebida¹⁰¹.

Pero si los datos no están disponibles o si la solicitud no reúne las condiciones previstas, el titular de los datos también podrá denegar la solicitud o pedir su modificación en un plazo de cinco días hábiles, en el caso de que la solicitud sea para dar respuesta a una emergencia pública, o de quince días hábiles, en otros casos de necesidad excepcional. Asimismo, si la solicitud se basa en la necesidad de dar respuesta a una emergencia pública, el titular de datos podrá también denegar la solicitud cuando previamente se hayan facilitado los datos con la misma finalidad a otra Administración pública y no se haya notificado la destrucción de los mismos en cuyo caso deberá indicar la identidad de dicha pública¹⁰².

Las Administraciones públicas a las que se les haya facilitado los datos no los podrán utilizar de manera incompatible con la finalidad para la que fueron solicitados. Además, deberán aplicar medidas técnicas y organizativas que protejan los derechos y libertades de los interesados cuando existan datos personales. Finalmente, deberán destruir los datos tan pronto como dejen de ser necesarios para la finalidad indicada e informarán al titular de datos de su destrucción¹⁰³.

La propuesta de Ley de datos prevé que la puesta a disposición de datos en respuesta a una emergencia pública sea gratuita. No obstante, también se contempla que cuando la solicitud se base en otros supuestos de necesidad excepcional, el titular de datos puede pedir una compensación por facilitar datos que no podrá exceder de los costes técnicos y organizativos soportados para dar cumplimiento a la solicitud más un margen razonable¹⁰⁴.

Por último, se establece que los datos obtenidos por las Administraciones públicas no podrán ser facilitados para su reutilización así como que la Directiva de datos abiertos no será de aplicación a los

¹⁰¹ Artículo 18 propuesta Ley de datos.

¹⁰² Artículo 18 propuesta Ley de datos.

¹⁰³ Artículo 19 propuesta Ley de datos.

¹⁰⁴ Artículo 20 propuesta Ley de datos.

datos que obren en poder de las Administraciones públicas obtenidos según lo dispuesto en la propuesta de Ley de datos¹⁰⁵.

4.- REFLEXIONES FINALES: DISPONIBILIDAD Y CALIDAD DE LOS DATOS COMO FUNDAMENTO DEL DESARROLLO DE LA INTELIGENCIA ARTIFICIAL

Como hemos tenido oportunidad de exponer a lo largo de estas páginas, en los últimos años se han ido aprobando diversas normas que persiguen incrementar el volumen de datos disponibles para ser utilizados, entre otras finalidades, en el desarrollo de la inteligencia artificial. De este modo, a través de los distintos mecanismos previstos, las Administraciones Públicas pueden contribuir a incrementar el volumen de datos a disposición de empresas y centros de investigación para diseñar nuevos algoritmos o evaluar su funcionamiento y al mismo tiempo las propias Administraciones Públicas pueden utilizar datos de empresas y centros de investigación.

Estas normas pueden tener un impacto significativo en la disponibilidad de datos para el desarrollo de la inteligencia artificial. No obstante, pueden no ser suficientes para avanzar en el desarrollo de la inteligencia artificial en la medida en que, generalmente y al margen de alguna referencia esporádica, las distintas normas analizadas no prevén mecanismos específicos para garantizar la calidad de los datos dejando este extremo en manos de cada Administración pública.

En relación con esta cuestión, no podemos desconocer que para que los datos puedan ser útiles para las Administraciones Públicas y de interés y relevancia para las empresas y centros de investigación para impulsar la inteligencia artificial es necesario garantizar que los datos sean de calidad (JANSSEN, M., BROUS, ESTEVEZ, BARBOSA, y JANOWSKI, 2020). De hecho, el uso de datos de escasa calidad por los algoritmos puede llevar a decisiones erróneas o a resultados injustos o discriminatorios. En esta dirección se manifiesta enfáticamente la propuesta de Ley de inteligencia artificial cuando afirma que “si el sistema de IA no está entrenado con datos de buena calidad,

¹⁰⁵ Artículo 17.3 propuesta Ley de datos.

no cumple los requisitos oportunos en términos de precisión o solidez, o no se diseña y prueba debidamente antes de introducirlo en el mercado o ponerlo en servicio, puede señalar a personas de manera discriminatoria, incorrecta o injusta” por lo que reconoce que “Los conjuntos de datos de entrenamiento, validación y prueba deben ser lo suficientemente pertinentes y representativos, carecer de errores y ser completos en vista de la finalidad prevista del sistema”¹⁰⁶.

Para garantizar la calidad de los datos es importante que las Administraciones Públicas se doten de una gobernanza de datos adecuada (AL-RUITHE, BENKHELIFA y HAMEED, 2019). En efecto, para garantizar la calidad de los datos es necesario que las Administraciones Públicas cuenten con unos principios, valores y normas que pauten cómo deben interaccionar los que crean, gestionan y conservan los datos, los que definen su acceso y uso, los que acceden, usan y reutilizan los datos en una organización en el proceso de toma de decisiones sobre los datos aportando cada uno de ellos sus distintas visiones y los recursos diversos que ostentan (estratégicos, archivísticos, tecnológicos, normativos, económicos, etc.) (CERRILLO I MARTÍNEZ y CASADESÚS DE MINGO, 2021). En esta dirección también se pronuncia de manera clara la propuesta de Ley de inteligencia artificial cuando afirma que “Es preciso instaurar prácticas adecuadas de gestión y gobernanza de datos para lograr que los conjuntos de datos de entrenamiento, validación y prueba sean de buena calidad”¹⁰⁷.

En la medida en qué estos datos no solo sean de utilidad para las propias Administraciones Públicas si no también para las empresas y los centros de investigación, para garantizar la calidad de los datos, las Administraciones Públicas podrían impulsar entornos de colaboración con las empresas y los centros de investigación en los que poder utilizar mecanismos que permitan que todos ellos puedan

¹⁰⁶ Considerandos 38 y 44 propuesta de Ley de inteligencia artificial. Posteriormente, el considerando 44 reconoce que “Muchos sistemas de IA necesitan datos de alta calidad para funcionar correctamente, en especial cuando se emplean técnicas que implican el entrenamiento de modelos, con vistas a garantizar que el sistema de IA de alto riesgo funciona del modo previsto y en condiciones de seguridad y no se convierte en la fuente de alguno de los tipos de discriminación prohibidos por el Derecho de la Unión”.

¹⁰⁷ Considerando 44 propuesta de Ley de inteligencia artificial.

aportar los recursos necesarios para garantizar la calidad de los datos que puedan ser utilizados o reutilizados en el uso y el desarrollo de la inteligencia artificial.

BIBLIOGRAFÍA CITADA

- AL-RUITHE, M., BENKHELIFA, E. y HAMEED, K. (2019): "A systematic literature review of data governance and cloud data governance", *Personal and Ubiquitous Computing*, 839–859. doi:doi.org/10.1007/s00779-017-1104-3
- BELTRÁN CASTELLANOS, J. M. (2022): "Alcance de la obligación, seguridad jurídica y responsabilidad en la reutilización de la información del sector público", en VALERO TORRIJOS, J. y MARTÍNEZ GUTIÉRREZ, R. (eds.), *Datos abiertos y reutilización de la información del sector público* (pp. 127-158), Granada, Comares.
- CERRILLO I MARTÍNEZ, A. (2012): "The re-use of public sector information in Europe and its impact on transparency", *European Law Journal*, nº 18(6), 770–792.
- CERRILLO I MARTÍNEZ, A. (2020): "La apertura de datos públicos: un marco normativo en continua evolución", *Consultor de los ayuntamientos y de los juzgados: Revista técnica especializada en administración local y justicia municipal*, nº 3, 21-34.
- CERRILLO I MARTÍNEZ, A. y CASADESÚS DE MINGO, A. (2021): "Data governance for public transparency", *El profesional de la información*, nº 30(4). doi:https://doi.org/10.3145/epi.2021.jul.02
- CERRILLO I MARTÍNEZ, A. y GALÁN GALÁN, A. (2006): *La reutilización de la información del sector público*, Granada, Comares.
- COMISIÓN EUROPEA (2018), «Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones de 25 de abril de 2018 *Hacia un espacio común europeo de datos* [COM(2018) 232 final]»,
- COTINO HUESO, L. (2020): "Ética, valores y principios del «open data» y los retos futuros de la apertura de datos públicos", *Consultor de los ayuntamientos y de los juzgados: Revista técnica especializada en administración local y justicia municipal*, nº 3, 147-166.
- FERNÁNDEZ SALMERÓN, M. (2022): "Acceso a la información pública y datos abiertos: A vueltas con el ámbito objetivo de la disciplina sobre reutilización y su armonización con la normativa sobre transparencia", en VALERO TORRIJOS, J. y MARTÍNEZ GUTIÉRREZ, R. (eds.), *Datos abiertos y reutilización de la información del sector público* (pp. 1-20), Granada, Comares.

- GÓMEZ MANRESA, M. F. (2022): "Acuerdos exclusivos y reutilización de la información del sector público", en VALERO TORRIJOS, J. y MARTÍNEZ GUTIÉRREZ, R. (eds.), *Datos abiertos y reutilización de la información del sector público* (pp. 187-213), Granada, Comares.
- HINTZ, A., DENCİK, L. y WAHL-JORGENSEN, K. (2019): *Digital citizenship in a datafied society*, Cambridge, Polity Press.
- JANSSEN, K. y DUMORTIER, J. (2003): "Towards a European framework for the re-use of public sector information: a long and winding road", *International Journal of Law and Information Technology*, n° 11(2).
- JANSSEN, M., BROUS, P., ESTEVEZ, E., BARBOSA, L. S. y JANOWSKI, T. (2020): "Data governance: Organizing data for trustworthy Artificial Intelligence", *Government Information Quarterly*, n° 37(3), 101493. doi:<https://doi.org/10.1016/j.giq.2020.101493>
- LÓPEZ-TARRUELLA MARTÍNEZ, A. (2022): "La necesidad de una licencia gubernamental abierta en el ordenamiento español para favorecer el acceso a los datos", en VALERO TORRIJOS, J. y MARTÍNEZ GUTIÉRREZ, R. (eds.), *Datos abiertos y reutilización de la información del sector público* (pp. 159-186), Granada, Comares.
- MARTÍNEZ GUTIÉRREZ, R. (2014): "La responsabilidad patrimonial en el acceso, la difusión y la reutilización de la información", en VALERO TORRIJOS, J. y FERNÁNDEZ SALMERÓN, M. (eds.), *Régimen jurídico de la transparencia del sector público: del Derecho de acceso a la reutilización de la información* (pp. 427-458), Cizur Menor, Thomson-Aranzadi.
- MAYER-SCHÖNBERGER, V. y CUKIER, K. (2013): *Big data. La revolución de los datos masivos*, Madrid, Turner.
- NOVECK, B. (2011): "What's in a Name? Open Gov and Good Gov." *Huffington Post*.
- OECD (2008): "Recommendation of the Council for enhanced access and more effective use of public sector information [C(2008)36]".
- SÁNCHEZ GARCÍA, A. (2022): "El valor económico de la información pública y la regulación del precio de su reutilización", en VALERO TORRIJOS, J. y MARTÍNEZ GUTIÉRREZ, R. (eds.), *Datos abiertos y reutilización de la información del sector público* (pp. 215-252), Granada, Comares.
- THE ECONOMIST (2017): "Data is giving rise to a new economy. How is it shaping up?", *The Economist*.
- VALERO TORRIJOS, J. (2022): "El régimen jurídico de los datos de alto valor: dificultades y retos para su utilización", en VALERO TORRIJOS, J. y MARTÍNEZ GUTIÉRREZ, R. (eds.), *Datos abiertos y reutilización de la información del sector público* (pp. 81-102), Granada, Comares.

Capítulo XI

La contratación de soluciones de inteligencia artificial¹

ISABEL GALLEGO CÓRCOLES²

*Catedrática de Derecho administrativo
Universidad de Castilla-La Mancha*

Resumen: La contratación de soluciones de inteligencia artificial entraña diferencias sustanciales respecto a la adquisición de otros objetos contractuales. Por un lado, la contratación pública tiene gran potencial en este ámbito como instrumento de estímulo a la innovación. Por otro, debe garantizar que el uso de la IA respete escrupulosamente los derechos y libertades de la ciudadanía. A lo largo de este trabajo se reflexiona sobre las pautas necesarias para articular con éxito ambas funciones tanto en la fase del diseño y preparación de la licitación como en la de inclusión de determinado contenido obligatorio en el contrato.

Abstract: The procurement of artificial intelligence solutions entails substantial differences compared to the procurement of other contractual objects. On the one hand, public procurement in this area can be an instrument to stimulate innovation. On the other hand, it must ensure that the use of AI respects the rights and freedoms of citizens. The aim of this paper is to provide guidelines on how to articulate both functions, both in the design and preparation phase of the tender and in the inclusion of certain mandatory content in the contract.

Palabras-clave: INTELIGENCIA ARTIFICIAL, CONTRATACIÓN PÚBLICA, INNOVACIÓN, TRANSPARENCIA

Key-words: AI, PUBLIC PROCUREMENT, INNOVATION, TRANSPARENCY.

SUMARIO.- 1.- LA CONTRATACIÓN PÚBLICA COMO IMPULSOR Y GARANTE DEL USO DE SOLUCIONES BASADAS EN INTELIGENCIA ARTIFICIAL.-2.- CUESTIONES CLAVE EN EL DISEÑO Y PREPARACIÓN DEL PROCEDIMIENTO DE

¹ El presente trabajo se ha realizado en el marco del proyecto de investigación INNOVACION PARA UNA SALUD DE VANGUARDIA: COMPRA PUBLICA, TECNOLOGIA, SOSTENIBILIDAD AMBIENTAL Y FACTORES SOCIOECONOMICOS (IN-SALVAN) (PID2021-127828NB-I00) financiado por el Ministerio de Ciencia e Innovación y del proyecto “Administración Pública e Inteligencia Artificial: regulación y uso de la AI en el ámbito de la contratación pública” (TED2021-130682B-I00), financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea NextGenerationEU/PRTR

² ORCID: 0000-0003-4018-3115.

CONTRATACIÓN.- 2.1.- La conformación del equipo que impulsará la licitación.- 2.2.- La determinación de las necesidades a satisfacer y la delimitación del objeto del contrato: ¿cuándo adquirir una solución de inteligencia artificial?.- 2.3.- La elección del procedimiento de licitación.- 2.4.- Algunos aspectos a valorar en las ofertas.- 2.5.- La adecuada delimitación del equipo de trabajo: la necesaria incorporación de juristas.- 2.6.- La necesidad de prever posibles incidencias en la ejecución.- 3.-ASPECTOS RELEVANTES DEL CONTENIDO DEL CONTRATO.- 3.1.- La necesidad de cumplir las exigencias de transparencia, motivación y libre competencia: ¿debe exigirse la cesión de la titularidad del código fuente?.- 3.2.- Mecanismos para garantizar la supervisión, auditoría y control de los sistemas algorítmicos.- 3.3.- Cláusulas relativas a la protección de datos.-4.-CONCLUSIONES.- 5.- BIBLIOGRAFÍA CITADA

1.- LA CONTRATACIÓN PÚBLICA COMO IMPULSOR Y GARANTE DEL USO DE SOLUCIONES BASADAS EN INTELIGENCIA ARTIFICIAL

Como es sobradamente conocido, el empleo de la inteligencia artificial por parte de las Administraciones Públicas suscita importantes retos jurídicos, derivados de la necesaria adaptación de las técnicas de protección de derechos y libertades a una nueva realidad tecnológica con especial capacidad para vulnerar estos. Sobre el elenco garantías esenciales que ha de respetar la incorporación de soluciones de IA a la actuación pública existe un importante consenso doctrinal³ y recientemente también, cierto reconocimiento normativo⁴. Desde esta pers-

³ Sin ánimo de exhaustividad, CERRILLO MARTÍNEZ, A. (2019): “El impacto de la inteligencia artificial en el Derecho administrativo ¿Nuevos conceptos para nuevas realidades técnicas?”, *Revista General de Derecho Administrativo*, n. 50; VALERO TORRIJOS, J. (2019): “Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración”, *Revista Catalana de Dret Public*, n° 58, HUERGO LORA, A. (2020): “Una aproximación a los algoritmos desde el Derecho Administrativo”, en HUERGO LORA, A. (Dir.): *La regulación de los algoritmos*, Thomson-Aranzadi, Madrid, pp. 23-87; GAMERO CASADO, E. (2021): “Compliance (o cumplimiento normativo) de desarrollos de Inteligencia Artificial para la toma de decisiones administrativas”, en *Diario La Ley*, n° 50, 19 de abril de 2021; MARTÍN DELGADO, I. (2022): “La aplicación del principio de transparencia a la actividad administrativa algorítmica”, en esta misma obra,

⁴ La reciente Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación, dispone que en el marco de la Estrategia Nacional de Inteligencia

pectiva, en la medida en la que gran parte del desarrollo de soluciones de inteligencia artificial que incorpore la Administración pública será inevitablemente privado, y a falta de una regulación específica del uso de la IA, la contratación pública debe adoptar el papel de integración las distintas garantías que se perfilan en el Ordenamiento jurídico⁵, actuando como garante de la adquisición una inteligencia artificial centrada en las personas⁶. Todo ello sin descuidar su tradicional función de protección del correcto funcionamiento del mercado⁷.

Por otro lado, dado que la demanda pública tiene un impacto significativo en este sector⁸, la contratación pública está llamada a desempeñar

Artificial, de la Carta de Derechos Digitales y de las iniciativas europeas en torno a la Inteligencia Artificial, las administraciones públicas favorecerán la puesta en marcha de mecanismos para que los algoritmos involucrados en la toma de decisiones que se utilicen en las administraciones públicas tengan en cuenta criterios de minimización de sesgos, transparencia y rendición de cuentas, siempre que sea factible técnicamente. En estos mecanismos se incluirán su diseño y datos de entrenamiento, y abordarán su potencial impacto discriminatorio. Para lograr este fin, se promoverá la realización de evaluaciones de impacto que determinen el posible sesgo discriminatorio. Añade el precepto, que las administraciones públicas, en el marco de sus competencias en el ámbito de los algoritmos involucrados en procesos de toma de decisiones, priorizarán la transparencia en el diseño y la implementación y la capacidad de interpretación de las decisiones adoptadas por los mismos.

⁵ Vid. MIRANZO DÍAZ, J. (2020): “Inteligencia artificial y contratación pública” en MARTÍN DELGADO, I. / MORENO MOLINA, J.A. (Dirs.): *Administración electrónica, transparencia y contratación pública*, Iustel, Madrid, pp.105-142.

⁶ Tomo la expresión del trabajo OLUKA NAGITTA ,P./ ADOKO OBICCI, P./ AWOOR, E. (2022): “Human-centered artificial intelligence for the public sector: The gate keeping role of the public procurement professional”, *Procedia Computer Science*, n.º 200(1), pp. 1084 ss. Vid. igualmente SÁNCHEZ GRAELLS, A. (2022): “Protecting procurement’s AI gatekeeping role in domestic law, and trade agreements? en www.howtocrackanut.com.

⁷ Como señala VALERO TORRIJOS, J. (2021), “Inteligencia artificial y contratación pública”, *Observatorio de Contratación Pública*, www.obcp.es [última visita: 21/06/2022] se ha garantizar el servicio del interés general – evitando la cautividad de las Administraciones Públicas– la intensa colaboración público-privada en el desarrollo de IA y creación, permitiendo la igualdad de los concurrentes en la contratación para satisfacer principios tales como la transparencia y no discriminación en el diseño.

⁸ Así, en las Conclusiones del Consejo «Inversión pública a través de la contratación pública: recuperación sostenible y reactivación de una economía de la

un papel de impulsor y ordenador del mercado de IA. Dicho en otros términos, el uso estratégico de la contratación pública resulta insoslayable⁹, dada su capacidad para estimular y desarrollar el ecosistema de inteligencia artificial y el sector privado¹⁰. En este sentido, la contratación pública es una de las herramientas que posee el poder público para financiar la innovación y garantizar una IA fiable¹¹. Así, la Estrategia Nacional de Inteligencia Artificial¹², prevé que la Administración pública pueda “actuar como demandante de soluciones tecnológicas, impulsando el desarrollo de soluciones tecnológicas, impulsando así el desarrollo de nuevos programas y capacidades por el sector privado”¹³. Además,

UE resiliente» (2020/C 412 I/01), se pide los Estados miembros que elaboren políticas y estrategias de compra integrales que, en particular, deben centrarse en sectores en los que la demanda pública tiene un impacto significativo, como los medicamentos, los productos sanitarios o las tecnologías de la información, teniendo asimismo en cuenta aspectos relativos a la resiliencia, la gestión de riesgos y la seguridad del suministro.

⁹ Sobre el uso estratégico de la contratación pública, vid. por todos, GIMENO FELIU, J.M. (2014): *El nuevo paquete legislativo comunitario sobre contratación pública. El contrato público como herramienta del liderazgo institucional de los poderes públicos*, Aranzadi, Cizur Menor. MORENO MOLINA J.A. (2018): *Hacia una compra pública responsable y sostenible: novedades principales de la Ley de Contratos del Sector Público 9/2017*, Tirant lo Blanch, Valencia.

Igualmente es de interés GIMENO FELIU, J.M. (2018): “La contratación pública como estrategia: el fomento de la innovación”, *Revista Ius Publicum*, 1.

En relación con un mercado tecnológico concreto, vid. GIMENO FELIU, J.M./GARCÍA-ÁLVAREZ GARCÍA, G./TEJEDOR BIELSA, J. (2022): *La compra pública de tecnología en la nube en España*, Aranzadi, pp.123-126.

¹⁰ McBRIDE, K./ VAN NOORDT, C./ MISURACA, G./ HAMMERSCHMID, G. (2021): “Towards a Systematic Understanding...”, *op. cit.*, pp.11.

¹¹ GRUPO INDEPENDIENTE DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL (2019): *Policy and Investment Recommendations for Trustworthy AI*, p.20.

¹² Dentro del eje estratégico 5, “potenciar el uso de la IA en la Administración Pública y en las misiones estratégicas nacionales.”

¹³ Como se señala en el citado documento, “las necesidades de Inteligencia Artificial son numerosas en el sector público, en ámbitos como empleo, sanidad, justicia, migraciones; y su satisfacción permitirá a su vez la innovación en Inteligencia Artificial. Para ello, se podrán utilizar instrumentos de contratación, tales como la asociación para la innovación, en virtud de la cual un órgano de la Administración Pública plantea un objetivo o problema, dejando abierta la forma en que las empresas proporcionarán soluciones y, por tanto, impulsando la innovación privada en ámbitos en que no exista un sistema o procedimiento ya probado para abordar la cuestión. Para ello se tendrán en cuenta las orientacio-

las normas sobre contratación pública y las prácticas en este ámbito tienen una gran influencia en los mercados, especialmente en sus primeras etapas de desarrollo. Así se destaca en las *Orientaciones para la Compra de IA (Guideliness for AI Procurement)*¹⁴ del Foro Económico Mundial. Como se señala en este documento, la industria debate en este momento establecer sus propios estándares respecto a estas tecnologías; la autoridad institucional del gobierno y su credibilidad pueden ayudar a establecer unas líneas básicas en este ámbito.

En todo caso, la adquisición de soluciones de inteligencia artificial no es una adquisición de tecnología tradicional. Como se ha subrayado¹⁵, existen una serie de retos vinculados a la compra pública de inteligencia artificial vinculados a: a) la necesidad de que se usen datos a través del *machine learning*, con las implicaciones que ello tiene respecto al cumplimiento de la normativa sobre protección de datos; b) propiedad de los datos; c) potenciales sesgos de la inteligencia artificial; d) los derechos de propiedad intelectual y los secretos comerciales y su relación con la transparencia de los algoritmos. De ahí que a nivel internacional existan distintas iniciativas que tratan de establecer una serie de orientaciones respecto de la adquisición de soluciones basadas en la inteligencia artificial.

En el ámbito geográfico estrictamente europeo, los primeros Estados que han adoptado guías propias sobre compra pública de

nes del Digital Europe Programme, así como el seguimiento a la implementación del Libro Blanco sobre la Inteligencia Artificial, especialmente en lo que se refiere a la adopción de la IA en el sector público”.

¹⁴ El documento, fechado en 2019 está disponible en <https://www3.weforum.org>. Con un contenido similar vid. también del Foro Económico Mundial *AI Procurement in a Box: AI Government Procurement Guidenes, Toolkit*, Junio de 2020. En estos documentos ha colaborado la Oficina para la Inteligencia Artificial de Reino Unido. Elaborado por esta resulta de interés *Guidelines for AI procurement. A summary of best practices addressing specific challenges of acquiring Aritificial intelligence in public sector* (junio, 2020). Se trata de un documento breve, pero con múltiples referencias a documentación de interés. Disponible en <https://www.gov.uk>

¹⁵ McBRIDE, K./ VAN NOORDT, C./ MISURACA, G./ HAMMERSCHMID, G. (2021): “Towards a Systematic Understanding on the Challenges of Procuring Artificial Intelligence in the Public Sector”. <https://doi.org/10.31235/osf.io/un649>

inteligencia artificial han sido Estonia¹⁶ y Reino Unido¹⁷. De especial interés son, sin embargo, las “cláusulas tipo para la adquisición de sistemas algorítmicos fiables” desarrolladas en su día por la Ciudad de Ámsterdam (en adelante, *Cláusulas de la Ciudad de Ámsterdam*)¹⁸, que describen las obligaciones que se imponen a los proveedores para garantizar un uso seguro y responsable de los algoritmos. En este sentido, abordan los aspectos más importantes en el diseño de la compra de algoritmos: gestión del riesgo, calidad de los datos, transparencia, titularidad de los datos y auditorías. Sin alcanzar ni mucho menos este nivel de desarrollo, en nuestro entorno, en la “Estrategia municipal de algoritmos y datos para el impulso ético de la inteligencia artificial” del Ayuntamiento de Barcelona se prevé que las licitaciones públicas de sistemas o soluciones basados en IA deberán garantizar que, además de cumplir con las normativas relativas a la contratación pública, cumplen también con los valores éticos expuestos en esta estrategia, así como las normas/reglamentos/instrucciones que de esta se deriven. Dentro de las actuaciones previstas para el periodo 2021-2023, la actuación novena se refiere a “cláusulas para la licitación de soluciones basadas en IA”¹⁹.

¹⁶ Documentación disponible, solo parcialmente en inglés, en Instructional materials | Krattide veebileht (kratid.ee).

¹⁷ *Guidelines for AI procurement* (2020), disponible en Guidelines for AI procurement–GOV.UK (www.gov.uk).

¹⁸ Las cláusulas tipo del Ayuntamiento de Ámsterdam están disponibles en <https://www.amsterdam.nl/innovatie/digitalisering-technologie/algoritmen-ai/contractual-terms-for-algorithms/>. A través de dichas cláusulas tipo se describe la información que se solicitará a los proveedores de forma que se garantice un uso seguro y responsable de los algoritmos.

¹⁹ Como se señala en el documento, la contratación pública es uno de los instrumentos clave para que el Ayuntamiento pueda no solo regular la aplicación de la IA en el ámbito municipal, sino influir en el comportamiento de actores externos. Por ello, resulta fundamental incluir cláusulas relacionadas con los derechos digitales en la licitación de soluciones basadas en IA. Con el fin de facilitar el trabajo de los cuerpos técnicos y directivos a la hora de redactar los contratos públicos, como complemento de la guía de aplicación de la IA en el Ayuntamiento de Barcelona, se elaborarán cláusulas de contratación pública normalizadas para asegurar el cumplimiento de los principios rectores descritos en esta medida de gobierno.

Por su parte, en el ámbito del Derecho de la Unión Europea, dentro de la Acción 6 del Libro Blanco sobre la inteligencia artificial²⁰ se prevé que la Comisión prepare un “Programa de adopción de la IA” específico que respaldará la contratación pública de sistemas de inteligencia artificial, y ayudará a transformar los propios procesos de esta contratación²¹. Si bien ello aparentemente estará más ligado con el papel de impulsor de la contratación de inteligencia artificial, también la Comisión europea está auspiciando una iniciativa²² para garantizar la adquisición de soluciones de IA fiable. En este sentido, el proyecto consiste en adaptar las ya citadas Cláusulas de la Ciudad de Ámsterdam para convertirlas en un estándar europeo que esté alineado con la propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial, en adelante RIA)²³. Como se señala en el documento, a pesar de que la propuesta de Ley de IA

²⁰ COMISIÓN EUROPEA, *Libro Blanco sobre la inteligencia artificial-un enfoque europeo orientado a la excelencia y la confianza*, Bruselas, 19.2.2020, COM(2020) 65 final.

²¹ En concreto, la Comisión iniciará conversaciones por sector abiertas y transparentes, en las que dará prioridad a la atención sanitaria, las administraciones rurales y los operadores de servicios públicos, para presentar un plan de acción que facilite el desarrollo, la experimentación y la adopción de la inteligencia artificial. Las conversaciones por sector se emplearán para preparar un «Programa de adopción de la IA» específico que respaldará la contratación pública de sistemas de inteligencia artificial, y ayudará a transformar los propios procesos de esta contratación.

²² La *Proposal for standard contractual clauses for the procurement of artificial intelligence by public organisations version 0.9* está disponible en <https://living-in.eu/groups/solutions/ai-procurement> última visita: 21/06/2022].

²³ La propuesta establece una serie de obligaciones de los proveedores de sistemas de IA de alto riesgo. En concreto, los proveedores de sistemas de IA de alto riesgo:

- a) velarán por que sus sistemas de IA de alto riesgo cumplan los requisitos definidos en la norma para estos sistemas: sistemas gestión de riesgos, datos y gobernanza de datos, documentación técnica, registros, transparencia y comunicación de información a los usuarios, vigilancia humana, precisión, solidez y ciberseguridad,
- b) contarán con un sistema de gestión de la
- c) elaborarán la documentación técnica del sistema de IA de alto riesgo;
- d) cuando estén bajo su control, conservarán los archivos de registro que sus sistemas de IA de alto riesgo generen automáticamente;

es objeto de negociaciones en curso y el texto de la Ley de IA no es definitivo, contiene las mejores prácticas en el ámbito de la regulación de la IA que pueden utilizarse en la contratación de IA.

Partiendo de este contexto, en este trabajo –a la luz de las guías y orientaciones que se acaban de reseñar y de las principales conclusiones de nuestra doctrina científica– se intentan dar algunas pautas que permitan guiar los procedimientos de contratación tanto desde la perspectiva de lograr un uso estratégico de la contratación –en la que la compra pública adopta un papel de instrumento catalizador de la IA– como desde la vertiente de la garantía de derechos y libertades individuales. Para ello, se establecerán algunas consideraciones primero en relación con el diseño y preparación de la licitación y posteriormente se examinarán las cuestiones relativas al contenido obligatorio del contrato que han de plasmar los pliegos de condiciones. Consideraciones que en todo caso deberán modularse, atendiendo al principio de proporcionalidad, en función del riesgo potencial de la concreta solución de IA que se pretenda adquirir.

En las conclusiones finales se reflexionará sobre si estas pautas pueden sustituir a una regulación *ad hoc* del uso de la IA por la Administración pública.

-
- e) se asegurarán de que los sistemas de IA de alto riesgo sean sometidos al procedimiento de evaluación de la conformidad oportuno antes de su introducción en el mercado o puesta en servicio;
 - f) cumplirán las obligaciones de registro que prevé la norma;
 - g) adoptarán las medidas correctoras necesarias cuando un sistema de IA de alto riesgo no cumpla los requisitos establecidos en el proyecto;
 - h) informarán a las autoridades nacionales competentes del Estado miembro donde hayan comercializado o puesto en servicio el sistema de IA y, en su caso, al organismo notificado de los casos de no conformidad y de las medidas correctoras adoptadas;
 - i) colocarán el marcado CE en sus sistemas de IA de alto riesgo para indicar que cumplen lo dispuesto en el presente Reglamento.
 - j) demostrarán, a solicitud de la autoridad nacional competente, que sus sistemas de IA de alto riesgo cumplen los requisitos establecidos.
- Sobre esta propuesta, vid. los trabajos de GÓMEZ GUTIÉRREZ, E. y HUERGO LORA, A., en esta misma obra.

2.- CUESTIONES CLAVE EN EL DISEÑO Y PREPARACIÓN DEL PROCEDIMIENTO DE CONTRATACIÓN

2.1.- La conformación del equipo que impulsará la licitación

Un elemento fundamental para que la contratación de una solución de IA se desarrolle con éxito es que se perfile de forma adecuada el equipo que llevará a cabo el proyecto en el seno de la propia Administración compradora.

En este sentido, resulta fundamental que las ideas se desarrollen y las decisiones se adopten por parte de un equipo multidisciplinar²⁴: estos equipos deben tener conocimientos sólidos en las políticas relativas al área o campo (por ejemplo, justicia) en el que se aplicará la solución de IA, además incorporar el conocimiento necesario en aprendizaje automático / ciencia de datos; ingeniería de datos; tecnología (software y hardware); compras; derecho²⁵. Aunque durante todas las fases de la licitación quizá no sea necesaria la participación de todos estos profesionales, resultará una buena práctica tener prevista la composición del equipo desde un principio.

Por lo demás, se ha señalado que esta visión multidisciplinar es clave cuando llegue el momento de evaluar las ofertas recibidas en respuesta a la convocatoria²⁶, si bien en nuestro Derecho deberán tenerse en cuenta las limitaciones que establece la LCSP. Así, no podrán formar parte de la mesa de contratación el personal que haya participado en la redacción de la documentación técnica del contrato [art.326 LCSP]²⁷. No obstante, los informes de valoración de las

²⁴ WORLD ECONOMIC FORUM, *Guidelines for AI Procurement*, p. 10.

²⁵ *Ibidem*.

²⁶ WORLD ECONOMIC FORUM, *Guidelines for AI Procurement*, p. 11.

²⁷ Salvo en los supuestos a que se refiere la Disposición adicional segunda, relativa a la contratación de las Corporaciones Locales.

En relación con la posibilidad de que el personal de la unidad promotora forme parte del órgano de contratación, es de interés el Informe JCCA del Estado 3/2018, que tiene ocasión de afirmar que no cabe establecer apriorísticamente soluciones a todos los casos concretos de cada órgano de contratación, puesto que la medida de la intervención de cada empleado público en la redacción del

proposiciones podrán ser emitidos por las personas que hayan participado en la redacción de la documentación técnica del contrato²⁸. Además, la mesa podrá contar con el asesoramiento de técnicos o expertos independientes con conocimientos acreditados en las materias relacionadas con el objeto del contrato [art.326.5 LCSP]. Finalmente, deberá tenerse en cuenta que en los procedimientos abiertos y restringidos, la valoración de los criterios cuya cuantificación dependa de un juicio de valor corresponderá, en los casos en que proceda por tener atribuida una ponderación mayor que la correspondiente a los criterios evaluables de forma automática, a un comité formado por expertos con cualificación apropiada o a un organismo técnico especializado²⁹.

De esta forma, la trascendencia de una adecuada conformación del equipo por parte de la Administración pública es esencial, dado el amplio espectro de ámbitos del saber que están implicados en la implantación de este tipo de soluciones. La adecuada composición de los equipos es un reto al que deben enfrentarse las Administraciones más innovadoras. Como se ha señalado, la transformación digital de la Administración precisa de una modalidad diferente de selección, organización y gestión de los recursos humanos³⁰. Todo ello sin perjuicio de que, como se ha destacado, en los casos en los que la Administración carezca de perfiles profesionales de este tipo, pueda resultar

pliego técnico puede variar según el caso de que se trate. De este modo, un mero conocimiento de las condiciones del pliego no generaría la prohibición de formar parte de la mesa y, sin embargo, una participación efectiva en la confección del mismo, dando instrucciones o haciendo observaciones de obligado cumplimiento, sí vedaría la participación en el órgano de asistencia. Consecuentemente no cabe decir en términos generales si el personal de la unidad proponente del contrato puede o no participar en la mesa de contratación, pues tal circunstancia dependerá de en qué medida haya intervenido realmente en la redacción del pliego técnico.

²⁸ Vid. Informe JCCA del Estado 3/2018.

²⁹ Vid. Art.146.2 a) LCSP. Sobre la interpretación de este precepto es de interés el Informe JCCA del Estado 11/2020.

³⁰ MARTÍN DELGADO, I. (2022): “La aplicación del principio...”, *op. cit.*, y bibliografía allí citada.

aconsejable acudir a la externalización del servicio de asesoramiento técnico en la contratación pública de estas soluciones³¹.

2.2.- La determinación de las necesidades a satisfacer y la delimitación del objeto del contrato: ¿cuándo adquirir una solución de inteligencia artificial?

2.2.1.- Planteamiento

Aunque pueda parecer una obviedad, una de las decisiones más trascendentes dentro del proceso de licitación es la relativa a determinar si la IA es la solución para resolver alguna de las necesidades que ha detectado la Administración pública. Para su adopción han de tenerse en cuenta una serie de factores complejos, que no están presentes en otro tipo de compras públicas. Hasta tal punto es así, que en algunos países como Estonia o Reino Unido se ha tratado de guiar esta decisión a través de la publicación de distintos cuestionarios³².

En todo caso, la contratación de soluciones de inteligencia artificial requiere que previamente se definan las necesidades a satisfacer [art.1 LCSP] de forma que el objeto del contrato resulte idóneo para cubrir dichas necesidades [art.28 LCSP]. Así, la elaboración de la documentación ha de realizarse con especial esmero por parte del

³¹ Vid. GAMERO CASADO, E. (2022): “Supervisión, auditoría y control jurídico en la contratación pública de soluciones de robotización e inteligencia artificial para soporte a la toma de decisiones”, *www.obcp.es.*, para quien esta opción implica realizar un control externo a la empresa desarrolladora, compatible con la exigencia de juristas en los equipos de trabajo de esta última, y que podría ser el mejor modelo de desarrollo de estas soluciones, al fortalecer la posición de la Administración y diferenciar entre un control, auditoría o supervisión interno (de la propia empresa contratista sobre su sistema) y otro externo, independiente y objetivo (el de la Administración pública, realizado a su vez por una empresa asesora cualificada).

³² McBRIDE, K./ VAN NOORDT, C./ MISURACA, G./ HAMMERSCHMID, G. (2021): “Towards a Systematic Understanding on the Challenges of Procuring Artificial Intelligence in the Public Sector”. <https://doi.org/10.31235/osf.io/un649>, p.11. Vid., por ejemplo, en Reino Unido la Guía *Assesing if artificial intelligence is the right solution*, (2019), disponible en <https://www.gov.uk/guidance/assessing-if-artificial-intelligence-is-the-right-solution>.

órgano de contratación. Debe recordarse que en el expediente se justificará adecuadamente “la necesidad de la Administración a la que se pretende dar satisfacción mediante la contratación de las prestaciones correspondientes; y su relación con el objeto del contrato, que deberá ser directa, clara y proporcional”³³. Para que exista esta relación será preciso tanto conocer de forma precisa el mercado como evaluar de forma preliminar los riesgos que entraña la adopción de la solución basada en inteligencia artificial, así como verificar la disponibilidad de datos de calidad y perfilar cuestiones relativas a su gobernanza.

Solo tras la realización de estos pasos -lo que podrá hacerse de manera simultánea- se estará en condiciones de elaborar la memoria justificativa del contrato (art.116 LCSP), que será objeto de publicidad a través del perfil de contratante (art.63 LCSP). Completado el expediente de contratación, se dictará resolución motivada por el órgano de contratación aprobando el mismo y disponiendo la apertura del procedimiento de adjudicación. Esta resolución deberá ser objeto de publicación en el perfil de contratante (art.117 LCSP). De esta forma, contra las distintas decisiones que integran el procedimiento de contratación, especialmente en esta fase inicial, podrá ejercitarse los distintos derechos de impugnación³⁴, bien entendido que será en todo caso preciso que exista legitimación por parte del recurrente³⁵.

³³ Art116.4.d) LCSP.

³⁴ VALERO TORRIJOS, J. (2019): “Las garantías jurídicas..”, *op. cit.*, p.87, ha insistido en que la decisión de poner en funcionamiento un sistema de inteligencia artificial debería ser adoptada formalmente mediante un acto administrativo, requerimiento que constituye no solo una exigencia del ejercicio de la competencia y su carácter irrenunciable, sino que, además, es una condición inexcusable para facilitar la adecuada defensa de los derechos e intereses legítimos, ya que, en definitiva, de no ser así, se estaría dificultando una eventual impugnación de la decisión administrativa de poner en marcha un concreto sistema de inteligencia artificial. Pues bien, en los casos de externalización de la solución de inteligencia artificial es posible que se ejerza el control desde el momento inicial de la toma de decisión. En todo caso, de cara a una eventual -y deseable- regulación del uso de la IA en la Administración, deberá coordinarse los distintos procedimientos -eventual aprobación del sistema/contratación- de tal forma que los procedimientos no se dilaten más de lo necesario.

³⁵ La legitimación solo exige la presencia de un interés legítimo, que no obstante se interpreta en un sentido amplio. Sirva como ejemplo la *Resolución 45/2017 Órgano Administrativo de Recursos Contractuales de la Comunidad Autónoma de Euskadi*, que admitió la legitimación de un sindicato para recurrir los pliegos

2.2.2.- El conocimiento del mercado: la posibilidad de realizar consultas preliminares

Como en cualquier tipo de compra vinculada a la innovación³⁶, resultará útil un acercamiento previo al mercado a fin de conocer las distintas posibilidades que este ofrece. De hecho, en el ámbito de la contratación de soluciones de inteligencia artificial, el conocimiento del mercado es un elemento clave para el éxito del proyecto³⁷. Desde esta perspectiva, la realización de consultas preliminares del mercado, reguladas en el art.115 LCSP, permitirán recabar información de interés sobre las posibilidades que ofrece la tecnología para la satisfacción de la necesidad pública.

En efecto, la contratación de innovación exige una buena preparación por parte de comprador público; con la finalidad de adquirir un mejor conocimiento del mercado en cuestión, los compradores públicos pueden ponerse en contacto con los posibles proveedores

que regían la contratación de un servicio de limpieza, por entender que contradice los acuerdos previos con los sindicatos en los que la entidad contratante se comprometía a no externalizar nuevos servicios.

³⁶ En este sentido, se ha señalado que las consultas preliminares de mercado resultan esenciales en los supuestos de compra pública de innovación. Vid. VALCÁRCEL, P. (2018): “La consultas preliminares del mercado como mecanismo para favorecer las “compras inteligentes”, *Revista Española de Derecho Administrativo*, núm. 191 [consultada versión electrónica], MIÑO LÓPEZ, A. (2017): “Preliminary market consultations in innovation procurement: a principled approach and incentives for anticompetitive behaviours”, *Ius publicum*, DE GUERRERO MAN-SO, C. (2018): “Las consultas preliminares del mercado: una herramienta para mejorar la eficiencia en la contratación” en GIMENO FELIU, J.M. (dir.): *Estudio sistemático de la ley de contratos del sector público*, Aranzadi, pp.1047-1071.

La innovación se define normativamente como “la introducción de un producto, servicio o proceso nuevos o significativamente mejorados, que incluye, aunque no se limita a ellos, los procesos de producción, edificación o construcción, un nuevo método de comercialización o un nuevo método de organización de prácticas empresariales, la organización del lugar de trabajo o las relaciones exteriores, entre otros con el objetivo de ayudar a resolver desafíos de la sociedad o a apoyar la Estrategia Europa 2020 para un crecimiento inteligente, sostenible e integrador” (art.2.1.22 Directiva 2014/24/UE del Parlamento Europeo y del Consejo, de 26 de febrero de 2014, sobre contratación pública y por la que se deroga la Directiva 2004/18/CE)

³⁷ McBRIDE, K./ VAN NOORDT, C./ MISURACA, G./ HAMMERSCHMID, G. (2021): “Towards a Systematic Understanding...”, *op. cit.*, pp. 11.

en consultas previas del mercado, que pueden servir para recabar información, en particular sobre la estructura de los precios y las capacidades del mercado. Por lo demás, las consultas preliminares pueden ser útiles tanto para comprobar la tecnología disponible antes de redactar especificaciones descriptivas (evitando así favorecer soluciones concretas) como para establecer exigencias funcionales ambiciosas y realistas³⁸. También para conocer la estructura del mercado, de forma en función de su madurez pueda seleccionarse el procedimiento de licitación más adecuado³⁹. El aprendizaje obtenido de las consultas ayudará a definir mejor el planteamiento del problema y puede ser útil para determinar la viabilidad de la IA como solución a la necesidad pública⁴⁰. Además, aunque no sea esta la finalidad principal del art.115 LCSP, la consulta preliminar puede permitir de forma simultánea dar conocimiento a la ciudadanía del proyecto, instrumentalizándose como un cauce más amplio de

³⁸ Anuncio de la Comisión Europea “Orientaciones sobre la contratación pública en materia de innovación”, C(2018) 3051 final, Bruselas 15.2.2018, pp. 33-34. En todo caso, el empleo de especificaciones descriptivas no es el más aconsejable en el ámbito de la innovación. Como señalan estas orientaciones, las especificaciones técnicas descriptivas son más apropiadas cuando el poder adjudicador conoce a la perfección el potencial del mercado. Incluso en esos casos, dejar una parte del rendimiento abierto al proceso de innovación podría ayudar a obtener el resultado requerido.

³⁹ GALLEGO CÓRCOLES, I. (2019): “Las consultas preliminares del mercado: algunas reflexiones”, en GIMENO FELIU, J.M. (dir.): *Observatorio de los contratos públicos 2018*, Aranzadi, pp. 131.180.

⁴⁰ Por ejemplo, en la británica *Guidelines for AI procurement*, p.26. En el mismo sentido, la *Guidelines for AI Procurement* del Foro Económico Mundial es muy expresiva: “pre-market engagement is also often essential in helping you to describe your problem and narrow down the task that AI may be able to assist with”. (p. 8)

Las consultas preliminares de mercado son relativamente frecuentes en la práctica española. En relación con la adquisición de tecnología, sirva como ejemplo la consulta pública del Ayuntamiento de Barcelona para el Desarrollo de un Sistema Gestión de Big Data Social (Lote 1) e Implantación de Mecanismos de Lenguaje Natural e Inteligencia Artificial en la gestión del Área de Derechos Sociales. Vid. la información relativa a la misma en https://contractaciopublica.gencat.cat/ecofin_pscp/AppJava/es_ES/notice.pscp?idDoc=39996639&reqCode=viewCtn&

participación⁴¹. Así, en principio, la LCSP no cierra el círculo de los posibles interesados a los que dirigir “consultas previas” a la licitación. De esta forma, la consulta a los destinatarios de la prestación puede resultar extremadamente útil para perfilar con mayor precisión las necesidades del órgano de contratación⁴². Bien es verdad que este tipo de consultas, cuando tengan como destinatario exclusivo a usuarios, no inciden en el juego de la competencia entre las empresas, por lo que se sitúan al margen de la aplicación de todo el conjunto de garantías del art. 115 LCSP, pudiéndose desarrollar con mayor flexibilidad. Sin perjuicio, obviamente, de que se pueda canalizar la participación de los usuarios durante el proceso de consultas preliminares del mercado del art. 115 LCSP, en aras de una mayor economía procedimental.

2.2.3.- Evaluación preliminar de los riesgos

La decisión de iniciar un procedimiento destinado a la adquisición de una solución basada en IA deberá partir de una valoración inicial de los posibles riesgos que conlleva la utilización de esta tecnología

⁴¹ Tal y como establece el art.115 LCSP, antes de iniciarse la consulta, el órgano de contratación publicará en el perfil de contratante ubicado en la Plataforma de contratación del Sector Público o servicio de información equivalente a nivel autonómico, el objeto de la misma, cuando se iniciara esta y las denominaciones de los terceros que vayan a participar en la consulta, a efectos de que puedan tener acceso y posibilidad de realizar aportaciones todos los posibles interesados. Posteriormente, cuando el órgano de contratación haya realizado las consultas a que se refiere el presente artículo, hará constar en un informe las actuaciones realizadas. En el informe se relacionarán los estudios realizados y sus autores, las entidades consultadas, las cuestiones que se les han formulado y las respuestas a las mismas. Este informe estará motivado, formará parte del expediente de contratación, y estará sujeto a las mismas obligaciones de publicidad que los pliegos de condiciones, publicándose en todo caso en el perfil del contratante del órgano de contratación

⁴² ALHOLA, K. / SALO, A. / ANTIKAINEN, R. / BERG, A. (2017): “Promoting Public Procurement of Sustainable Innovations: Approaches for Effective Market Dialogue”, en *Global Public Procurement Theories and Practices*, Thai, Khi V. (Ed.), Springer, p. 97.

y de cómo mitigar estos riesgos⁴³, de modo que la decisión se fundamente en un enfoque proporcionado⁴⁴.

De hecho, dados los riesgos que entraña la contratación de soluciones de inteligencia artificial, se recomienda que los pliegos solo exijan la implementación de una solución de inteligencia artificial cuando existan fuertes indicios de que esta tecnología solventará el problema que se intenta resolver⁴⁵. La necesidad de adquirir un sistema de inteligencia artificial debe derivar de un análisis de distintas alternativas, de forma que se compare con otras potenciales soluciones -siempre que el proyecto de inteligencia artificial no este orientado específicamente a la investigación e innovación-. Si durante la fase de evaluación resulta evidente que otra solución que no incorpore IA satisface la necesidad pública, las Guías sobre contratación de soluciones de inteligencia artificial recomiendan emplear la opción alternativa⁴⁶. En este mismo sentido, ya había advertido ZLOTNIK⁴⁷ que antes de plantear una solución basada en IA se debe realizar un estudio de alternativas, en el que se comparen los riesgos asumidos, así como los recursos y el tiempo necesario para la construcción de una solución clásica, basada en programación tradicional, con la adopción de una solución basada en IA.

2.2.4.- ¿Existen datos de calidad?

Finalmente, en la fase de preparación de la licitación debe analizarse cuidadosamente las cuestiones relativas a la viabilidad y gobernanza de datos. La disponibilidad de los datos es un requisito previo para la implementación de cualquier solución de inteligencia

⁴³ WORLD ECONOMIC FORUM, *Guidelines for AI Procurement*, p.9.

⁴⁴ Para un ejemplo de herramienta de evaluación de los riesgos de IA puede consultarse en FORO ECONÓMICO MUNDIAL (2021): *Manual para los gobiernos sobre adquisiciones de sistemas de inteligencia artificial*, disponible en WEF_AI_Procurement_in_a_Box_Workbook_Espanol_2021.pdf

⁴⁵ WORLD ECONOMIC FORUM, *Guidelines for AI Procurement*, p.8.

⁴⁶ WORLD ECONOMIC FORUM, *Guidelines for AI Procurement*, p.8.

⁴⁷ ZLOTNIK, A. (2019): “Inteligencia Artificial en las Administraciones Públicas: definiciones, evaluación de viabilidad de proyectos y áreas de aplicación”, *Boletín*, n.o 84, pp.24.

artificial⁴⁸. Además, se ha señalado que el acceso a los datos solo debe concederse tras un examen cuidadoso de las cuestiones relacionadas con la gobernanza de datos⁴⁹, que debe abarcar todas las actividades de datos relacionadas con el proyecto: acceso de los datos por parte de los miembros del proyecto, lugar de almacenamiento y cuestiones relativas al consentimiento⁵⁰. Además, como se ha propuesto, la licitación debe fomentar enfoques tecnológicos innovadores que hagan un uso menos intrusivo de los datos⁵¹.

En definitiva, el análisis de los datos existentes y de las cuestiones ligadas a su gobernanza es uno de los primeros pasos que deben desarrollarse en la preparación de la licitación, pues si en este examen aparecen obstáculos insalvables, no se justificará que continúe siquiera el diseño de la contratación.

2.3.- La elección del procedimiento de licitación

El procedimiento de licitación es el cauce formal que permite conseguir el resultado óptimo respecto a cada concreta adquisición. De ahí la trascendencia de elegir la herramienta más idónea en cada supuesto determinado.

Aunque el procedimiento abierto es el procedimiento que, en sus distintas modalidades, se usa con mayor frecuencia en la práctica en todo tipo de contrataciones⁵², en más de una ocasión no resultará el

⁴⁸ Como señala ZLOTNIK, A. (2019): “Inteligencia Artificial...”, op.cit. p.27, la viabilidad técnica y organizativa de una determinada solución de IA requiere, al menos y entre otro requisitos, la existencia de datos representativos y suficientes.

⁴⁹ WORLD ECONOMIC FORUM, *Guidelines for AI Procurement*, p. 10.

⁵⁰ REINO UNIDO (2019): *Guidelines for AI Procurement*.

⁵¹ REINO UNIDO (2019): *Guidelines for AI Procurement*.

⁵² Según el Informe Anual de Supervisión de la Contratación Pública de la OIRESCON, en 2021, la suma de procedimiento abierto y procedimiento abierto simplificado supone en el total de las licitaciones el 78,88% en importe económico y el 71,94% en número de procedimientos.

De un total de 162.086 expedientes, durante el año 2021 se licitaron 1224 expedientes relativos a procedimientos negociados con publicidad/licitaciones con negociación, 17 diálogos competitivos y 2 relativos a asociaciones para la innovación.

El informe, fechado en octubre de 2022, está disponible en <https://www.hacienda.gob.es/RSC/OIREscon/informe-anual-supervision-2022/ias2022-modulo1.pdf>

procedimiento idóneo para la adquisición de soluciones de inteligencia artificial, en tanto en cuanto no es el más adecuado para la compra de innovación⁵³. En este sentido, debe recordarse que pueden utilizarse el procedimiento de licitación con negociación [art.167 a] LCSP] y el diálogo competitivo [art.173.3 LCSP] “cuando la prestación del objeto del contrato incluya un proyecto o soluciones innovadoras” y la asociación para la innovación⁵⁴ cuando se trate “del desarrollo de productos, servicios u obras innovadores y la compra ulterior de los suministros, servicios u obras resultantes” [art.177.1 LCSP].

Aunque existe coincidencia de los supuestos habilitantes para su utilización, el procedimiento de licitación con negociación y el diálogo competitivo son procedimientos diferentes porque responden a puntos de partida y a objetivos distintos, lo que determina sus diferencias de tramitación procedimental⁵⁵. El procedimiento de licitación con negociación es idóneo en los supuestos en los que el poder adjudicador puede definir con cierta precisión el objeto del contrato, si bien este tiene las características de diversidad de planteamientos de ejecución y de complejidad que aconsejen la negociación con los proveedores. En cambio, en el diálogo competitivo, el poder adjudicador no está siquiera en condiciones de definir la prestación, de forma que

⁵³ Sobre esta cuestión, vid. GALLEGO CÓRCOLES, I. (2014): “Los procedimientos de adjudicación en la nueva Directiva sobre contratación pública”, en *Observatorio de los contratos públicos 2013*, GIMENO FELIU, J.M. (dir.), Aranzadi, Cizur Menor.

⁵⁴ Vid. VALCÁRCEL FERNÁNDEZ, P. (2018): “La compra pública de innovación (CPI): el procedimiento de adjudicación de asociación para la innovación”, en GAMERO CASADO, E./GALLEGO CÓRCOLES, I., *Tratado de Contratos del Sector Público*, t.2, Tirant lo Blanch, Valencia, pp.1669-1712.

⁵⁵ BLANCO LÓPEZ, F. (2011): “El procedimiento negociado de adjudicación de los contratos administrativos. La negociación como elemento esencial y configurador del procedimiento”, *Revista Contratación Administrativa Práctica*, N° 111, p.34.

En relación con estos procedimientos, vid. HERNANDO RYDINGS, M. (2018): “Procedimiento de licitación con negociación y diálogo competitivo”, en GAMERO CASADO, E./GALLEGO CÓRCOLES, I., *Tratado de Contratos del Sector Público*, t.2, Tirant lo Blanch, Valencia, pp.1629-1668.

Sobre la potencialidad de la utilización del procedimiento de licitación con negociación en la compra pública de tecnología en la nube, vid. GIMENO FELIU, J.M./ GARCÍA-ÁLVAREZ GARCÍA, G./TEJEDOR BIELSA, J. (2022): *La compra pública de tecnología...*, pp. 135 y ss.

necesita informarse previamente con el mercado provisor antes de licitar. Solo finalizado el proceso de diálogo, el poder adjudicador está capacitado para fijar sus condiciones y las características esenciales de la contratación, convirtiéndose entonces la licitación en un procedimiento formalizado sin negociación⁵⁶. Y es que la finalidad del diálogo será determinar y definir los medios adecuados para satisfacer las necesidades públicas (art.175.1 LCSP). La diferencia, por tanto, entre diálogo competitivo y procedimiento de licitación con negociación es que solo en este último caso el poder adjudicador puede especificar las características requeridas con carácter previo a la licitación. Y desde una perspectiva procedimental, el procedimiento de licitación con negociación es relativamente más rígido que el diálogo competitivo. Existe una expresa prohibición de negociar los criterios de adjudicación y los requisitos mínimos en el ámbito del procedimiento de licitación con negociación (art.169.5 LCSP), limitación que no se recoge expresamente en el ámbito del diálogo competitivo. Es más, como es sabido, durante el diálogo se podrán debatir todos los aspectos del contrato con los candidatos seleccionados (art.175.1 LCSP). Además, en el ámbito del diálogo competitivo se permiten, con ciertas cautelas, negociaciones tras la adjudicación (art.176.3 LCSP). Estas están prohibidas en el ámbito del procedimiento de licitación con negociación.

Por su parte, el procedimiento de asociación para la innovación podrá emplearse cuando la necesidad de un producto, servicio u obra innovadores que no puede ser satisfecha mediante la adquisición de productos, servicios u obras ya disponibles en el mercado (art.177.1 LCSP). Así, la asociación para la innovación es un procedimiento que tiene como finalidad el desarrollo de productos, servicios u obras innovadores y la compra ulterior de los suministros, servicios u obras resultantes, siempre que correspondan a los niveles de rendimiento y a los costes máximos acordados entre los órganos de contratación y los participantes (art.177.1 LCSP). En definitiva, permite aunar en un mismo procedimiento un contrato de servicios de I+D y un posterior contrato de obra, suministros o servicios.

De esta forma, los tipos procedimentales no son técnicas totalmente intercambiables aunque la ausencia de una clasificación jerárquica

⁵⁶ *Ibidem*.

entre los distintos procedimientos redunde a favor de un incremento de la flexibilidad. Así lo demuestra la práctica en la adjudicación de proyectos de colaboración-público privada en el ámbito de la Unión Europea. Generalmente, el procedimiento abierto y el restringido se han utilizado para procedimientos simples, cuando el poder adjudicador puede especificar todos los requisitos de antemano. En la medida en que los proyectos son más complejos y la necesidad de implicar a los licitadores para definir las soluciones óptimas crece, los poderes adjudicadores tienden a usar procedimientos más flexibles: diálogo competitivo y procedimiento de licitación con negociación⁵⁷. De esta forma, en los supuestos de compra pública innovadora -como lo será en más de una ocasión la adquisición de soluciones de inteligencia artificial-, antes de elegir el concreto procedimiento, los poderes adjudicadores tendrán que tomar en consideración aspectos tales grado de conocimiento del mercado, si son necesarios trabajos de investigación y desarrollo, si las especificaciones pueden desarrollarse, si existe necesidad de adquirir la solución a escala comercial, el número de suministradores potenciales y la estructura del mercado, y el tiempo y los recursos disponibles para la licitación⁵⁸.

2.4.- Algunos aspectos a valorar en las ofertas

Como insiste la doctrina más cualificada, a la hora de establecer los criterios de adjudicación, estos deben vincularse a la lógica de la calidad-precio. Y en consecuencia, el precio deberá ser un criterio, entre otros, y no el principal, por lo que la ponderación deberá ser

⁵⁷ EUROPEAN PPP EXPERTISE CENTRE, *Procurement of PPP and the use of Competitive Dialogue in Europe. A review of public sector practices across the EU*, págs. 15 y ss. Disponible en: <http://www.eib.org/epcc/resources/epcc-procurement-and-cd-public.pdf> (fecha de consulta: 10-10-2022).

⁵⁸ En este sentido, resulta sumamente interesante el documento *Public Procurement Innovation Guidance for Public Authorities*, publicado por PROCUREMENT OF INNOVATION PLATFORM, disponible en https://www.innovation-procurement.org/fileadmin/editor-content/Guides/PPI-Platform_Guide_new-final_download.pdf (fecha de consulta: 10-10-2022).

pensada y motivada, superándose así una visión economicista de la contratación⁵⁹.

La especificidad de la contratación de soluciones basadas en IA implica que deban tenerse en cuenta determinados aspectos en la valoración de las ofertas tales como el documento de evaluación de riesgos que puede solicitarse a los potenciales licitadores así como un plan de formación.

En primer lugar, como se señala en la *Guía del Foro Económico Mundial*, entre de los elementos a valorar en la oferta puede solicitarse a los licitadores un documento evaluación de riesgos de la IA⁶⁰, ya que este ofrecerá información valiosa acerca de lo consciente que es cada licitador de los riesgos potenciales. En este sentido, si previamente se han identificado una serie de riesgos significativos en la evaluación inicial, se debería solicitar a los licitadores que expliquen cómo se tratarán de mitigar⁶¹.

En todo caso, la evaluación de los riesgos debe ser permanente⁶², lo que deberá tener reflejo en el contenido obligacional del pliego⁶³.

Por lo demás, a la hora de garantizar la supervisión humana del algoritmo⁶⁴ resultará esencial la formación que reciban los empleados públicos pues, de otro modo, esta supervisión no será real. De esta forma, al margen de que pueda incluirse como contenido obligacional un contenido mínimo, la mejoras que se ofrezcan en relación con la misma podrán ser objeto de valoración⁶⁵.

2.5.- La adecuada delimitación del equipo de trabajo: la necesaria incorporación de juristas

Como se ha señalado *supra*, el desarrollo de un proyecto completo de IA requiere un equipo multidisciplinar con conocimientos en las

⁵⁹ Entre otros trabajos, GIMENO FELIU, J.M./ GARCÍA-ÁLVAREZ GARCÍA, G./ TEJEDOR BIELSA, J. (2022): *La compra pública de tecnología...*, p. 138.

⁶⁰ WORLD ECONOMIC FORUM, *Guidelines for AI Procurement*, p.9.

⁶¹ *Ibidem*.

⁶² *Ibidem*.

⁶³ Vid. *Infra*.

⁶⁴ Vid. *Infra*.

⁶⁵ WORLD ECONOMIC FORUM, *Guidelines for AI Procurement*, p.14.

distintas disciplinas independientes que cubre la inteligencia artificial. De forma paralela, los pliegos deben asegurar que el equipo que ejecutará el proyecto dispone de las competencias y cualificaciones necesarias⁶⁶. No parece que esta sea la realidad de la práctica española, tal y como ha denunciado GAMERO CASADO⁶⁷ que, tras un cribado de los pliegos que se han venido publicando, ha puesto de manifiesto la ausencia de perfiles jurídicos en la ejecución del contrato. Así, como señala el citado autor, aunque en los equipos se suelen incluir a ingenieros, informáticos, consultores e incluso expertos en comunicación, no están presentes juristas, lo que constituye un grave problema, ya que dificultan que se tomen en consideración suficientemente los –intensos– condicionamientos jurídicos que rodean la implantación de estas soluciones.

En definitiva, otro de los aspectos clave que deberán tomarse en consideración en el diseño de la contratación es la exigencia al licitador de la conformación de un equipo en el que estén presentes todos los perfiles necesarios para llevar a cabo con éxito el proyecto.

2.6.- La necesidad de prever posibles incidencias en la ejecución

El desarrollo de una solución de inteligencia artificial en ocasiones arrojará resultados que no se adecuen a la necesidad inicial. De hecho, las dificultades que eventualmente surjan en la ejecución del contrato pueden incluso hacer aconsejable su terminación anticipada o la adaptación del contenido de la prestación. No obstante, el régimen de modificación de los contratos que establecen los arts.203 y ss. LCSP, por imposición del Derecho europeo, dificulta la introducción de modificaciones en el contrato si estas no están previstas en los pliegos de condiciones⁶⁸. De modo que la falta de previsión -en el momento de

⁶⁶ WORLD ECONOMIC FORUM, *Guidelines for AI Procurement*, p.13.

⁶⁷ Vid. GAMERO CASADO, E. (2022): “Supervisión, auditoria y control jurídico...”, *op. cit.*

⁶⁸ Sobre esta cuestión, GALLEGO CÓRCOLES, I. (2015): “La modificación de los contratos en la cuarta generación de Directivas sobre contratación pública” en GIMENO FELIU, J.M. et alt., *Las nuevas directivas de contratación pública: (ponencias sectoriales X Congreso Asociación Española Profesores de Derecho Administrativo)*, Aranzadi, Cizur Menor, pp.107-167 y HORGUÉ BAE-NA, C., “Modificación de los contratos” en GAMERO CASADO, E./GALLEGO

licitación- de las incidencias a las que puede abocar el desarrollo de un contrato complejo desencadenarán en la práctica situaciones de difícil solución. De todo ello es suficientemente ilustrativa la STJUE de 7 de septiembre de 2016, C-549/14, *Frogne*.

En el supuesto que resuelve la sentencia *Frogne*, en la ejecución de un contrato para el suministro de un sistema de comunicaciones complejo, y ante evidentes dificultades de ejecución, las partes deciden poner fin de forma anticipada al contrato. Una de las empresas que no participó en la licitación recurre dicho acuerdo, entendiendo que la reducción del objeto contractual suponía una modificación esencial del contrato contraria a las Directivas sobre contratación. Frente a ello, la Administración argumentaba que cuando surgen dificultades en la ejecución de un contrato —lo que no es inhabitual en ciertos tipos de contratos, como los referentes al desarrollo de sistemas informáticos—, es preciso reconocer un amplio margen de apreciación al poder adjudicador, a fin de permitirle llegar a una solución razonable para las dificultades de ejecución. En caso contrario, se obliga al poder adjudicador, bien a renunciar a adaptaciones razonables, bien a romper el contrato, con los riesgos y pérdidas que esto conlleva.

Frente a estos argumentos, el TJUE concluyó que con posterioridad a la adjudicación de un contrato público no es posible introducir en él una modificación sustancial sin abrir un nuevo procedimiento de adjudicación del contrato, ni siquiera en el supuesto de que esa modificación constituya, objetivamente, una solución de compromiso que implique renunciaciones recíprocas de ambas partes y pretenda poner fin a un conflicto de resultado incierto, nacido de las dificultades que la ejecución del contrato plantea. Solo cabría una conclusión diferente en el caso de que la documentación de dicho contrato estableciera la facultad de adaptar determinadas condiciones del mismo, incluso importantes, con posterioridad a su adjudicación y determinara el modo de aplicar esa facultad. En este sentido, el Tribunal tuvo ocasión de afirmar que:

“el propio hecho de que ciertos contratos públicos puedan considerarse desde el primer momento, a causa de su objeto, contratos que

presentan un carácter aleatorio permite anticipar el riesgo de que surjan dificultades en la fase de ejecución de los mismos. Por lo tanto, para un contrato de este tipo, incumbe al poder adjudicador, no sólo recurrir al procedimiento de adjudicación más adaptado, sino además determinar con precaución el objeto del contrato. [...]

En efecto, aunque el principio de igualdad de trato y la obligación de transparencia deben respetarse incluso en relación con los contratos públicos especiales, ello no impide tomar en consideración los rasgos específicos de tales contratos. La conciliación de este imperativo jurídico y de esta necesidad concreta requiere, por una parte, un respeto estricto de las condiciones del contrato establecidas en la documentación del mismo hasta que finalice la fase de ejecución de ese contrato, pero también, por otra parte, la posibilidad de establecer expresamente en esa documentación la facultad del poder adjudicador de adaptar determinadas condiciones de dicho contrato, incluso importantes, con posterioridad a su adjudicación. Al establecer expresamente esa facultad y determinar el modo de aplicarla en dicha documentación, el poder adjudicador garantiza que todos los operadores económicos interesados en participar en ese contrato tengan conocimiento de ello desde el principio y se encuentren así en pie de igualdad en el momento de formular su oferta (véase, por analogía, la sentencia de 29 de abril de 2004, Comisión/CAS Succhi di Frutta, C-496/99 P, EU:C:2004:236, apartados 112, 115, 117 y 118)⁶⁹.

A la vista de la doctrina contenida en la sentencia, resultará aconsejable que cuando se liciten herramientas basadas en IA de cierta complejidad se articule la ejecución del contrato por fases sucesivas y se prevea, si es necesario, posibles adaptaciones del objeto contractual.

En otro orden de consideraciones, pueden existir igualmente dificultades cuando se produce la extinción de contratos que incluyan software externo crítico para el funcionamiento de la Administración, y en el que no se haya cedido la propiedad intelectual del mismo. En un contexto como el descrito, la STSJ Asturias de 17 de noviembre de 2020 (ECLI:ES:TSJAS:2020:2595) ha admitido la validez de cláusulas de garantía de mantenimiento del uso de la aplicación hasta que ésta sea reemplazada por la Administración⁷⁰.

⁶⁹ Considerandos 36 y 37.

⁷⁰ Sobre esta sentencia, vid. MIRANZO DÍAZ, J. (2021): “Apuntes sobre la contratación de sistemas de software”, *www.obcp.es*. En el supuesto, el uso del software se mantuvo hasta veintidós meses después de la finalización del contrato. La sentencia destaca que el uso de la aplicación de los programas, caso de fina-

3.- ASPECTOS RELEVANTES DEL CONTENIDO DEL CONTRATO

3.1.- *La necesidad de cumplir las exigencias de transparencia, motivación y libre competencia: ¿debe exigirse la cesión de la titularidad del código fuente?*

3.1.1.- Planteamiento

Como insiste la doctrina más especializada⁷¹, las obligaciones de transparencia y motivación a las que están sometidas nuestras Administraciones públicas no cesan en los supuestos en los que se realicen actuaciones basadas en IA. En particular, los arts. 5.1LTBG y 35LPAC no quedan desplazados. De hecho, otras normas autonómicas o sectoriales refuerzan las garantías que estos preceptos contienen⁷². Por lo demás, el art. 11.1 i) de Real Decreto 203/2021, de 30 de marzo, por el que se aprueba el Reglamento de actuación y funcionamiento del sector público por medios electrónicos, ha ampliado la información que las Administraciones Públicas deben difundir a través de las sedes electrónicas, incluyendo ahora la relación actualizada de las actuaciones administrativas automatizadas vinculadas a los servicios, procedimientos y trámites. Para ello, se informará de la descripción del diseño y funcionamiento de la actuación administrativa automatizada, de

lización del contrato, estaba expresamente contemplado en la cláusula 5.4 PPT que no establecía tiempo máximo ni pago alguno por la utilización del software.

⁷¹ Entre otros, PONCE SOLÉ, J. (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad; algoritmos y procedimiento administrativo debido tecnológico”, *Revista General de Derecho Administrativo*, n° 50, p. 40. En esta obra, vid. GAMERO CASADO, E., “Las garantías de la legislación administrativa y el derecho a la tutela judicial efectiva: actividad automatizada, motivación de los actos y control judicial”.

⁷² En este sentido, el art. 16 de Ley 1/2022, de 13 de abril, de la Generalitat, de la Transparencia y Buen Gobierno de la Comunitat Valenciana, contempla entre las obligaciones de publicidad activa relativas a información de relevancia jurídica “la relación de sistemas algorítmicos o de inteligencia artificial que tengan impacto en los procedimientos administrativos o la prestación de los servicios públicos con la descripción de manera comprensible de su diseño y funcionamiento, el nivel de riesgo que implican y el punto de contacto al que poder dirigirse en cada caso, de acuerdo con los principios de transparencia y explicabilidad”.

los mecanismos de rendición de cuentas y transparencia, así como de los datos utilizados en su configuración y aprendizaje⁷³.

Las garantías de transparencia y motivación hallan cierta concreción en la Carta de Derechos Digitales aprobada por el Gobierno de España a través de la Secretaría de Estado de Digitalización e Inteligencia Artificial con la colaboración de un grupo de expertos⁷⁴. Así, los apartados 6 b) y c) del art. XVIII prevén que la ley pueda regular las condiciones de transparencia y el acceso al código fuente y reconocen el derecho de los ciudadanos a obtener una motivación comprensible⁷⁵. Estas obligaciones, conviene insistir, no nacen de este texto no vinculante⁷⁶, sino que ya estaban presentes en nuestro Ordenamiento -como en otros de nuestro entorno-.

Por lo demás, la necesidad de que en el diseño de la contratación se prevea la forma en la que las obligaciones de transparencia y rendición de cuentas se cumplirán se reitera en las diferentes *Guías* que

⁷³ Sobre el interés de esta norma y sobre las dificultades prácticas de su aplicación, vid. CERRILLO I MARTÍNEZ, A. (2019): “Robots, asistentes virtuales..”, *op. cit.*, p.287.

⁷⁴ Disponible en https://www.lamoncloa.gob.es/presidente/actividades/Documentos/2021/140721-Carta_Derechos_Digitales_RedEs.pdf (Última visita: 30/05/2022).

⁷⁵ En concreto se dispone que: “se promoverán los derechos de la ciudadanía en relación con la inteligencia artificial reconocidos en esta Carta en el marco de la actuación administrativa reconociéndose en todo caso los derechos a [...]: B) Transparencia sobre el uso de instrumentos de inteligencia artificial y sobre su funcionamiento y alcance en cada procedimiento concreto y, en particular, acerca de los datos utilizados, su margen de error, su ámbito de aplicación y su carácter decisorio o no decisorio.

La ley podrá regular las condiciones de transparencia y el acceso al código fuente, especialmente con objeto de verificar que no produce resultados discriminatorios.

c) Obtener una motivación comprensible en lenguaje natural de las decisiones que se adopten en el entorno digital, con justificación de las normas jurídicas relevantes, tecnología empleada, así como de los criterios de aplicación de las mismas al caso. El interesado tendrá derecho a que se motive o se explique la decisión administrativa cuando esta se separe del criterio propuesto por un sistema automatizado o inteligente”.

⁷⁶ Sobre la ausencia del carácter vinculante de la Carta y sus implicaciones vid. MARTÍN DELGADO, I. (2022), “La aplicación del principio...”, *op. cit.*

se han elaborado por distintas instituciones respecto a la contratación pública de soluciones de IA⁷⁷. También en la doctrina⁷⁸.

Siendo ello así, el debate que se plantea con cierta intensidad es el relativo a determinar si las obligaciones de transparencia y motivación obligan a establecer mecanismos que permitan dar acceso a terceros al código fuente⁷⁹. La solución que se adopte al respecto habrá de plasmarse tanto en las cláusulas relativas a la eventual cesión de derechos de propiedad intelectual, como las referentes a los límites del principio de confidencialidad⁸⁰.

⁷⁷ Vid. WORLD ECONOMIC FORUM, *Guidelines for AI Procurement*, pp. 13-15.

⁷⁸ Entre otros, McBRIDE, K./ VAN NOORDT, C./ MISURACA, G./ HAMMERS-CHMID, G. (2021): “Towards a Systematic Understanding...”, *op. cit.*, pp.11.

⁷⁹ MARTÍN DELGADO, I. (2022), “La aplicación del principio...”, *op. cit.*.

⁸⁰ En relación estos, el art. 133.1 LCSP dispone que “sin perjuicio de lo dispuesto en la legislación vigente en materia de acceso a la información pública y de las disposiciones contenidas en la presente Ley relativas a la publicidad de la adjudicación y a la información que debe darse a los candidatos y a los licitadores, los órganos de contratación no podrán divulgar la información facilitada por los empresarios que estos hayan designado como confidencial en el momento de presentar su oferta. El carácter de confidencial afecta, entre otros, a los secretos técnicos o comerciales, a los aspectos confidenciales de las ofertas y a cualesquiera otras informaciones cuyo contenido pueda ser utilizado para falsear la competencia, ya sea en ese procedimiento de licitación o en otros posteriores. El deber de confidencialidad del órgano de contratación así como de sus servicios dependientes no podrá extenderse a todo el contenido de la oferta del adjudicatario ni a todo el contenido de los informes y documentación que, en su caso, genere directa o indirectamente el órgano de contratación en el curso del procedimiento de licitación. Únicamente podrá extenderse a documentos que tengan una difusión restringida, y en ningún caso a documentos que sean públicamente accesibles.

El deber de confidencialidad tampoco podrá impedir la divulgación pública de partes no confidenciales de los contratos celebrados, tales como, en su caso, la liquidación, los plazos finales de ejecución de la obra, las empresas con las que se ha contratado y subcontratado, y, en todo caso, las partes esenciales de la oferta y las modificaciones posteriores del contrato, respetando en todo caso lo dispuesto en la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal.”

Sobre la problemática aplicación de este principio, vid. RAZQUIN LIZARRGA M.M. (2018): “El principio de confidencialidad en la contratación pública”, en GIMENO FELIU, J.M. (dir.): *Estudio sistemático de la ley de contratos del sector público*, Aranzadi, pp.1047-1071.

Además, el acceso al código fuente puede también implicaciones desde la perspectiva del cumplimiento de la normativa sobre contratación pública: la ausencia de cesión de los derechos de propiedad intelectual puede suponer que las ampliaciones y desarrollos no incluidos en el objeto inicial del contrato deban ser realizados por el mismo contratista, lo que, en algunas ocasiones como veremos, puede suponer una vulneración del principio de libre concurrencia.

3.1.2.- El eventual acceso al código fuente en las Guías sobre contratación de soluciones de inteligencia artificial

Desde una perspectiva general, y aunque parte de nuestra doctrina sostiene que existe obligación de otorgar acceso al código fuente⁸¹, como señala MARTIN DELGADO no necesariamente ha de equipararse transparencia y motivación con apertura total del sistema algorítmico. Transparencia y motivación se vinculan a la legitimidad del sistema, la rendición de cuentas y el control de la decisión adoptada; y para cumplir tales fines no siempre resultará preciso dar a conocer el interior del sistema, el código fuente. Lo relevante en términos jurídicos -continúa el autor- es difundir la existencia del sistema, los datos de los que se nutre, la información relativa a su operatividad técnica, las razones por las que se ha adoptado la decisión. De esta forma, las exigencias de transparencia y motivación no estarían vinculadas tanto a dar a conocer sistemáticamente la forma en la que opera el sistema desde el punto de vista tecnológico, sino a la interacción humana con el *output* del algoritmo y los criterios

⁸¹ Entre otros, BOIX PALOP A. (2020): “Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la Administración para la adopción de decisiones”, *Revista de Derecho Público: Teoría y Método*, Vol. I, pp. 223-270 y GÓMEZ JIMÉNEZ, M.L. (2021): *Automatización procedimental y sesgo electrónico: el procedimiento administrativo electrónico desde la Inteligencia Artificial*, Thomson-Reuters Aranzadi, Cizur Menor, p. 150.

de diseño de los *inputs* para mantener la exigencia de inteligibilidad de la decisión⁸².

Los distintos documentos que en otros Ordenamientos se han desarrollado sobre contratación de soluciones de inteligencia artificial contienen la propuesta que se ha reseñado: no es preciso en todo caso dar acceso al código fuente para salvaguardar las garantías de transparencia y motivación. Así, en la *Guía* del Reino Unido, en relación con eventual acceso al código fuente se señala que es posible que los proveedores no deseen revelar detalles del funcionamiento interno de su solución para proteger su propiedad intelectual y la ventaja comercial. Además, se subraya que durante el diseño y el despliegue del sistema de IA es probable que se desarrolle un nuevo algoritmo, o se adapte uno ya existente (por ejemplo, reentrenado a través de sus datos), por lo que la Administración de valorar si el proveedor o ella misma debe ser la propietaria de los derechos de propiedad intelectual, para lo que resultará fundamental determinar quién puede explotar mejor la propiedad intelectual generada.

Por su parte, las *cláusulas contractuales de la Ciudad de Ámsterdam* establecen una regulación muy detallada de las cuestiones relacionadas con la explicabilidad y transparencia de los algoritmos. Para ello, se introduce una diferenciación entre transparencia procedimental, transparencia técnica y explicabilidad⁸³. El concepto de “transparencia procedimental” hace referencia a la información sobre la finalidad del sistema algorítmico y el proceso seguido en su desarrollo los datos utilizados en ese contexto, incluyendo en todo caso

⁸² *Ibidem*.

En un sentido en parte coincidente, señala PONCE SOLÉ, J. (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico”, *Revista General de Derecho Administrativo*, n.º 50, que en los casos en que la transparencia plena (esto es, el acceso completo al código fuente de un algoritmo) no es posible o deseable, se ha planteado la utilidad de las *auditorías de algoritmos* para poder comprobar aspectos como la licitud de las premisas o principios básicos que ejecuta el algoritmo (si son o no discriminatorias), la regularidad del funcionamiento del sistema, la regularidad del funcionamiento y aplicación del algoritmo. Para el autor, este es un campo a desarrollar en el futuro.

⁸³ Sobre el concepto de explicabilidad, vid. igualmente MARTÍN DELGADO, I. (2022), “La aplicación del principio ...”, *op. cit.*

una descripción de las opciones e hipótesis realizadas, las categorías de datos utilizadas, la forma en que se prevé la intervención humana en el método utilizado para identificar los riesgos, los riesgos identificados y las medidas adoptadas para mitigar los riesgos, así como las partes que participaron en el desarrollo del sistema algorítmico y sus funciones.

Por su parte, la “transparencia técnica” es aquella que permite comprender el funcionamiento técnico del sistema algorítmico, lo que incluye en todo caso la divulgación del código fuente del sistema algorítmico, las especificaciones técnicas y los datos utilizados en su desarrollo, la información técnica sobre cómo se han obtenido y editado los datos utilizados en el desarrollo del sistema algorítmico, la información sobre el método de desarrollo utilizado y el proceso de desarrollo emprendido, la justificación de la elección de un modelo concreto y sus parámetros, y la información sobre el rendimiento del sistema algorítmico.

Por su parte, las cláusulas definen “explicabilidad” como la capacidad de explicar a nivel individual por qué un sistema algorítmico conduce a una decisión o resultado concreto, lo que incluirá en cualquier caso una indicación clara de los factores clave que han llevado a un sistema algorítmico a un resultado particular y los cambios en la entrada que deben hacerse para llegar a un resultado diferente. En este sentido, se añade que hacer que un sistema algorítmico sea explicable incluye el suministro de toda la información técnica y de otro tipo necesaria para explicar, en los procedimientos de recurso o en otros procedimientos legales, cómo se ha llegado a una decisión y para ofrecer a la otra parte y a cualquier otra parte interesada la oportunidad de evaluar la forma en que se ha llegado a una decisión, a fin de otorgar a terceros una protección legal realista.

Dentro de este marco conceptual, la regulación que las *cláusulas de la Ciudad de Amsterdam* establecen en su artículo 5 sobre la transparencia de los algoritmos imponen al contratista distintas obligaciones dependiendo si la información está incluida en los con-

ceptos de “transparencia procedimental”, “transparencia técnica” o “explicabilidad”⁸⁴.

⁸⁴ El tenor de la cláusula es el siguiente (traducción de la autora a partir del inglés):
ARTÍCULO 5–TRANSPARENCIA DEL SISTEMA ALGORÍTMICO

5.1 A petición de [NOMBRE DE LA ORGANIZACIÓN], el contratista proporcionará a NOMBRE DE LA ORGANIZACIÓN transparencia en el procedimiento. NOMBRE DE LA ORGANIZACIÓN tendrá derecho a compartir la información proporcionada por el contratista en ese contexto con terceros y a divulgarla. A petición de NOMBRE DE LA ORGANIZACIÓN, el contratista la inscribirá en un registro de sistemas algorítmicos que será designado por NOMBRE DE LA ORGANIZACIÓN.

5.2 A petición de NOMBRE DE LA ORGANIZACIÓN, el contratista facilitará a NOMBRE DE LA ORGANIZACIÓN la Transparencia Técnica para que NOMBRE DE LA ORGANIZACIÓN pueda llevar a cabo la auditoría a la que se refiere el artículo 8. NOMBRE DE LA ORGANIZACIÓN solo solicitará y utilizará dicha información si y en la medida necesaria para los fines del artículo 8. NOMBRE DE LA ORGANIZACIÓN, de conformidad con el artículo 5.2, mantendrá la confidencialidad de la información comercial confidencial que se le proporcione y la destruirá después de una auditoría como la mencionada en el artículo 8, a menos que una obligación legal por parte de NOMBRE DE LA ORGANIZACIÓN se oponga a la confidencialidad o a la destrucción o que NOMBRE DE LA ORGANIZACIÓN necesite la información en el contexto de un litigio con el contratista o un tercero.

5.3 A los efectos del artículo 5.2, el contratista podrá optar por no entregar el código fuente del sistema algorítmico a NOMBRE DE LA ORGANIZACIÓN, sino a un tercero independiente que será designado y contratado por NOMBRE DE LA ORGANIZACIÓN y que realizará la auditoría mencionada en el artículo 8 en nombre de NOMBRE DE LA ORGANIZACIÓN. Cualquier coste adicional en el que se incurra como consecuencia de ello será a cargo del contratista. NOMBRE DE LA ORGANIZACIÓN podrá exigir al contratista el pago de un anticipo en relación con los costes del tercero independiente.

5.4 NOMBRE DE LA ORGANIZACIÓN deberá ser capaz de explicar en todo momento el funcionamiento del sistema algorítmico (Explicabilidad). El contratista estará obligado a prestar toda su colaboración para que el sistema algorítmico sea explicable y a facilitar a NOMBRE DE LA ORGANIZACIÓN toda la información que le sea requerida al respecto. NOMBRE DE LA ORGANIZACIÓN tendrá derecho a compartir la información proporcionada por el contratista en ese contexto con terceros y a divulgarla.

5.5 Durante la vigencia del contrato, las exigencias descritas en este artículo constituirán obligaciones para realizar dichas exigencias y, salvo que las partes acuerden lo contrario, NOMBRE DE LA ORGANIZACIÓN no adeudará honorarios adicionales al contratista en consideración al cumplimiento de dichas obligaciones. Tras la expiración del plazo del contrato, las demandas descritas en este artículo constituirán una obligación de realizarlas en la medida de las posibi-

En primer lugar, a través de la introducción de concepto de “transparencia procedimental” se establece una categoría de transparencia que permite divulgar información relevante sobre el sistema algorítmico sin perjudicar desproporcionadamente los intereses del contratista⁸⁵. Así, el contratista está obligado a documentar la información incluida en el ámbito de la definición de “transparencia procedimental” y a compartir dicha información con la Administración contratante, que tendrá la posibilidad revelar la información proporcionada y compartirla con terceros. La regulación se basa en que esta información es esencial para que administración pueda rendir cuentas a la ciudadanía en relación con el uso del sistema algorítmico.

Por otro lado, se dispone que el contratista proporcionará “transparencia técnica” solo a petición de Administración. La finalidad de esta modalidad de transparencia es permitir a la Administración comprobar si el contratista ha cumplido sus obligaciones contractuales a través de las correspondientes auditorías. Pues bien, en muchas ocasiones, la información proporcionada dentro del concepto de transparencia técnica estará protegida por derechos de propiedad intelectual. Por ello las cláusulas obligan a la Administración a preservar el carácter confidencial de esta información y a destruirla una vez finalizada la auditoría. En todo caso, no existe obligación de mantener la confidencialidad o de destruir la información si un deber legal impuesto a la Administración se opone a ello, o si la Administración requiere la información en el contexto de un litigio con un contratista o un tercero. Además, con la finalidad de reforzar los intereses del contratista, se prevé que este pueda optar por no proporcionar el código fuente del sistema algorítmico, sino entregarlo a un tercero independiente que realizará la auditoría en nombre de la Administración y, en ese contexto, analizará el código fuente del sistema algorítmico.

lidades de las Partes y NOMBRE DE LA ORGANIZACIÓN deberá abonar un honorario adicional al contratista como contraprestación por los servicios que éste deba prestar al respecto.

5.6 En todas las situaciones en las que la información facilitada en base a los distintos apartados de este artículo se superponga, NOMBRE DE LA ORGANIZACIÓN será libre de elegir el régimen que le resulte más favorable.

⁸⁵ *Standard Clauses for Procurement of Trustworthy Algorithmic Systems, City of Amsterdam.*

Finalmente se establece que la Administración debe poder explicar en todo momento el funcionamiento del sistema algorítmico, y que el contratista tiene la obligación de prestar su plena colaboración para que el sistema algorítmico sea explicable y de facilitar a la Administración toda la información que sea necesaria para ello. La Administración tendrá derecho a compartir la información proporcionada por el contratista en ese contexto con terceros y a divulgarla⁸⁶.

Como cláusula de cierre del sistema se prevé una especie de prelación entre los distintos tipos de transparencia. Así, como se señala en el documento, es probable que, en muchas situaciones, la información proporcionada en el contexto de la transparencia procedimental, transparencia técnica, y explicabilidad coincidan parcialmente. Por ello en las *cláusulas* se establece que, en ese caso, la Administración podrá elegir en todo momento el régimen que le resulte más favorable. Con ello se pretende evitar que el contratista pueda ampararse los límites a la transparencia técnica para impedir que la Administración haga explicable un sistema algorítmico. Por lo tanto, si cierta información entra en el ámbito de la transparencia técnica y es necesaria para hacer un sistema algorítmico explicable, la Administración será libre de revelar la información relevante y compartirla con terceros.

En definitiva, la regulación que se incluye en las cláusulas de la Ciudad de Ámsterdam se articula de forma que, sin llegar a imponer en todo caso la entrega del código fuente, contienen elevadas garantías de transparencia y motivación. En la práctica, quizá la principal debilidad es que no termina de perfilar con precisión las obligaciones que asume el contratista, que dependerá en cada caso concreto de la posibilidad de garantizar la explicabilidad de las decisiones sin divulgar sus secretos comerciales.

Frente a este modelo, en el proyecto de cláusulas que auspicia Comisión Europea las garantías respecto a la transparencia y motivación se establecen con un carácter marcadamente abierto. Como ya se ha indicado *supra*, este proyecto se alinea con la regulación sustantiva

⁸⁶ Como se explica en las cláusulas-tipo, esta exigencia se basan en distintas decisiones del Consejo de Estado neerlandés, que exige cuando una Administración adopta una decisión usando un sistema algorítmico, que esta sea capaz de explicar cómo este ha llegado a un resultado determinado.

del RIA⁸⁷. Y este, en su artículo 70, en principio obliga a las autoridades nacionales a respetar la confidencialidad de la información y los

⁸⁷ Así, el artículo 13 RIA se refiere “transparencia y comunicación de información a los usuarios” en los siguientes términos:

“Los sistemas de IA de alto riesgo se diseñarán y desarrollarán de un modo que garantice que funcionan con un nivel de transparencia suficiente para que los usuarios interpreten y usen correctamente su información de salida. Se garantizará un tipo y un nivel de transparencia adecuados para que el usuario y el proveedor cumplan las obligaciones oportunas previstas en el capítulo 3 del presente título.

2. Los sistemas de IA de alto riesgo irán acompañados de las instrucciones de uso correspondientes en un formato digital o de otro tipo adecuado, las cuales incluirán información concisa, completa, correcta y clara que sea pertinente, accesible y comprensible para los usuarios.

3. La información a que se refiere el apartado 2 especificará:

a) la identidad y los datos de contacto del proveedor y, en su caso, de su representante autorizado;

b) las características, capacidades y limitaciones del funcionamiento del sistema de IA de alto riesgo, y en particular:

i) su finalidad prevista;

ii) el nivel de precisión, solidez y ciberseguridad mencionado en el artículo 15 con respecto al cual se haya probado y validado el sistema de IA de alto riesgo y que puede esperarse de este, así como las circunstancias conocidas o previsibles que podrían afectar al nivel de precisión, solidez y ciberseguridad esperado;

iii) cualquier circunstancia conocida o previsible, asociada a la utilización del sistema de IA de alto riesgo conforme a su finalidad prevista o a un uso indebido razonablemente previsible, que pueda dar lugar a riesgos para la salud y la seguridad o los derechos fundamentales;

iv) su funcionamiento en relación con las personas o los grupos de personas en relación con los que se pretenda utilizar el sistema;

v) cuando proceda, especificaciones relativas a los datos de entrada, o cualquier otra información pertinente en relación con los conjuntos de datos de entrenamiento, validación y prueba usados, teniendo en cuenta la finalidad prevista del sistema de IA;

c) los cambios en el sistema de IA de alto riesgo y su funcionamiento predeterminados por el proveedor en el momento de efectuar la evaluación de la conformidad inicial, en su caso;

d) las medidas de vigilancia humana a que se hace referencia en el artículo 14, incluidas las medidas técnicas establecidas para facilitar la interpretación de la información de salida de los sistemas de IA por parte de los usuarios;

e) la vida útil prevista del sistema de IA de alto riesgo, así como las medidas de mantenimiento y cuidado necesarias para garantizar el correcto funcionamiento de dicho sistema, también en lo que respecta a la actualización del software.

datos obtenidos en el ejercicio de sus funciones⁸⁸, todo ello -es preciso entender- sin perjuicio de las estipulaciones que se establezcan en el ámbito de cada concreto procedimiento de contratación. Sea como fuere, el art. 6 del borrador de cláusulas tipo se limita a disponer que⁸⁹:

“6.1. El Proveedor garantiza que el Sistema de IA se diseñará y desarrollará de manera que se garantice que el funcionamiento del Sistema de IA sea lo suficientemente transparente para que [NOMBRE DEL ORGANISMO PÚBLICO] pueda interpretar el interpretar los resultados del sistema y utilizarlos adecuadamente.

6.2. <opcional>Durante la vigencia del contrato, el Proveedor está obligado a asistir a [NOMBRE DEL ORGANISMO PÚBLICO] a la primera solicitud de [NOMBRE DEL ORGANISMO PÚBLICO] para explicar a nivel individual cómo el Sistema de IA llegó a una decisión o un resultado concreto. En cualquier caso, incluirá una indicación clara de los factores clave que han llevado al sistema de IA a un resultado concreto y los cambios en los datos de entrada que deben realizarse para llegar a un resultado diferente.

6.3. <opcional>La obligación descrita en el artículo 6.2 comprende en todo caso el suministro a [NOMBRE DEL ORGANISMO PÚBLICO] de toda la información técnica y de otro tipo necesaria para explicar cómo el sistema de IA ha llegado a una decisión o un resultado concreto y ofrecer a la otra parte y a cualquier otra partes interesada la oportunidad de verificar la forma en que el sistema de IA ha llegado a una decisión o resultado concreto. El Proveedor concede a [NOMBRE DEL ORGANISMO PÚBLICO] el derecho a utilizar, compartir y divulgar esta información, siempre

⁸⁸ En concreto, el art.70.1 RIA establece que las autoridades nacionales competentes y los organismos notificados involucrados en la aplicación del presente Reglamento respetarán la confidencialidad de la información y los datos obtenidos en el ejercicio de sus funciones y actividades de modo que se protejan, en particular: a) los derechos de propiedad intelectual y la información empresarial confidencial o los secretos comerciales de una persona física o jurídica, incluido el código fuente, salvo en los casos contemplados en el artículo 5 de la Directiva 2016/943 relativa a la protección de los conocimientos técnicos y la información empresarial no divulgados (secretos comerciales) contra su obtención, utilización y revelación ilícitas; b) la aplicación eficaz del presente Reglamento, en particular a efectos de inspecciones, investigaciones o auditorías; c) los intereses públicos y de seguridad nacional;

⁸⁹ c) la integridad de las causas penales o los procedimientos administrativos.

Traducción de la autora.

y cuando sea necesario para informar a las personas físicas sobre el funcionamiento del Sistema de IA y/o en cualquier procedimiento judicial.

6.4. <opcional> Sin perjuicio de lo dispuesto en el artículo 4, las obligaciones mencionadas en el artículo 6.2 y el artículo 6.3 incluyen el código fuente del Sistema de IA, las especificaciones técnicas utilizadas en el desarrollo del sistema de IA, los conjuntos de datos, la información técnica sobre cómo se han obtenido y editado los conjuntos de datos información sobre el método de desarrollo utilizado y el proceso de desarrollo llevado a cabo proceso de desarrollo emprendido, la justificación de la elección de un determinado modelo y sus parámetros, e información sobre el rendimiento del sistema de IA”.

Como se puede comprobar, la solución adoptada por el borrador es menos satisfactoria que la que incluía -con sus límites- las cláusulas de la Ciudad de Ámsterdam. De hecho, el contenido obligatorio -apartado 6.1- no parece suficiente para garantizar las exigencias de transparencia y motivación que dispone nuestro Derecho. Como mínimo, deberá preverse que el licitador facilite a la Administración la información necesaria para hacer efectivas las obligaciones de transparencia activa que dispone el ya citado artículo 11.1 i) de Real Decreto 203/2021, de 30 de marzo, por el que se aprueba el Reglamento de actuación y funcionamiento del sector público por medios electrónicos. Sin perjuicio de que además se incorpore la obligación de facilitar toda la información necesaria para garantizar la explicabilidad del sistema algorítmico.

3.1.3.- El principio de competencia y entrega del código fuente

Más allá de su trascendencia en las relaciones entre Administración y ciudadanía, la eventual opacidad de los algoritmos plantea dificultades desde la perspectiva del respeto al principio de libre competencia, principio ínsito en todo sistema de contratación pública⁹⁰. En este sentido, la *Guía sobre contratación de inteligencia artificial* del Reino Unido subraya la necesidad de un “enfoque explicable” para

⁹⁰ Recuérdese que el artículo 18.1 de la Directiva 2014/24/UE, relativo a los principios de la contratación, dispone que los poderes adjudicadores tratarán a los operadores económicos en pie de igualdad y sin discriminaciones, y actuarán de manera transparente y proporcionada.

el desarrollo de la IA, de forma que sus resultados puedan ser interpretados tanto por el comprador público como otros proveedores, lo que incrementará la posibilidad de que la continuación o ampliación del sistema inicial pueda atribuirse a otros empresarios, limitándose así el riesgo de cautividad del proveedor⁹¹. Hasta tal punto es así que SANCHEZ-GRAELLS⁹² entiende que la cesión del código fuente a la Administración constituye un requisito básico en la contratación de IA si el poder adjudicador quiere poder cumplir legalmente sus obligaciones en virtud de la legislación de la UE sobre contratación pública, dada la incertidumbre sobre el futuro desarrollo, integración o sustitución de las soluciones de IA en el momento de su contratación. Para llegar a esta conclusión, el autor parte de la STJUE de 28 de mayo de 2020, C-796/18, *ISE*⁹³, cuyo objeto es determinar si una concreta cesión de software puede encuadrarse como cooperación interadministrativa -y, por tanto, realizarse al margen de un procedimiento de contratación-.

Los hechos que originan la sentencia *ISE* son los siguientes: el Land de Berlín, que dispone del mayor cuerpo de bomberos en Alemania, cedió gratuitamente su software a la ciudad de Colonia. Este software fue adquirido a la empresa Sopra Steria y, en los términos del acuerdo de cesión, se estipula que la ciudad de Colonia está obligada ceder gratuitamente los módulos especializados o adicionales a dicho software. Pues bien, ISE, empresa que desarrolla y comercializa software, solicita que el acuerdo de cesión sea invalidado por vulnerar la normativa sobre contratación pública. A partir de estos hechos, el Tribunal de Justicia deba analizar, en primer lugar, si se da la nota

⁹¹ En la guía se señalan además buenas prácticas como la adopción de estándares abiertos, licencias libres y condiciones de dominio público.

⁹² SÁNCHEZ GRAELLS, A. (2022): “Public procurement and [AI] source code transparency, a (downstream) competition issue (RE C-796/18), en www.howto-crackanut.com

⁹³ Hemos tenido ocasión de estudiar con detalle su contenido en GALLEGO CÓRCOLES, I. (2021): “Transferencia de competencias y fórmulas de cooperación horizontal en la jurisprudencia del Tribunal de Justicia del año 2020”, en GIMENO FELIU, J.M.: *Observatorio de los contratos públicos 2020*, Aranzadi, pp. 449-477.

de onerosidad que es caracterizadora de los contratos públicos⁹⁴. En segundo lugar, si el contenido del acuerdo se incardina en el concepto de cooperación. Una vez otorgada una respuesta afirmativa a ambas cuestiones, el Tribunal examina si el acuerdo puede provocar un efecto de falseamiento de la competencia, en cuyo caso la cesión del software no sería lícita. Es en este contexto en el que se afirma que:

“para garantizar el respeto de los principios de la contratación, tal como se enuncian en el artículo 18 de la Directiva 2014/24, incumbe al tribunal remitente comprobar, en primer lugar, que tanto el Land de Berlín como la ciudad de Colonia disponen del código fuente del *software* «IGNIS Plus»; en segundo lugar, que, en el supuesto de que organicen una licitación para garantizar el mantenimiento, la adaptación o el desarrollo de dicho *software*, esos poderes adjudicadores comunican tal código fuente a los candidatos y licitadores potenciales y, en tercer lugar, que el acceso a ese único código fuente basta para garantizar que los operadores económicos interesados en la adjudicación del contrato en cuestión son tratados de modo transparente, equitativo y no discriminatorio”.

En definitiva, en realidad lo que censura el Tribunal no es que el Land de Berlín no hubiese adquirido la propiedad intelectual sobre el software, sino que al cederlo a otra Administración, amplía *de facto* el ámbito subjetivo del contrato, pues la Administración cesionaria acudirá también a la misma empresa para realizar desarrollos del software, desarrollos estos -es preciso señalar- de elevado valor económico⁹⁵. Y aunque la doctrina contenida en la sentencia se limita a

⁹⁴ En este sentido resulta de gran relevancia el hecho de que el contrato de cesión del *software* y el contrato de cooperación tienen carácter sinalagmático en la medida en que la cesión gratuita del *software* “IGNIS Plus” genera una obligación recíproca de desarrollo de este, desde el momento en que la óptima utilización del sistema de gestión de intervenciones y su permanente adaptación a las necesidades exigen tal evolución, que se concreta en la financiación de módulos complementarios que seguidamente han de cederse gratuitamente al otro socio. Sobre el concepto de onerosidad en el ámbito de la contratación pública, vid. VILALTA REIXACH, M.. (2012): *La encomienda de gestión*, Aranzadi, 2012, pp. 293 y 294.

⁹⁵ En efecto, en 2018, el Land de Berlín había publicado un anuncio previo para licitar el desarrollo del programa, por un montante de 3 500 000 euros. El anuncio de la ciudad de Colonia por anuncio de licitación (restringida) para la adaptación, puesta en marcha y mantenimiento del software cedido por el Land de Berlín tuvo un valor estimado de 2 000 000 de euros.

este específico supuesto, no debe desdeñarse su contenido, a la vista que la falta de cesión de los derechos de propiedad intelectual puede suponer un obstáculo a determinadas fórmulas de reutilización de aplicaciones informáticas.

En todo caso, más allá de los supuestos de cooperación interadministrativa, la cesión de tecnología entre entes públicos suscita interesantes cuestiones desde la óptica del Derecho de la competencia. Desde una perspectiva jurisprudencial es de interés la STJUE de 7 de noviembre de 2019, As. C-687/17 P. En aquel caso determinadas empresas solicitaron a la Comisión que declarase que la financiación de la Plataforma de Licitación Holandesa TEDNed constituía una ayuda de Estado ilegal, ya que especialmente en lo que respecta al módulo de presentación de ofertas, constituiría una actividad económica. Por tanto, a diferencia de la sentencia *ISE*, se trataría de una cesión *gratuita*, a la que no se aplican las Directivas sobre contratación pública. Sea como fuere, tanto la Comisión como el Tribunal General y finalmente el Tribunal de Justicia rechazan que la financiación de la Plataforma de licitación constituya una ayuda de Estado ilegal. Así, en la medida en que dicha Plataforma está conectada al ejercicio del poder público, el Tribunal de Justicia determinó que las actividades desarrolladas por TedNed carecen de naturaleza económica⁹⁶.

Por lo demás, puede plantearse igualmente qué relevancia tendrá la configuración de los derechos de propiedad intelectual cuando se encadenen sucesivos contratos vinculados a la solución contratada inicialmente. En otros términos, puede plantearse la duda de si, en caso de que los derechos de propiedad intelectual no se hayan cedido a

⁹⁶ Desde una perspectiva doctrinal es de gran interés la posición mantenida por GIMENO FELIU, J.M. (2019): *La Ley de Contratos del Sector Público 9/2017. Sus principales novedades, los problemas interpretativos y posibles soluciones*, Thomson-Reuters, pp. 197 y ss, quien sí considera que supuestos como el descrito podría existir una vulneración del Derecho de la competencia. En sentido opuesto, son igualmente relevantes las aportaciones de MARTÍN DELGADO, I. (2020): “Innovación tecnológica e innovación administrativa en la contratación pública”, en Isaac Martín Delgado y Jose Antonio Moreno (dir.), *Administración electrónica, transparencia y contratación pública*, Iustel, pp. 42 y ss. y MIRANZO DÍAZ, J. (2022): “El ‘modelo’ de plataformas de contratación pública en España: la autonomía administrativa y el derecho de la competencia”, *Revista Catalana de Dret Públic*, 64, pp. 140-157.

la Administración, puede esta adjudicar contratos sucesivos al mismo empresario a través de un procedimiento negociado sin publicidad. El art.168.a.2ª LCSP permite el uso del procedimiento negociado sin publicidad cuando por razones técnicas o artísticas o por cualquier otra razón relacionada con la protección de derechos de exclusividad, el contrato sólo puede encomendarse a un operador determinado. En todo caso, siempre que se trate de recurrir al procedimiento negociado sin publicidad porque sólo exista objetivamente un operador económico que pueda ejecutar el contrato, los poderes adjudicadores que se acojan a esta excepción deben motivar por qué no existen otras alternativas, como por ejemplo la utilización de otros canales de distribución incluso fuera del Estado miembro del poder adjudicador o la toma en consideración de obras, suministros o servicios desde el punto de vista funcional⁹⁷ (considerando 50). De esta forma, es evidente que la aplicación de este supuesto es fuertemente excepcional⁹⁸.

Cuando la situación de exclusividad se deba a razones técnicas, éstas deben definirse y justificarse rigurosamente para cada caso concreto. Entre estas razones cabe citar la práctica imposibilidad técnica de que otro operador económico alcance los resultados necesarios, o la necesidad de utilizar conocimientos técnicos, herramientas o medios específicos que sólo estén a disposición de un único operador económico. También pueden derivarse razones técnicas de los requisitos específicos en materia de interoperabilidad o de seguridad que deban cumplirse a fin de garantizar la idoneidad e las obras, suministros o servicios que vayan a contratarse⁹⁹. En todo caso, la ausencia de com-

⁹⁷ Considerando 50 de la Directiva 2014/24/UE del Parlamento Europeo y del Consejo, de 26 de febrero de 2014, sobre contratación pública y por la que se deroga la Directiva 2004/18/CE.

⁹⁸ Además, subraya el Informe 2/2016, de 6 de abril, de la Junta Consultiva de Contratación Administrativa de la Generalitat de Catalunya (Comisión Permanente) hay que tener en cuenta que esta causa de utilización del procedimiento negociado no sólo implica una restricción a la concurrencia y a la publicidad, sino también a la propia regla establecida en el TRLCSP relativa a la necesidad de solicitar ofertas, al menos, a tres empresas capacitadas para la realización del objeto del contrato, siempre que eso sea posible. Es justamente por este motivo que se tiene que utilizar en circunstancias muy excepcionales que deben motivarse y acreditarse suficientemente en el expediente de contratación.

⁹⁹ Considerando 50 de la Directiva 2014/24/UE del Parlamento Europeo y del Consejo, de 26 de febrero de 2014, sobre contratación pública y por la que se

petencia que permite la utilización del procedimiento negociado sin publicidad no debe ser consecuencia de una situación de exclusividad creada artificialmente por el propio poder adjudicador¹⁰⁰.

Por tanto, todos estos elementos han de tenerse en cuenta a la hora de valorar la licitud de una adjudicación a través del procedimiento negociado sin publicidad de un contrato relacionado con la contratación de una solución de IA. Y aunque la solución que se dé a cada caso concreto es ciertamente casuística, a la vista de la doctrina de los Tribunales de Recursos Contractuales, un parámetro interpretativo que puede resultar de especial trascendencia para discernir si puede utilizarse lícitamente el procedimiento negociado sin publicidad para la adjudicación de estos contratos estrechamente vinculados con otros anteriores es determinar si la adjudicación del primer contrato se realizó de forma competitiva¹⁰¹. De esta forma, cuando se trata de la contratación de soluciones de IA será muy relevante el diseño de la contratación inicial, en aspectos tales como el procedimiento de licitación empleado o el grado de detalle de las prescripciones técnicas. De esta forma, deberá promoverse al máximo la concurrencia, ya en la primera licitación. Y en el caso de que fueran necesarios desarrollos posteriores, deberán apurarse igualmente las posibilidades de recurrir de nuevo

deroga la Directiva 2004/18/CE.

¹⁰⁰ En relación con la aplicación de este criterio en la contratación de software, vid. entre otras la Resolución 13/2019, de 30 de enero, del Tribunal Catalán de Contratos del Sector Público, el Acuerdo 383/2021, de 26 de agosto del Tribunal Administrativo de Contratación Pública de Madrid y Resolución 307/2020, de 5 de marzo, del Tribunal Administrativo Central de Recursos Contractuales. En esta última resolución, el Tribunal tiene ocasión de afirmar que “aunque el órgano de contratación haya cumplido en cuanto a la forma y al fondo con los trámites necesarios para poder acudir legalmente a un procedimiento negociado sin publicidad, se recomienda la utilización preferente del procedimiento abierto, procedimiento catalogado por el legislador como ordinario, siempre que puedan existir dudas sobre la posible exclusividad en la prestación. En el momento en que se ha de acudir al mercado a preguntar sobre esta circunstancia, parece claro que no existe una evidente situación de exclusividad. Por tanto, en aras a no perjudicar la concurrencia e igualdad entre licitadores, es preferible acudir al procedimiento abierto.”

¹⁰¹ Sobre esta cuestión, GALLEGO CÓRCOLES, I. (2017): “La utilización del procedimiento negociado sin publicidad por razones de exclusividad técnica”, *Contratación Administrativa Práctica*, núm. 149.

a un procedimiento competitivo. Solo cuando pueda justificarse suficientemente que esta opción no es viable podrá acudir al procedimiento negociado sin publicidad.

3.1.4.- La regulación de la eventual cesión de derechos de propiedad intelectual en la LCSP

Finalmente, debe señalarse que la LCSP establece ciertas pautas sobre la cesión de derechos de propiedad intelectual, si bien permite que la regulación que se establezca en los pliegos sobre esta cuestión sea muy flexible.

En primer lugar, el art. 122 LCSP dispone que los pliegos podrán también especificar si va a exigirse la transferencia de derechos de propiedad intelectual o industrial, sin perjuicio de lo establecido en el artículo 308 respecto de los contratos de servicios.

El artículo 308 de la LCSP, por su parte, prevé que “salvo que se disponga otra cosa en los pliegos de cláusulas administrativas o en el documento contractual, los contratos de servicios que tengan por objeto el desarrollo y la puesta a disposición de productos protegidos por un derecho de propiedad intelectual o industrial llevará aparejada la cesión de este a la Administración contratante. En todo caso, y aun cuando se excluya la cesión de los derechos de propiedad intelectual, el órgano de contratación podrá siempre autorizar el uso del correspondiente producto a los entes, organismos y entidades pertenecientes al sector público”. Con carácter general, por tanto, en los contratos que lleven aparejada la creación de productos algorítmicos la propiedad intelectual se cederá a la Administración Pública, lo cual -señala MIRANZO¹⁰²- es una práctica habitual en las licitaciones que incluyen la realización o producción de nuevos sistemas por el adjudicatario¹⁰³. No obstante, como se desprende

¹⁰² MIRANZO DÍAZ, J. (2020): “Inteligencia artificial y contratación pública...”, *op.cit.*, p.

¹⁰³ Sobre la cesión de derechos de propiedad intelectual, resulta de interés la RTA-CRC 823/2019, de 11 de julio de 2019.

del tenor literal del art. 308 LCSP, la norma admite pacto en contrario¹⁰⁴.

Frente a lo anterior, los contratos en los que la prestación no suponga una creación o desarrollo de un sistema nuevo, sino la puesta a disposición, o adaptación, de un sistema de IA preexistente, deberían ser calificado como contrato de suministros¹⁰⁵. En estos supuestos, la norma no ofrece regla general sobre la transmisión o no de los derechos de propiedad intelectual, de forma que, tal y como recoge el art.122.2 LCSP, los pliegos podrán también especificar si va a exigirse la transferencia de derechos de propiedad intelectual o industrial¹⁰⁶.

Las estipulaciones que se establezcan respecto a la propiedad intelectual pueden tener un impacto estratégico nada desdeñable, en tanto en cuanto pueden afianzar o no el papel de las Administraciones públicas como sujetos catalizadores de la innovación. De hecho, el art.177 LCSP -en el ámbito del procedimiento de asociación para la innovación-, insiste en que el órgano de contratación ha de definir las disposiciones aplicables a los derechos de propiedad intelectual e industrial. A menudo, el órgano de contratación necesitará adquirir la licencia de uso, y no los derechos de propiedad intelectual. Y en este sentido, CERQUEIRA¹⁰⁷ ha destacado que si se permite a los innovadores mantener su propiedad intelectual, la Administración ofrecerá al sector privado, especialmente a las PYMES, incentivos adecuados

¹⁰⁴ Mantiene una interpretación diferente VESTRI, G. (2021): “La inteligencia artificial ante el desafío de la transparencia algorítmica. Una aproximación desde la perspectiva jurídico-administrativa”, *Revista Aragonesa de Administración Pública*, n° 56, p.384.

¹⁰⁵ El artículo 16.b) LCSP dispone que se considerarán contratos de suministro, los que tengan por objeto la adquisición y el arrendamiento de equipos y sistemas de telecomunicaciones o para el tratamiento de la información, sus dispositivos y programas, y la cesión del derecho de uso de estos últimos, en cualquiera de sus modalidades de puesta a disposición, a excepción de los contratos de adquisición de programas de ordenador desarrollados a medida, que se considerarán contratos de servicios. En relación con la interpretación de este precepto, vid. Informe 58/2018 de Junta Consultiva de Contratación Administrativa del Estado.

¹⁰⁶ MIRANZO DÍAZ, J. (2020): “Inteligencia artificial y contratación pública...”, *op.cit.*

¹⁰⁷ CERQUEIRA GOMES, P. (2014): “The innovative Innovation Partnerships under the new EU directives (2014)”, *Public Procurement Law Review*, 23(4), p.218.

para completar su solución innovadora, lo que parece ser el espíritu del considerando 49 de la Directiva.

En este sentido, en su día ya se subrayó que la reserva a favor de la Administración de los derechos de propiedad intelectual puede resultar un freno a la innovación, aunque en ocasiones esta reserva puede justificarse¹⁰⁸. En definitiva, a la hora de valorar la conveniencia o no de exigir la cesión de la propiedad intelectual habrá de tenerse en cuenta la relevancia que la promoción de la innovación tiene en el diseño de cada contratación.

Finalmente, hay que tener en cuenta que el art.157.2 LRJSP dispone que las administraciones podrán declarar de fuentes abiertas las aplicaciones, desarrolladas por sus servicios o que hayan sido objeto de contratación y de cuyos derechos de propiedad intelectual sean titulares, cuando de ello se derive una mayor transparencia en el funcionamiento de la Administración Pública. Frente a algunas interpre-

¹⁰⁸ Así en la Comunicación de la Comisión “la contratación precomercial: impulsar la innovación para dar a Europa servicios públicos de alta calidad y sostenibles”, Bruselas, 14.12.2007, COM(2007) 799 final ya se destacaba que el “desarrollo exclusivo” significa que el comprador público se reserva para su propio uso todos los resultados y beneficios del desarrollo (incluidos los derechos de propiedad intelectual). Esto hace que las empresas que desarrollan los productos o servicios no puedan utilizarlos con otros posibles clientes, lo que suele traducirse en precios más elevados. No obstante, el desarrollo exclusivo puede justificarse en algunos casos como, por ejemplo, cuando el comprador público «necesita» derechos exclusivos sobre los resultados de los proyectos (por ejemplo, en campos relacionados con la defensa o la seguridad que requieren mantener en secreto los resultados) o cuando el comprador público «es» el único cliente interesado (por ejemplo, en el desarrollo de equipos especiales específicos). Según los expertos, los compradores públicos europeos tienden a utilizar el desarrollo exclusivo, No obstante, en la mayoría de los casos, la «exclusividad» de los resultados del proyecto no es imprescindible para los compradores públicos, ya que ellos no son más que uno de los muchos usuarios potenciales de la solución desarrollada. Por otra parte, los compradores públicos suelen infravalorar los costes y los esfuerzos adicionales necesarios para cosechar los beneficios de los resultados. La carga de los altos costes y los riesgos del desarrollo exclusivo no suele estar justificada salvo cuando los compradores públicos tengan un mandato y planes concretos para explotar comercialmente los resultados de la investigación. En estos casos, el desarrollo exclusivo puede obstaculizar la innovación. La asignación exclusiva de derechos al comprador público elimina los incentivos para que las empresas inviertan más en comercialización.

taciones¹⁰⁹, no parece que el precepto obligue a que la Administración adquiera los derechos de propiedad intelectual sobre todas las aplicaciones: más bien el precepto se limita a admitir que solo en estos casos podrá declararse una aplicación de fuente abierta.

3.2.- Mecanismos para garantizar la supervisión, auditoría y control de los sistemas algorítmicos

3.2.1.- Planteamiento

El art.41 de la Ley 40/2015, de 1 de octubre, de régimen jurídico del sector público dispone que “en caso de actuación administrativa automatizada deberá establecerse previamente el órgano u órganos competentes, según los casos, para la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente”. Este precepto halla cierto nivel de desarrollo en el ámbito estatal, de modo que se obliga al titular del órgano administrativo competente por razón de la materia o del órgano ejecutivo competente del organismo o entidad de derecho público, según corresponda, a establecer medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos de las personas interesadas¹¹⁰.

Aunque puedan existir diferentes interpretaciones acerca del alcance del art.41 de la Ley 40/2015, como señala GAMERO, la norma articula la obligación de que los sistemas de actuación administrativa automatizada e IA sean objeto de “supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente”¹¹¹. El precepto se alinea así con las exigencias que se vienen expresando por la literatura especializada, que reclama la necesidad

¹⁰⁹ Para PONCE, de este precepto se desprende que las administraciones deben utilizar programario abierto o declararlo de fuentes abiertas. Vid. PONCE SOLÉ, J. (2019): “Inteligencia artificial, Derecho administrativo..”, *op. cit.*.

¹¹⁰ Art.13 del Real Decreto 203/2021, de 30 de marzo, por el que se aprueba el Reglamento de actuación y funcionamiento del sector público por medios electrónicos.

¹¹¹ GAMERO CASADO, E. (2021): “Compliance (o cumplimiento normativo)...”, *op. cit.*

de que estas soluciones tecnológicas se sometan a supervisión humana, auditoría y control, desde la fase de diseño hasta su entrada en servicio efectiva, y, además, con revisiones periódicas que verifiquen su adecuado funcionamiento¹¹².

3.2.2.- Supervisión humana de los algoritmos

Como ya se ha señalado, el art.41 Ley 40/2015, de 1 de octubre, de régimen jurídico del sector público, exige que los casos de actuación automatizada los sistemas de soporte estén sometidos a supervisión. Esta supervisión, tal y como ha sido destacado en la doctrina¹¹³ y se reconoce en las normas de *soft law*¹¹⁴, consiste en una supervisión *humana*. Como subraya GAMERO CASADO¹¹⁵, esta supervisión humana no se circunscribe al momento inicial de implementación de la solución de IA: debe proyectarse sobre todas las fases de su ciclo de vida; debe alcanzar también la fase de ejecución del contrato, de forma que se revise periódicamente el funcionamiento del sistema, actualizando y perfeccionando los datos de que se nutre y observando los resultados para evitar desviaciones.

En realidad, el artículo 14 RIA prevé que los sistemas de IA de alto riesgo se diseñen y desarrollen de modo que puedan ser vigilados de manera efectiva por personas físicas durante el período que estén en uso, lo que incluye dotarlos de una herramienta de interfaz humano-máquina adecuada, entre otros aspectos¹¹⁶. En este sentido,

¹¹² GAMERO CASADO, E. (2021): “Compliance (o cumplimiento normativo)...”, *op. cit.*

¹¹³ Por todos, PONCE SOLÉ, J. (2019): “Inteligencia artificial, Derecho administrativo..”, *op. cit.*

¹¹⁴ En el ámbito nacional, el art. XXV de la Carta de Derechos Digitales reconoce el derecho a solicitar una supervisión e intervención humana.

¹¹⁵ GAMERO CASADO, E. (2021): “Compliance (o cumplimiento normativo)...”, *op. cit.*

¹¹⁶ El precepto añade que el objetivo de la vigilancia humana será prevenir o reducir al mínimo los riesgos para la salud, la seguridad o los derechos fundamentales que pueden surgir cuando un sistema de IA de alto riesgo se utiliza conforme a su finalidad prevista o cuando se le da un uso indebido razonablemente previsible, en particular cuando dichos riesgos persisten a pesar de aplicar otros requisitos establecidos en el presente capítulo.

La vigilancia humana se garantizará de una de las siguientes maneras o de ambas:

como explica el considerando 48 de la propuesta, los sistemas de IA de alto riesgo deben diseñarse y desarrollarse de tal modo que su funcionamiento pueda ser vigilado por personas físicas. A tal fin, el proveedor del sistema debe definir las medidas adecuadas de vigilancia humana antes de su introducción en el mercado o puesta en servicio. Cuando proceda, dichas medidas deben garantizar, en concreto, que el sistema presente limitaciones operativas que el propio sistema no pueda desactivar, que responda al operador humano, y que las personas físicas a quienes se haya encomendado la vigilancia humana posean las competencias, la formación y la autoridad necesarias para desempeñar esa función.

De manera paralela a la regulación proyectada en el art. 14 RIA¹¹⁷, el art. 7 del borrador de cláusulas auspiciado por la Comisión Europea establece que:

a) el proveedor definirá las medidas de vigilancia humana y, cuando sea técnicamente viable, las integrará en el sistema de IA de alto riesgo antes de su introducción en el mercado o puesta en servicio;

b) el proveedor definirá las medidas de vigilancia humana, que serán adecuadas para que las lleve a cabo el usuario, antes de la introducción del sistema de IA de alto riesgo en el mercado o de su puesta en servicio.

¹¹⁷ En concreto, los apartados 4 y 5 del art. 14 RIA disponen que: “4. Las medidas mencionadas en el apartado 3 permitirán que las personas a quienes se encomienda la vigilancia humana puedan, en función de las circunstancias:

a) entender por completo las capacidades y limitaciones del sistema de IA de alto riesgo y controlar debidamente su funcionamiento, de modo que puedan detectar indicios de anomalías, problemas de funcionamiento y comportamientos inesperados y ponerles solución lo antes posible;

b) ser conscientes de la posible tendencia a confiar automáticamente o en exceso en la información de salida generada por un sistema de IA de alto riesgo («sesgo de automatización»), en particular con aquellos sistemas que se utilizan para aportar información o recomendaciones con el fin de que personas físicas adopten una decisión;

c) interpretar correctamente la información de salida del sistema de IA de alto riesgo, teniendo en cuenta en particular las características del sistema y las herramientas y los métodos de interpretación disponibles;

d) decidir, en cualquier situación concreta, no utilizar el sistema de IA de alto riesgo o desestimar, invalidar o revertir la información de salida que este genere;

e) intervenir en el funcionamiento del sistema de IA de alto riesgo o interrumpir el sistema accionando un botón específicamente destinado a tal fin o mediante un procedimiento similar.

“7.1 El proveedor garantiza que el sistema de IA se diseñará y desarrollará de tal manera, incluso con herramientas adecuadas de interfaz hombre-máquina, que pueda ser supervisado eficazmente por personas físicas.

7.2. La supervisión humana tendrá como objetivo prevenir o minimizar los riesgos para la salud seguridad o los derechos fundamentales que puedan surgir cuando el sistema de IA se utilice de acuerdo con su finalidad prevista o en condiciones de uso indebido razonablemente previsible, en particular cuando dichos riesgos persistan a pesar de la aplicación de otros requisitos establecidos en las presentes cláusulas.

7.3. El proveedor garantiza que, antes de la entrega, se tomarán las medidas adecuadas integradas en el Sistema de IA y adoptadas para garantizar la supervisión humana. Estas medidas, que podrían incluir, entre otras, la formación de los empleados de ORGANIZACIÓN PÚBLICA, permitirán a las personas a las que se asigne la supervisión humana asignada hagan lo siguiente, según las circunstancias:

a. comprender plenamente las capacidades y limitaciones del sistema de IA y ser capaces de supervisar debidamente su funcionamiento, de modo que los indicios de anomalías disfunciones y rendimientos inesperados puedan ser detectados y se aborden tan pronto como sea posible;

b. ser consciente de la posible tendencia a confiar o exagerar automáticamente.

b. ser consciente de la posible tendencia a confiar o sobrepasar automáticamente los resultados obtenidos por el sistema de IA (“sesgo de automatización”), en particular si el sistema de IA se utiliza para proporcionar información o recomendaciones para las decisiones que deben tomar las personas físicas;

c. ser capaz de interpretar correctamente los resultados del sistema de IA, teniendo en cuenta teniendo en cuenta, en particular, las características del sistema y las herramientas y métodos de interpretación disponibles;

d. poder decidir, en cualquier situación concreta, no utilizar el sistema de IA o ignorar, anular o invertir los resultados del sistema de IA del Sistema de IA;

e. poder intervenir en el funcionamiento del Sistema de IA o interrumpir el

5. En el caso de los sistemas de IA mencionados en el punto 1, letra a), del anexo III, las medidas que figuran en el apartado 3 garantizarán además que el usuario no actúe ni tome ninguna decisión sobre la base de la identificación generada por el sistema, salvo que un mínimo de dos personas físicas la hayan verificado y confirmado,”

e. poder intervenir en el funcionamiento del sistema de IA o interrumpirlo mediante un botón de “parada” o un procedimiento similar”.

Prácticamente este artículo reitera el contenido del art.14 RIA¹¹⁸, con la única salvedad que añade que entre las prestaciones del contrato puede incluirse la formación al personal de la Administración pública, lo que resultará imprescindible para garantizar que este personal tiene capacidad real de supervisar el sistema.

Por lo demás, el borrador de cláusulas tipo va más a allá que las cláusulas de la Ciudad de Ámsterdam que le han servido de precedente. En estas, el requisito de la supervisión humana se encuentra incluido en el de la explicabilidad. Parece no obstante más alineado con nuestro Derecho que se incluya también esta vertiente de la supervisión incluida en lo que actualmente es un borrador auspiciado por la Comisión, que no se ciñe a una decisión concreta sino al funcionamiento del sistema.

¹¹⁸ En concreto, los apartados 4 y 5 del art. 14 RIA disponen que: “4. Las medidas mencionadas en el apartado 3 permitirán que las personas a quienes se encomienda la vigilancia humana puedan, en función de las circunstancias:

a) entender por completo las capacidades y limitaciones del sistema de IA de alto riesgo y controlar debidamente su funcionamiento, de modo que puedan detectar indicios de anomalías, problemas de funcionamiento y comportamientos inesperados y ponerles solución lo antes posible;

b) ser conscientes de la posible tendencia a confiar automáticamente o en exceso en la información de salida generada por un sistema de IA de alto riesgo («sesgo de automatización»), en particular con aquellos sistemas que se utilizan para aportar información o recomendaciones con el fin de que personas físicas adopten una decisión;

c) interpretar correctamente la información de salida del sistema de IA de alto riesgo, teniendo en cuenta en particular las características del sistema y las herramientas y los métodos de interpretación disponibles;

d) decidir, en cualquier situación concreta, no utilizar el sistema de IA de alto riesgo o desestimar, invalidar o revertir la información de salida que este genere;

e) intervenir en el funcionamiento del sistema de IA de alto riesgo o interrumpir el sistema accionando un botón específicamente destinado a tal fin o mediante un procedimiento similar.

5. En el caso de los sistemas de IA mencionados en el punto 1, letra a), del anexo III, las medidas que figuran en el apartado 3 garantizarán además que el usuario no actúe ni tome ninguna decisión sobre la base de la identificación generada por el sistema, salvo que un mínimo de dos personas físicas la hayan verificado y confirmado,”

3.2.3.- Control y auditoría del sistema

Como ha señalado la doctrina, la supervisión no debe reservarse en exclusiva a la toma de la decisión en sentido estricto, sino que también puede producirse *ex ante* al analizar los riesgos inherentes al uso de un sistema de inteligencia administrativa en la fase de resolución de un procedimiento administrativo o prever mecanismos de certificación o auditoría o *ex post* al articular mecanismos ad hoc para el control de las decisiones adoptadas de manera automatizada¹¹⁹. De manera muy especial en los supuestos en que se implanten sistemas basados en el *machine learning*, las exigencias y garantías que se implementen deberían ser no solo objeto de aprobación inicial, sino que, además, tendrían que estar sometidas a sistemas de evaluación periódicos, a fin de garantizar que la propia evolución de los algoritmos sigue ajustándose a las premisas iniciales o, en su caso, han de ser actualizadas a fin de que las evaluaciones iniciales conserven plenamente su sentido¹²⁰.

Aunque no es evidente a qué se refiere el art. 41 LRJSP cuando utiliza los términos “control” y “auditoría”, lo cierto es que las obligaciones sobre “sistemas de gestión de riesgos” y “auditorías” que se prevén en el proyecto de cláusulas tipo de la Comisión Europea responden bien a esta lógica de un control previo a la implantación y a un sistema de seguimiento. Como en tantas otras ocasiones, el sistema instaurado parte del que en su día estableció la Ciudad de Ámsterdam.

Las cláusulas de la ciudad de Ámsterdam imponen al contratista, en primer lugar, la obligación de elaborar una estrategia de gestión de riesgos. En este sentido se dispone que:

Artículo 6– Estrategia de gestión de riesgos en el desarrollo del sistema algorítmico

6.1 En el desarrollo del sistema algorítmico, el contratista aplicará una estrategia de gestión de riesgos común y actualizada, adecuada a la naturaleza del sistema algorítmico. A efectos de la aplicación de dicha estrategia de gestión de riesgos, el contratista identificará los principales riesgos

¹¹⁹ CERRILLO I MARTÍNEZ, A. (2019): “Robots, asistentes virtuales..”, *op. cit.*, p. 296.

¹²⁰ VALERO TORRIJOS, J. (2019): “Las garantías jurídicas..”, *op. cit.*, p. 91.

que puedan producirse cuando [NOMBRE DE LA ORGANIZACIÓN] utilice el sistema algorítmico y adoptará medidas para que los riesgos identificados sean gestionables. Al identificar los riesgos, el contratista prestará atención, en su caso, al riesgo de incumplimiento de una o varias de las obligaciones mencionadas en el artículo 2 y en el artículo 4, a los riesgos relacionados con la discriminación, a la posibilidad de que [NOMBRE DE LA ORGANIZACIÓN] controle el sistema algorítmico y a la protección de datos.

6.2 El contratista aplicará y documentará la estrategia de gestión de riesgos descrita en el artículo 6.1 de forma que la auditoría a la que se refiere el artículo 8 pueda verificar si el contratista ha cumplido la obligación descrita en el artículo 6.1.

6.3 Si el contratista ha desarrollado el sistema algorítmico antes de la celebración del contrato, y en la medida en que lo haya hecho, el contratista garantiza que las medidas descritas en el artículo 6.1 y/o en el artículo 6.2 ya han sido adoptadas. El contratista presentará pruebas de ello a petición de [NOMBRE DE LA ORGANIZACIÓN].

Por su parte, el art.9 RIA¹²¹, prevé que se establezca, documente y mantenga un sistema de gestión de riesgos asociado a los sistemas de IA de alto riesgo.

De manera paralela, el artículo 2 del borrador de las cláusulas tipo de la auspiciado por la Comisión Europea dispone que el proveedor garantizará que, antes de la entrega del sistema de IA, se establezca un sistema de gestión de riesgos en relación con el sistema de IA. Este sistema comprenderá, como mínimo, los siguientes elementos:

¹²¹ Como señala el precepto, el sistema de gestión de riesgos consistirá en un proceso iterativo continuo que se llevará a cabo durante todo el ciclo de vida de un sistema de IA de alto riesgo, el cual requerirá actualizaciones sistemáticas periódicas. Constará de las siguientes etapas:

- a) la identificación y el análisis de los riesgos conocidos y previsibles vinculados a cada sistema de IA de alto riesgo;
- b) la estimación y la evaluación de los riesgos que podrían surgir cuando el sistema de IA de alto riesgo en cuestión se utilice conforme a su finalidad prevista y cuando se le dé un uso indebido razonablemente previsible;
- c) la evaluación de otros riesgos que podrían surgir a partir del análisis de los datos recogidos con el sistema de seguimiento posterior a la comercialización al que se refiere el artículo 61;
- d) la adopción de medidas oportunas de gestión de riesgos con arreglo a lo dispuesto en los apartados siguientes.

- a. identificación y análisis de los riesgos conocidos y previsibles para la salud seguridad y los derechos fundamentales que puedan surgir a la luz de la finalidad prevista del sistema de IA;
- b. estimación y evaluación de los riesgos que pueden surgir cuando el sistema de IA se utiliza de acuerdo con su finalidad prevista y en condiciones de uso indebido razonablemente previsible;
- c. la evaluación de otros riesgos que puedan surgir;
- d. la adopción de medidas adecuadas de gestión de riesgos de acuerdo con las disposiciones de los apartados siguientes.

Se reitera así prácticamente el contenido del 9 RIU, si bien se añade la obligación de garantizar, antes de la entrega del Sistema de IA, que el sistema de IA se somete a pruebas para verificar si cumple con las cláusulas y si las medidas de gestión de riesgos mencionadas en él son efectivas a la luz de la finalidad prevista. Si lo solicita el organismo público, el proveedor está obligado a probar el sistema de IA en el entorno de este organismo¹²². Además, en las cláusulas tipo se establece la obligación de entregar este documento de gestión de riesgos. Lo que parece una práctica muy recomendable en nuestro Derecho, siempre que se adapte dicho documento a las exigencias del principio de proporcionalidad¹²³.

Por su parte, la *auditabilidad* se refiere a la capacidad de un sistema de IA de someterse a la evaluación de sus algoritmos, datos y

¹²² En este sentido, el 9.7 RIA dispone que “las pruebas de los sistemas de IA de alto riesgo se realizarán, según proceda, en cualquier momento del proceso de desarrollo y, en todo caso, antes de su introducción en el mercado o puesta en servicio”.

¹²³ En relación con esta cuestión, vid. *Standard Clauses for Procurement of Trustworthy Algorithmic Systems*, City of Amsterdam, p.21. La estrategia de gestión de riesgos a seguir debe ser adecuada a la naturaleza del sistema algorítmico. Esto significa que un sistema algorítmico que vaya a ser utilizado en una situación en la que la violación de los derechos de los ciudadano pueda ser potencialmente mayor, tendrá que ser sometido a una estrategia de gestión de riesgos más intensa que un sistema algorítmico que vaya a ser utilizado en una situación en la que no lo sea.

Por lo demás, como se señala en dicho documento, en las Directrices éticas para una IA fiable, elaboradas por el grupo de expertos de alto nivel sobre inteligencia artificial de la Comisión Europea (2019) se establece una lista de evaluación que puede ser parte del sistema de gestión de riesgos.

procesos de diseño. Constituye uno de los siete requisitos que cumplir cualquier sistema de IA fiable¹²⁴. En las cláusulas de la Ciudad de Ámsterdam se incorpora como exigencia que se realicen determinadas auditorías, a través de las cuales se pretende de verificar si el contratista, en la ejecución del contrato, ha cumplido todas las obligaciones impuestas por la Administración. Se trata por tanto de un control técnico, que ha de reflejarse en un informe que podrá ser objeto de publicación por parte de la Administración. Se prevé que la Administración pueda realizar auditorías anuales, realizadas total o parcialmente por auditores externos.

El contenido de las cláusulas tipo de la Ciudad de Ámsterdam se ha trasladado a la propuesta de las cláusulas tipo de la Comisión Europea de manera prácticamente literal, si bien con un contenido opcional¹²⁵. Ello es así por estrecha vinculación de este documento con

¹²⁴ Directrices éticas para una IA fiable, elaboradas por el grupo de expertos de alto nivel sobre inteligencia artificial de la Comisión Europea (2019), p.48.

¹²⁵ Artículo 18. Auditoría

1. El proveedor está obligado a cooperar en una auditoría u otro tipo de inspección que por o en nombre de [NOMBRE DEL ORGANISMO PÚBLICO] para evaluar si el proveedor cumple con sus obligaciones establecidas en estas cláusulas en todo momento. Dicha colaboración incluirá el suministro de toda la información requerida por [NOMBRE DEL ORGANISMO PÚBLICO], proporcionando una visión del sistema de gestión de riesgos implantado, la disponibilidad del personal del proveedor para entrevistas y facilitar el acceso a las instalaciones del proveedor o en nombre de [NOMBRE DEL ORGANISMO PÚBLICO] para evaluar si el proveedor cumple con sus obligaciones establecidas en estas cláusulas en todo momento. Dicha colaboración incluirá el suministro de toda la información requerida por [NOMBRE DEL ORGANISMO PÚBLICO], proporcionando una visión del sistema de gestión de riesgos sistema de gestión de riesgos implantado, la disponibilidad del personal del proveedor para entrevistas y facilitar el acceso a las instalaciones del proveedor.

2. [NOMBRE DEL ORGANISMO PÚBLICO] preparará, o hará que se prepare, un informe en el que se recojan las conclusiones de la auditoría. En el informe, [NOMBRE DEL ORGANISMO PÚBLICO] dejará constancia de la medida en que el proveedor cumple con las obligaciones derivadas del Contrato. Si [NOMBRE DEL ORGANISMO PÚBLICO ORGANISMO PÚBLICO] establece que el proveedor no cumple con las obligaciones previstas en este artículo, el proveedor estará obligado a subsanar los defectos identificados por [NOMBRE DEL ORGANISMO PÚBLICO] en el plazo razonable fijado por [NOMBRE DEL ORGANISMO PÚBLICO] en el informe. Si el proveedor no Si el proveedor no subsana los defectos identificados por [NOMBRE DEL ORGANISMO PÚBLICO]

el RIA, en el que se adopta como solución para los sistemas de IA de alto riesgo mencionados en el anexo III, “una evaluación integral de la conformidad ex ante mediante controles internos, combinada con una supervisión ex post estricta, dado que la intervención reguladora se encuentra en una fase temprana, que el sector de la IA es muy innovador y que apenas están empezando a acumularse los conocimientos necesarios para llevar a cabo auditorías. Para evaluar los sistemas de IA de alto riesgo «independientes» mediante controles internos, sería necesario cumplir ex ante de manera plena, efectiva y debidamente documentada todos los requisitos del Reglamento, así como los sólidos sistemas de gestión de la calidad y los riesgos y el seguimiento posterior a la comercialización”.¹²⁶ De esta forma, en la propuesta se

CO] en el plazo fijado en el informe para subsanar dichos defectos, el proveedor incurrirá en incumplimiento por ministerio de la ley.

3. [NOMBRE DEL ORGANISMO PÚBLICO] tendrá derecho a publicar las conclusiones del informe mencionado en el artículo 18.2.

4. [NOMBRE DEL ORGANISMO PÚBLICO] tendrá derecho a realizar, o hacer que se realice una auditoría una vez por año natural.

5. [NOMBRE DEL ORGANISMO PÚBLICO] puede decidir que la totalidad o parte de la auditoría realizada por un auditor independiente.

6. Los costes del auditor contratado por [NOMBRE DEL ORGANISMO PÚBLICO], si los hubiera, serán pagados por [NOMBRE DEL ORGANISMO PÚBLICO] pagará al proveedor unos honorarios razonables por los costes en los que deba incurrir en que incurra el proveedor en el marco de la auditoría. En ningún caso una disputa sobre el importe de dichos honorarios dará derecho al proveedor a suspender sus obligaciones en virtud de las presentes cláusulas. No se adeudarán dichos honorarios por parte de [NOMBRE DEL ORGANISMO PÚBLICO] si la auditoría revela que el proveedor ha incumplido sus obligaciones en virtud de las presentes cláusulas.

¹²⁶ Dispone el primer apartado del artículo 17 (sistema de gestión de la calidad) Los proveedores de sistemas de IA de alto riesgo establecerán un sistema de gestión de la calidad que garantice el cumplimiento del presente Reglamento. Dicho sistema se documentará de manera sistemática y ordenada mediante políticas, procedimientos e instrucciones escritas e incluirá, al menos, los siguientes aspectos:

- a) una estrategia para el cumplimiento reglamentario, incluido el cumplimiento de los procedimientos de evaluación de la conformidad y de los procedimientos de gestión de las modificaciones de los sistemas de IA de alto riesgo;
- b) las técnicas, los procedimientos y las actuaciones sistemáticas que se utilizarán en el diseño y el control y la verificación del diseño del sistema de IA de alto riesgo;
- c) las técnicas, los procedimientos y las actuaciones sistemáticas que se utilizarán en el desarrollo y el control y el aseguramiento de la calidad del sistema de IA de alto riesgo;

establece la obligación de que el proveedor establezca un sistema de gestión de la calidad, en los términos previstos en el art. 17 RIU, la necesidad realizar una evaluación de conformidad (artículo 19 RIU)¹²⁷ y de un plan de seguimiento posterior (art. 61 RIU)¹²⁸.

-
- d) los procedimientos de examen, prueba y validación que se llevarán a cabo antes, durante y después del desarrollo del sistema de IA de alto riesgo, así como la frecuencia con que tendrán lugar;
 - e) las especificaciones técnicas, incluidas las normas, que se aplicarán y, cuando las normas armonizadas pertinentes no se apliquen en su totalidad, los medios que se utilizarán para velar por que el sistema de IA de alto riesgo cumpla los requisitos establecidos en el capítulo 2 del presente título;
 - f) los sistemas y procedimientos de gestión de datos, lo que incluye su recopilación, análisis, etiquetado, almacenamiento, filtrado, prospección, agregación, conservación y cualquier otra operación relacionada con los datos que se lleve a cabo antes de la introducción en el mercado o puesta en servicio de sistemas de IA de alto riesgo y con ese fin;
 - g) el sistema de gestión de riesgos que se menciona en el artículo 9;
 - h) el establecimiento, la implantación y el mantenimiento de un sistema de seguimiento posterior a la comercialización con arreglo al artículo 61;
 - i) los procedimientos asociados a la notificación de incidentes graves y defectos de funcionamiento con arreglo al artículo 62;
 - j) la gestión de la comunicación con las autoridades nacionales competentes; las autoridades competentes, incluidas las sectoriales, que permiten acceder a datos o facilitan el acceso a ellos; los organismos notificados; otros operadores; los clientes, u otras partes interesadas;
 - k) los sistemas y procedimientos destinados a llevar un registro de toda la documentación e información pertinente;
 - l) la gestión de los recursos, incluida la seguridad de las medidas relacionadas con el suministro;
 - m) un marco de rendición de cuentas que defina las responsabilidades del personal directivo y de otra índole en relación con todos los aspectos enumerados en este apartado.

¹²⁷ Según el art. 19.1, evaluación de la conformidad, “los proveedores de sistemas de IA de alto riesgo se asegurarán de que sus sistemas sean sometidos al procedimiento de evaluación de la conformidad oportuno, de conformidad con el artículo 43, antes de su introducción en el mercado o puesta en servicio. Cuando dicha evaluación de la conformidad demuestre que los sistemas de IA cumplen los requisitos establecidos en el capítulo 2 del presente título, sus proveedores elaborarán una declaración UE de conformidad con arreglo al artículo 48 y colocarán el marcado CE de conformidad con arreglo al artículo 49”.

¹²⁸ Según el artículo 61 RIA, Los proveedores establecerán y documentarán un sistema de seguimiento posterior a la comercialización de forma proporcionada a la naturaleza de las tecnologías de inteligencia artificial y a los riesgos de los sistemas de IA de alto riesgo.

En nuestro Derecho, resulta preferible la opción de prever en los pliegos la obligación de someter el sistema de inteligencia artificial a auditorías periódicas, tal y como exige el art.41 LRJSP, siguiendo por tanto el modelo de la *Ciudad de Ámsterdam*. Por lo demás, el conocimiento sobre la forma de realizar auditorías algorítmicas es ya existente. Entre otros documentos, puede citarse el libro blanco sobre la auditoria de algoritmos, elaborado por las cortes de auditores de Finlandia Alemania, Países Bajos, Noruega y Reino Unido¹²⁹. En el ámbito estrictamente español, la AEPD ha elaborado un documento de requisitos para Auditorías de Tratamientos que incluyan IA¹³⁰.

2.El sistema de seguimiento posterior a la comercialización recabará, documentará y analizará de manera activa y sistemática datos pertinentes proporcionados por usuarios o recopilados a través de otras fuentes sobre el funcionamiento de los sistemas de IA de alto riesgo durante toda su vida útil, y permitirá al proveedor evaluar el cumplimiento de los requisitos establecidos en el título III, capítulo 2, por parte de los sistemas de IA.

3.El sistema de seguimiento posterior a la comercialización se basará en un plan de seguimiento posterior a la comercialización. El plan de seguimiento posterior a la comercialización formará parte de la documentación técnica a que se refiere el anexo IV. La Comisión adoptará un acto de ejecución en el que se establecerán disposiciones detalladas que constituyan un modelo para el plan de seguimiento posterior a la comercialización y la lista de elementos que deberán incluirse en él.

4.En el caso de los sistemas de IA de alto riesgo regulados por los actos legislativos a que hace referencia el anexo II, cuando ya se hayan establecido un sistema y un plan de seguimiento posteriores a la comercialización con arreglo a dicha legislación, los elementos descritos en los apartados 1, 2 y 3 se integrarán en dicho sistema y plan, según proceda.

El párrafo primero también se aplicará a los sistemas de IA de alto riesgo a que se refiere el punto 5, letra b), del anexo III introducidos en el mercado o puestos en servicio por entidades de crédito reguladas por la Directiva 2013/36/UE.

¹²⁹ Disponible en Auditing machine learning algorithms (auditingalgorithms.net). La información sobre el proyecto que difunde la National Audit Office del Reino Unido está disponible en How to audit artificial intelligence models–National Audit Office (NAO) insight

La Guía sobre auditoria de algoritmos elaborada por ISACA (Information Systems Audit and Control Association), *Auditing Artificial Intelligence*, está disponible en Auditing AI – the way for trustworthy AI | Futurium (europa.eu).

¹³⁰ Disponible en Requisitos para Auditorías de Tratamientos que incluyan IA (aepd.es).

3.3.- Cláusulas relativas a la protección de datos

Con carácter general, el art.122 LCSP dispone que los pliegos deberán mencionar expresamente la obligación del futuro contratista de respetar la normativa vigente en materia de protección de datos. Los riesgos de incumplir esta normativa, en el ámbito de los sistemas de inteligencia artificial que, precisamente, se basan en un uso masivo de datos, es ya conocida¹³¹. Por lo demás, como ya se ha adelantado, la AEPD ha elaborado una *Guía de adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial*¹³², que deberá ser obviamente tenida en cuenta por los órganos de contratación¹³³.

Por lo demás, en las cláusulas tipo de la Ciudad de Ámsterdam se establecen dos artículos relativos a los datos: el art.2, que regula la calidad de los datos y el art. 3 que se refiere a la propiedad de los datos.

En relación con la calidad de los datos, estas cláusulas impone al contratista dos obligaciones¹³⁴: a) asegurar que el tratamiento cumpla

¹³¹ Sobre esta cuestión, de manera sintética, vid. CERRILLO I MARTÍNEZ, A. (2019): “Robots, asistentes virtuales.”, *op. cit.*, pp.297-299. Con carácter más extenso, en esta misma obra, VALERO TORRIJOS, J., “La protección de datos de carácter personal: art.18 CE”.

¹³² La guía está fechada en Febrero de 2020. Disponible en <https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia.pdf>

¹³³ Vid. igualmente el documento requisitos para Auditorias de Tratamientos que incluyan IA, de la AEPD, al que hicimos referencia en el epígrafe anterior.

¹³⁴ El tenor del artículo 2 es el siguiente (traducción de la autora): “Si, y en la medida en que el sistema algorítmico se desarrolle sobre la base de datos proporcionados por NOMBRE DE LA ORGANIZACIÓN al contratista, éste tomará las medidas que razonablemente puedan esperarse de él para garantizar que los datos utilizados en el desarrollo del sistema algorítmico sean analizados, estructurados y/o editados

a. de acuerdo con un enfoque motivado, cuyo propósito incluye, sin limitación, evitar la distorsión, las inexactitudes, los errores, las equivocaciones y los prejuicios no deseados (“sesgo”) contruidos socialmente en dichos datos en la medida de lo posible;

b. de manera que se cumplan las leyes y reglamentos aplicables.

2.2 Si y en la medida en que el sistema algorítmico se desarrolle sobre la base de datos no proporcionados por NOMBRE DE LA ORGANIZACIÓN al contratista, éste garantizará que los datos utilizados en el desarrollo del sistema algorítmico serán recogidos, analizados, estructurados y/o editados

a. de acuerdo con un enfoque motivado, cuyo propósito incluye, sin limitación, evitar la distorsión, las inexactitudes, los errores, las equivocaciones y los prejuicios

las leyes y reglamentos aplicables¹³⁵ y b) garantizar que los datos utilizados en el desarrollo del sistema algorítmico serán recogidos, analizados, estructurados y/o editados de acuerdo “con un enfoque motivado”. La primera de dichas obligaciones debería estar prevista expresamente en los pliegos en virtud del art.122 LCSP. Incorporar la segunda, como se explicará a continuación, resultará igualmente de interés.

Como se detalla en el documento en la que se contienen las cláusulas de la Ciudad de Ámsterdam, la calidad de los datos utilizados puede ser crucial para el rendimiento de un sistema algorítmico. Cuando se recogen los datos, éstos pueden contener distorsiones, inexactitudes, errores y fallos contruidos socialmente. Las medidas que debe tomar el contratista deben centrarse en evitarlo. Otro objetivo del requisito de que la (recogida), el análisis, la estructuración y la edición de los datos deben tener lugar según un enfoque motivado es evitar el sesgo en los datos en la medida de lo posible¹³⁶. Así, se añade, los conjuntos de datos utilizados (tanto para la formación como para la realización del trabajo) pueden contener distorsiones históricas inadvertidas, estar incompletos o contener modelos de mala gobernanza. Si se mantiene esa distorsión, pueden producirse prejuicios y discriminaciones inadvertidos, directos o indirectos, contra determinados grupos o determinadas personas, lo que podría reforzar los prejuicios y la marginación en el uso del sistema algorítmico. Cualquier distorsión demostrable y discriminatoria debe ser identificada y, cuando sea posible, eliminada por el contratista lo antes posible.

Por otro lado, en el art.3 se regula lo referente a la propiedad de los datos¹³⁷, previendo acuerdos contractuales sobre el uso de los datos,

cios no deseados (“sesgo”) contruidos socialmente en dichos datos en la medida de lo posible;

b. de manera que se cumplan las leyes y reglamentos aplicables.”

¹³⁵ Entre estas se encuentran, obviamente, las normas sobre protección de datos.

¹³⁶ Añade el documento que lo que constituye un sesgo tendrá que evaluarse en cada caso individual. En cualquier caso, se considerará que existe un sesgo si el uso de los datos conduce, o puede conducir, a una forma de discriminación prohibida en virtud de las leyes y reglamentos aplicables, incluido el artículo 14 del Convenio Europeo de Derechos Humanos.

¹³⁷ Artículo 3–Derechos sobre los datos

3.1 Todos los derechos relativos a los datos que NOMBRE DE LA ORGANIZACIÓN deba proporcionar al contratista en virtud del Contrato corresponderán

su entrega y su destrucción¹³⁸. Así, el art.3 distingue la situación en la que la Administración proporciona datos al contratista (art.3.1) y de la situación en la que los datos se recogen o crean en el contexto de la ejecución del contrato (art.3.2). En ambos supuestos, el principio básico es que los derechos sobre los datos corresponden exclusivamente a la Administración. La diferencia entre el art.3.1 y el art.3.2 es que el art.3.2 incluye expresamente la posibilidad de establecer excepciones a dicho artículo. La razón por la que se incluye esta posibilidad es que es concebible que se produzcan situaciones en las que, dada la naturaleza de las disposiciones adoptadas y la confidencialidad de los datos recogidos, sea razonable y permisible que el contratista utilice los datos recogidos también para sus propios fines¹³⁹.

a NOMBRE DE LA ORGANIZACIÓN. El contratista no tendrá derecho a utilizar dichos datos para ningún otro fin que no sea la ejecución del Contrato. A petición de NOMBRE DE LA ORGANIZACIÓN, el contratista destruirá dichos datos y/o los entregará a NOMBRE DE LA ORGANIZACIÓN.

3.2 Todos los derechos relativos a los datos que se creen o recojan en el curso de la ejecución del Contrato corresponderán a NOMBRE DE LA ORGANIZACIÓN. A menos que las Partes acuerden lo contrario, el contratista no tendrá derecho a utilizar dichos datos para ningún otro fin que no sea la ejecución del Contrato. A petición de NOMBRE DE LA ORGANIZACIÓN, el contratista destruirá dichos datos y/o los entregará a NOMBRE DE LA ORGANIZACIÓN.

3.3 Los datos mencionados en los artículos 3.1 y 3.2 se entregarán en un formato de archivo común que designará NOMBRE DE LA ORGANIZACIÓN. Si, para convertir los datos al formato de archivo deseado por NOMBRE DE LA ORGANIZACIÓN, el contratista debe realizar un trabajo adicional, NOMBRE DE LA ORGANIZACIÓN pagará al contratista una tarifa razonable al respecto. En ningún caso una disputa sobre el importe de los honorarios a pagar por NOMBRE DE LA ORGANIZACIÓN al contratista podrá dar lugar a la suspensión por parte del contratista de sus obligaciones en virtud de estos términos y condiciones.

3.4 A petición de NOMBRE DE LA ORGANIZACIÓN, el contratista proporcionará pruebas de la destrucción de los datos.

¹³⁸ Por lo demás, el artículo 3 no establece disposiciones diferentes sobre la confidencialidad, ya que se supone que las condiciones genéricas de confidencialidad se habrán incluido en otra parte del acuerdo o en las condiciones generales de las que forman parte las cláusulas tipo. Si no es así, sería aconsejable seguir haciéndolo.

¹³⁹ Como ya hemos señalado, este artículo se ha incorporado en el art. 16 del borrador de las cláusulas tipo de la Comisión europea, si bien con carácter opcional.

El borrador de cláusulas tipo auspiciado por la Comisión Europea no contempla de modo exhaustivo las cuestiones relativas a la normativa sobre protección de datos personales, ya que estas se ocupan exclusivamente de las cuestiones que cubre el proyecto de RIA. No obstante, si incorpora, con carácter voluntario, una cláusula sobre propiedad de los datos¹⁴⁰ que es idéntica a la cláusula que con ese contenido se contiene en las cláusulas tipo de la Ciudad de Ámsterdam. Y que por lo demás, resulta conveniente introducir en los pliegos relativos a la contratación de soluciones de inteligencia artificial dadas las dificultades y dudas que sobre este aspecto pueden surgir tras la finalización del contrato¹⁴¹.

4.- CONCLUSIONES

Como se indicó al principio de este trabajo, la contratación pública puede desempeñar un trascendente papel en la incorporación de soluciones de IA desde múltiples perspectivas.

En primer lugar, la contratación pública puede servir de estímulo al despliegue de esta tecnología, si se emplea como elemento catalizador de la innovación. En la Estrategia Nacional de IA se vislumbra esta función, que deberá dotarse de concreto contenido y de la consecuente financiación. Cuando este sea uno de los elementos primordiales de la licitación, sobre esta función deberán pivotar sus aspectos decisivos, como el uso de un determinado procedimiento de licitación o la regulación de los derechos de propiedad intelectual.

En segundo lugar, y al hilo de lo anterior, no debe olvidarse que uno de los principios básicos de la contratación pública, el de garantía de competencia, no se repliega en los supuestos de contratación de tecnología. Por ello, al tratarse de un mercado estratégico todavía en evolución, la contratación pública tiene gran potencial de orientar a este mercado hacia las prácticas más satisfactorias desde el punto de vista del uso ético de la IA.

¹⁴⁰ Vid. cláusula 16.

¹⁴¹ Sobre este tipo de cuestiones ha tenido que pronunciarse la AEPD en las Resoluciones 1167/2011 de 24 de mayo de 2011 y 857/2007 de 13 de septiembre de 2007.

En tercer lugar, en ausencia de una regulación específica del uso de la IA en las Administraciones públicas, actualmente el procedimiento de licitación permite articular un cauce en cierto modo adecuado para plasmar algunas de garantías ínsitas en nuestro Ordenamiento. Así, entre otros aspectos, con carácter previo a la contratación deberán evaluarse riesgos y alternativas, además de justificarse suficientemente la necesidad de la contratación. Es posible además facilitar canales de participación antes de la contratación de la solución. Y es preciso prever ya las obligaciones de transparencia y explicabilidad de los algoritmos empleados. Además, la licitación formal permite que distintas decisiones de relevancia puedan ser objeto de impugnación. No obstante, todo ello no es óbice para que sea deseable que, como reclama la doctrina, se incorpore en nuestra legislación un procedimiento específico de aprobación del uso de algoritmos¹⁴², al margen del procedimiento de contratación. Porque en otro caso, como señala SÁNCHEZ GRAELLS, se corre el riesgo de que sean los órganos de contratación los que actúen de guardianes del uso fiable de la IA, cuando pueden carecer de la especialización y experiencia suficiente, adoptándose además decisiones que no cuenten con el necesario nivel de estandarización¹⁴³. Por tanto, lo aconsejable será que se articule un procedimiento de aprobación de algoritmos que esté coordinado suficientemente con el procedimiento de licitación o, que al menos, se articulen los instrumentos de gobernanza necesarios que permitan garantizar el uso adecuado de la discrecionalidad técnica en la adopción de distintas decisiones. En otro caso, por lo demás, no todas las Administraciones tendrán posibilidad de incorporar soluciones de IA con las garantías necesarias.

Por lo demás, y en ausencia de mecanismos como los que acabamos de describir, es difícil que en los supuestos de contratación de soluciones de IA que causen mayor riesgo pueda actualmente prescindirse de las garantías que respecto al contenido de los pliegos se establecen en los documentos internacionales sobre cláusulas tipo que

¹⁴² Vid., entre otros, CERRILLO I MARTÍNEZ, A. (2019): “Robots, asistentes virtuales..”, *op. cit.*, p.291 e MARTÍN DELGADO, I. (2022): “La aplicación del principio...”, *op. cit.*.

¹⁴³ SÁNCHEZ GRAELLS, A. (2022): “Interesting legislative proposal to make procurement of AI conditional on external checks”, en www.howtocrackanut.com

hemos tenido ocasión de analizar. Por un lado, porque la mayor parte de estas cláusulas derivan sin dificultad de la legislación general aplicable al uso de las nuevas tecnologías por parte de la Administración. Por otro, porque eludir las garantías que contienen estas cláusulas tipo puede comprometer la utilización futura del algoritmo cuando la regulación europea sobre IA sea una realidad.

BIBLIOGRAFÍA CITADA

- ALHOLA, K. / SALO, A. / ANTIKAINEN, R. / BERG, A. (2017): “Promoting Public Procurement of Sustainable Innovations: Approaches for Effective Market Dialogue”, en *Global Public Procurement Theories and Practices*, Thai, Khi V. (Ed.), Springer.
- BOIX PALOP A. (2020): “Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la Administración para la adopción de decisiones”, *Revista de Derecho Público: Teoría y Método*, Vol. I, pp. 223-270.
- CANALS, D. (2019): “El proceso normativo ante el avance tecnológico y de la transformación digital (inteligencia artificial, redes sociales y datos masivos)”, en *Revista General de Derecho Administrativo*, nº 50.
- CERQUEIRA GOMES, P. (2014): “The innovative Innovation Partnerships under the new EU directives (2014)”, *Public Procurement Law Review*, 23(4).
- CERRILLO I MARTÍNEZ, A. (2019): “Com obrir les caixes negres de les administracions públiques? Transparència i rendició de comptes en l’ús dels algorismes”. *Revista Catalana de Dret Públic*, núm. 58.
- CERRILLO I MARTÍNEZ, A. (2019): “Robots, asistentes virtuales y automatización de las administraciones públicas”, *Revista Galega de administración pública*, 61, pp.279-309.
- DE GUERRERO MANSO, C. (2018): “Las consultas preliminares del mercado: una herramienta para mejorar la eficiencia en la contratación” en GIMENO FELIU, J.M. (dir.): *Estudio sistemático de la ley de contratos del sector público*, Aranzadi, pp.1047-1071.
- GALLEGO CÓRCOLES, I. (2015): “La modificación de los contratos en la cuarta generación de Directivas sobre contratación pública” en GIMENO FELIU, J.M. et alt., *Las nuevas directivas de contratación pública: (ponencias sectoriales X Congreso Asociación Española Profesores de Derecho Administrativo)*, Aranzadi, Cizur Menor.

- GALLEGO CÓRCOLES, I. (2019): “Las consultas preliminares del mercado: algunas reflexiones”, en GIMENO FELIU, J.M. (dir.): *Observatorio de los contratos públicos 2018*, Aranzadi, pp. 131-180.
- GALLEGO CÓRCOLES, I. (2021): “Transferencia de competencias y fórmulas de cooperación horizontal en la jurisprudencia del Tribunal de Justicia del año 2020”, en GIMENO FELIU, J.M.: *Observatorio de los contratos públicos 2020*, Aranzadi, pp. 449-477.
- GAMERO CASADO, E. (2021): “Compliance (o cumplimiento normativo) de desarrollos de Inteligencia Artificial para la toma de decisiones administrativas”, en *Diario La Ley*, nº 50, 19 de abril de 2021.
- GAMERO CASADO, E. (2022): “Supervisión, auditoría y control jurídico en la contratación pública de soluciones de robotización e inteligencia artificial para soporte a la toma de decisiones”, *www.obcp.es*.
- GIMENO FELIU, J.M. (2014): *El nuevo paquete legislativo comunitario sobre contratación pública. El contrato público como herramienta del liderazgo institucional de los poderes públicos*, Aranzadi, Cizur Menor.
- GIMENO FELIU, J.M. (2018): “La contratación pública como estrategia: el fomento de la innovación”, *Revista Ius Publicum*, 1.
- GIMENO FELIU, J.M. (2019): *La Ley de contratos del Sector Público 9/2017. Sus principales novedades, los problemas interpretativos y posibles soluciones*, Thomson-Reuters.
- GIMENO FELIU, J.M./ GARCÍA-ÁLVAREZ GARCÍA, G./TEJEDOR BIEL-SA, J. (2022): *La compra pública de tecnología en la nube en España*, Aranzadi, Cizur Menor.
- GÓMEZ JIMÉNEZ, M.L. (2021): *Automatización procedimental y sesgo electrónico: el procedimiento administrativo electrónico desde la Inteligencia Artificial*, Thomson-Reuters Aranzadi, Cizur Menor, p. 150.
- GRUPO INDEPENDIENTE DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL (2019): *Policy and Investment Recommendations for Trustworthy AI*, p.20.
- GUTIÉRREZ DAVID, M.E. (2021): “Administraciones inteligentes y acceso al código fuente de los algoritmos públicos. Conjuro de cajas negras decisionales”, *Derecom*, nº 30, pp. 143-229.
- HERNANDO RYDINGS, M. (2018): “Procedimiento de licitación con negociación y diálogo competitivo”, en GAMERO CASADO, E./GALLEGO CÓRCOLES, I., *Tratado de Contratos del Sector Público*, t.2, Tirant lo Blanch, Valencia, pp.1629-1668.
- HORGUÉ BAENA, C.(2018): “Modificación de los contratos” en GAMERO CASADO, E./GALLEGO CÓRCOLES, I., *Tratado de Contratos del Sector Público*, t.3, Tirant lo Blanch, Valencia, pp. 2259-2312.

- HUERGO LORA, A. (2020): “Una aproximación a los algoritmos desde el Derecho Administrativo”, en HUERGO LORA, A. (Dir.): *La regulación de los algoritmos*, Thomson-Aranzadi, Madrid, pp. 23-87.
- MARTÍN DELGADO, I. (2020): “Innovación tecnológica e innovación administrativa en la contratación pública”, en Isaac Martín Delgado y Jose Antonio Moreno (dir.), *Administración electrónica, transparencia y contratación pública*, Iustel, pp. 42.
- MARTÍN DELGADO, I. (2022): “La aplicación del principio de transparencia a la actividad administrativa algorítmica”, en esta misma obra.
- McBRIDE, K./ VAN NOORDT, C./ MISURACA, G./ HAMMERSCHMID, G. (2021): “Towards a Systematic Understanding on the Challenges of Procuring Artificial Intelligence in the Public Sector”.
- MIÑO LÓPEZ, A. (2017): “Preliminary market consultations in innovation procurement: a principled approach and incentives for anticompetitive behaviours”, *Ius publicum*,
- MIRANZO DÍAZ, J. (2020): “Inteligencia artificial y contratación pública” en MARTÍN DELGADO, I. / MORENO MOLINA, J.A. (Dirs.): *Administración electrónica, transparencia y contratación pública*, Iustel, Madrid, pp.105-142.
- MIRANZO DÍAZ, J. (2021): “Apuntes sobre la contratación de sistemas de software”, *www.obcp.es*.
- MIRANZO DÍAZ, J. (2022): “El ‘modelo’ de plataformas de contratación pública en España: la autonomía administrativa y el derecho de la competencia”, *Revista Catalana de Dret Públic*, 64, pp. 140-157.
- MORENO MOLINA J.A. (2018): *Hacia una compra pública responsable y sostenible: novedades principales de la Ley de contratos del Sector Público 9/2017*, Tirant lo Blanch, Valencia.
- OLUKA NAGITTA, P./ ADOKO OBICCI, P./ AWOUR, E. (2022): “Human-centered artificial intelligence for the public sector: The gate keeping role of the public procurement professional”, *Procedia Computer Science*, n.º 200(1).
- PONCE SOLÉ, J. (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico”, *Revista General de Derecho Administrativo*, n.º 50.
- RAZQUIN LIZARRGA M.M. (2018): “El principio de confidencialidad en la contratación pública”, en GIMENO FELIU, J.M. (dir.): *Estudio sistemático de la ley de contratos del sector público*, Aranzadi, pp.1047-1071.
- SÁNCHEZ GRAELLS, A. (2022): “Interesting legislative proposal to make procurement of AI conditional on external cheks”, en *www.howtocrackanut.com*.

- SÁNCHEZ GRAELLS, A. (2022): “Protecting procurement’s AI gatekeeping role in domestic law, and trade agreements?” en www.howtocrackanut.com
- SÁNCHEZ GRAELLS, A. (2022): “Public procurement and [AI] source code transparency, a (downstream) competition issue (RE C-796/18), en www.howtocrackanut.com.
- VALCÁRCEL FERNÁNDEZ, P. (2018): “Definición del objeto contractual en la compra pública de innovación”, en FERNÁNDEZ SALMERÓN, M./MARTÍNEZ GUTIERREZ, R.: *Transparencia, innovación y buen gobierno en la contratación pública*, Tirant lo Blanch.
- VALCÁRCEL FERNÁNDEZ, P. (2018): “La compra pública de innovación (CPI): el procedimiento de adjudicación de asociación para la innovación”, en GAMERO CASADO, E./GALLEGO CÓRCOLES, I., *Tratado de Contratos del Sector Público*, t.2, Tirant lo Blanch, Valencia.
- VALCÁRCEL FERNÁNDEZ, P. (2019): “Tres dimensiones de la transparencia en la contratación pública: rendición de cuentas, respeto de los derechos de operadores económicos y mejora global de la gestión de este sector a través del “big data”, *Observatorio de los contratos públicos 2018* (coord. por Carmen de Guerrero Manso; José María Gimeno Feliú (dir.), Aranzadi Thomson Reuters, Cizur Menor, pp. 93-130.
- VALCÁRCEL FERNÁNDEZ, P. (2018): “La consultas preliminares del mercado como mecanismo para favorecer las “compras inteligentes”, *Revista Española de Derecho Administrativo*, núm. 191 [consultada versión electrónica].
- VALCÁRCEL FERNÁNDEZ, P. (2019): “The Relevance of Promoting Collaborative and Joint Cross Border Procurement for Buying Innovative Solutions”; en *Joint public procurement and innovation: lessons accross borders* (Gabriella M. Racca (ed. lit.), Christopher R. Yukins (ed. lit.), 2019, Bruylant, Bruselas, pp. 133-170.
- VALERO TORRIJOS, J. (2019): “Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración”, *Revista Catalana de Dret Public*, nº 58.
- VALERO TORRIJOS, J. (2021): “Inteligencia artificial y contratación pública”, *Observatorio de Contratación Pública*, www.obcp.es.
- VESTRI, G. (2021): “La inteligencia artificial ante el desafío de la transparencia algorítmica. Una aproximación desde la perspectiva jurídico-administrativa”, *Revista Aragonesa de Administración Pública*, nº 56, pp. 368-398.
- ZLOTNIK, A. (2019): “Inteligencia Artificial en las Administraciones Públicas: definiciones, evaluación de viabilidad de proyectos y áreas de aplicación”, *Boletic*, n.o 84, pp.24-32.

Capítulo XII

Empleo público e inteligencia artificial

RAMON GALINDO CALDÉS¹

*Profesor Agregado de Derecho Administrativo
Universitat Oberta de Catalunya*

Resumen: La creciente implantación de la automatización, los sistemas basados en datos e incluso de la IA plantea algunas dificultades en un empleo público basado en dinámicas de otros tiempos. Se trata de un proceso que avanza a diferentes velocidades y que impacta en el empleo público de diferentes maneras. La necesidad de nuevos tipos de empleados converge con una estructura de empleo público que afronta otros problemas, como el relevo generacional, la temporalidad o la falta de adaptación a las nuevas necesidades de la actividad pública. En este trabajo reflexionamos sobre algunas cuestiones claves que deben abordarse desde un punto de vista pragmático. La transformación del empleo público debe afrontarse teniendo en cuenta los condicionantes actuales, pero dando pasos para avanzar en una Administración cada vez más automatizada y en la que los datos estarán en el centro de su actividad. Hay por tanto que reflexionar sobre cómo transformar las plantillas actuales, pero también sobre los nuevos empleados públicos que deben incorporarse progresivamente a la Administración para que pueda cumplir sus fines.

Abstract: The growing implementation of automation, data-based systems and even AI poses some difficulties in public employment based on the dynamics of past times. It is a process that advances at different speeds and that impacts public employment in different ways. The need for new types of employees converges with a public employment structure that faces other problems, such as generational change, temporality or the lack of adaptation to the new needs of public activity. In this paper we reflect on some key issues that must be addressed from a pragmatic point of view. The transformation of public employment must be faced taking into account the current conditions, but taking steps to advance in an Administration that is increasingly automated and in which data will be at the centre of its activity. It is therefore necessary to reflect on how to transform the current staffs, but also on the new public employees who must gradually join the Administration so that it can fulfil its purposes.

Palabras-clave: empleo público, administración digital, inteligencia artificial

Key-words: Public employment, Digital Administration, Artificial intelligence

SUMARIO.- 1. TRES GRADOS, DIFERENTES AFECTACIONES.- 1.1. El impacto de la automatización en el empleo público: destrucción y redefinición de puestos y tareas.- 1.2. El uso de datos en la toma de decisiones.- 1.3 La sustitución del empleado en la toma de decisiones.- 1.3.1. La sustitución de las decisiones por un sistema de IA.- 1.3.2.

¹ ORCID: <https://orcid.org/0000-0003-2684-1425>

¿Hay una reserva de intervención de los empleados públicos?- 2. ¿HACIA UNA NUEVA SEGMENTACIÓN DE LOS EMPLEADOS PÚBLICOS?- 2.1.- Un impacto a diferentes niveles.- 2.1.1 Redefinición de puestos y tareas.- 2.1.2 Cambios necesarios en los empleados actuales.- 2.1.3 Nuevos empleados públicos.- 2.2.- La contratación pública: ¿personal especializado o recurso al sector privado?- 2.3.- ¿Dos clases de empleados? Si no es posible cambiar todo el empleo público.... ¿Lo hacemos con una parte?- 3. CUATRO PRESUPUESTOS NECESARIOS PARA LA ADAPTACIÓN DEL EMPLEO PÚBLICO.- 3.1.- Planificación.- 3.2.- Selección.- 3.3.- Transformación.- 3.4.- Los datos en la gestión de los recursos humanos.

1.- TRES GRADOS, DIFERENTES AFECTACIONES

A menudo se engloba bajo la “etiqueta” de IA cualquier sistema de procesamiento de datos basado en algoritmos. Cuando analizamos como estos sistemas impactan en el empleo público es necesario distinguir entre los sistemas que automatizan tareas, los que sirven de apoyo a la toma de decisiones, y los que pueden tomar decisiones de forma autónoma. Lo hacemos a continuación, con todas las prevenciones, y a los efectos de diferenciar los distintos efectos que el uso de los algoritmos y los datos en el sector público produce en el empleo público.

1.1.- El impacto de la automatización en el empleo público: destrucción y redefinición de puestos y tareas

La automatización supone la sustitución de tareas operadas por humanos a través de la tecnología, generalmente con el uso de algoritmos. La LAESP la define como la “actuación administrativa producida por un sistema de información adecuadamente programado sin necesidad de intervención de una persona física en cada caso singular. Incluye la producción de actos de trámite o resolutorios de procedimientos, así como de meros actos de comunicación”. La Ley General Tributaria (LGT) preveía ya hace casi dos décadas las notificaciones telemáticas y la actuación administrativa automatizada (art. 96 LGT), y de forma general la LAESP preveía sistemas de firma electrónica para la actuación administrativa automatizada (arts. 13.2 b) y 18 LAESP) o la automatización de registros electrónicos (arts. 24.4 y 25 LAESP). Pero también de forma explícita la posibilidad de

adoptar y notificar resoluciones administrativas automatizadas (art. 38.2 LAESP), indicando el órgano competente (art. 39 LAESP). Simplemente destacamos que este tipo de actuación no es nueva, aunque apenas se había desarrollado hasta los últimos años. La actual regulación sobre régimen jurídico también contempla la actuación administrativa automatizada (art. 41.1 LRJSP), y recientemente la normativa de desarrollo sobre Administración electrónica ha hecho lo propio, con pocas novedades (art. 13 RAFME).

Se pueden automatizar todo tipo de tareas repetitivas o con un margen de decisión estrecho. La automatización es un hecho en la producción industrial a través de la robotización o en actividades terciarias con la automatización de procesos. No es un hecho extraño en la actividad pública, ya que de forma creciente se han automatizado procesos o incluso se prestan servicios públicos operados por sistemas robotizados. En la Automatización Robótica de Procesos (RPA) se sustituyen tareas utilizando reglas predefinidas, con lo que se reduce drásticamente el tiempo de procesamiento respecto al trabajo manual. Su uso se está extendiendo de forma sostenida, por razones de eficacia y eficiencia bastante evidentes.

Esta automatización puede afectar a procedimientos –aunque no solo, ya que no toda actuación administrativa se enmarca en un procedimiento administrativo-. Se da en actuaciones claramente predefinidas y en el caso de procedimientos, absolutamente reglados. La casuística es enorme y solo destacaremos una distinción, entre actos de comunicación o de trámite, y actos cualificados y resolutorios. En los primeros se sustituye la intervención de un empleado público por un algoritmo. Es una situación que ha ido generalizándose y la vemos con naturalidad en procedimientos informativos (por ejemplo la expedición automatizada de certificados de todo tipo), trámites (por ejemplo, fiscales o laborales) o acuses de recibo de solicitudes (por ejemplo, matrículas escolares o solicitudes de ayudas y subvenciones). En los segundos ya se produce una “decisión” de la Administración, tanto a través de actos de trámite cualificados (como el inicio de un procedimiento sancionador) o de resoluciones administrativas, y sus correspondientes notificaciones. En ambos casos, recordemos, se trata de procedimientos reglados y sin margen de discrecionalidad.

Un ejemplo claro en este sentido son las actas de infracción automatizadas de la Inspección de Trabajo y Seguridad Social (que se introducen a través de una modificación del art. 53.1.a) del Real Decreto Legislativo 5/2000, de 4 de agosto, por el que se aprueba el texto refundido de la Ley sobre Infracciones y Sanciones en el Orden Social (LISOS)) o la resolución y notificación automatizadas en prestaciones del sistema de la Seguridad Social, procedimientos de afiliación, cotización y recaudación (nuevo art. 130 del Real Decreto Legislativo 8/2015, de 30 de octubre, por el que se aprueba el texto refundido de la Ley General de la Seguridad Social), modificaciones legislativas ambas introducidas en 2021. Se automatizan en estos casos procedimientos reglados en los que la intervención de los funcionarios se ha producido en la determinación de los supuestos en los que se va a aplicar dicha automatización². Existe intervención humana, pero se desplaza a la determinación de los supuestos, que necesariamente deberán excluir cualquier resolución en la que pueda existir el más mínimo margen de discrecionalidad.

En todos los casos apuntados sustituimos al empleado público por un algoritmo. O siendo más precisos, lo hacemos con una o más tareas que corresponden a un empleado. En un grado de menor afectación el empleado deja de desempeñar una de sus tareas –por ejemplo, la notificación en un procedimiento concreto- y en un grado máximo la automatización “vacía” las tareas correspondientes a un puesto de trabajo. La automatización también puede afectar a diversas tareas que pueden llevar a cabo “robots software” o conversacionales³, por ejemplo *chatbots* de atención al ciudadano. No hay duda de los bene-

² En este sentido, se entiende por “actividad automatizada, cualquier actuación realizada íntegramente a través de medios electrónicos por la Inspección de Trabajo y Seguridad Social en la que la intervención del personal con funciones inspectoras se produzca de forma indirecta.” (art. 43.1 del Real Decreto 928/1998, de 14 de mayo, por el que se aprueba el Reglamento general sobre procedimientos para la imposición de sanciones por infracciones de orden social y para los expedientes liquidatorios de cuotas de la Seguridad Social. (RISOS)). Esta intervención se concreta con la determinación de los supuestos por parte del Director del Organismo Estatal Inspección de Trabajo y Seguridad Social (art. 43.2 RISOS).

³ CERRILLO i MARTÍNEZ, A. (2021): “Robots, asistentes virtuales y automatización de las administraciones públicas”, *Revista Galega de Administración Pública* (REGAP), nº 61

ficios de automatizar, y también de que deben reforzarse las garantías de los ciudadanos para evitar posibles errores o sesgos que pueden generarse de inicio en el diseño de los algoritmos. Sin embargo, nos centramos aquí en observar cómo impacta dicha automatización sobre el empleo público, tanto colectiva, organizativa como individualmente.

1.2.- *El uso de datos en la toma de decisiones*

A menudo se etiquetan como IA sistemas que no lo son en sentido estricto, sino sistemas de apoyo a la toma de decisiones basados en datos. (sistemas de *business intelligence*, cuadros de mando, *big data*, etc.). En este caso la decisión sigue estando en manos de las personas responsables, pero disponer de la información que ofrecen estos sistemas puede producir una decisión más fundamentada. En definitiva, el titular de un órgano que debe resolver un procedimiento -o el instructor de un procedimiento-, dispone de mayor información para fundamentar su decisión. Incluso el sistema puede sugerir soluciones o de forma predictiva (a través del *big data*) pueden detectarse patrones que generen posteriormente actuaciones públicas, por ejemplo, en lo que se ha denominado como “policía predictiva”⁴ o en la clasificación de empresas según el riesgo de incumplimiento de la legislación laboral⁵.

La cuestión es si el uso de dichos sistemas produce mejores decisiones. La respuesta no puede ser otra que en principio sí, aunque dependerá del factor humano, de si quién decide sabe interpretar los datos o, en general, si sabe utilizar el sistema. La generalización del uso de los datos en las decisiones públicas y de la *datificación* de la actividad administrativa en general exige disponer de datos actualizados y de calidad. También implementar un sistema de gobernanza de datos, que garantice que todos los datos de una Administración son compartidos y se evitan *silos* de información. También comporta nue-

⁴ DÍAZ GONZÁLEZ, G. M. (2020): “Algoritmos y actuación policial: la policía predictiva”, en HUERGO LORA, A. (ed.): *La regulación de los algoritmos*, Cizur Menor, Aranzadi-Thomson-Reuters.

⁵ TODOLÍ SIGNES, A. (2020): “Retos legales del uso del big data en la selección de sujetos a investigar por la Inspección de Trabajo y de la Seguridad Social”, *Revista Galega de Administración Pública (REGAP)*, nº 59, pp. 79-102

vos problemas, como son la opacidad, protección de datos, seguridad, posibles sesgos, etc⁶.

Nos interesa en este punto ver en qué medida afecta al empleo público. Dos de los problemas que nos genera este contexto parecen claros. En primer lugar, la implementación y uso de dichos sistemas necesita de perfiles profesionales de los que en general no dispone la Administración. Por otro lado, existen importantes dificultades para reclutar nuevos efectivos que sí respondan a dichos perfiles especializados, ya que el sector público no es una prioridad para ellos.

1.3.- *La sustitución del empleado en la toma de decisiones*

En un tercer grado encontramos lo que propiamente podemos denominar sistemas de IA, en los que efectivamente se sustituyen las decisiones *humanas* por otras que el sistema toma de forma autónoma, concretamente si existe una reserva de intervención de los empleados públicos, o como se ha denominado, una *reserva de humanidad*⁷.

1.3.1- *La sustitución de las decisiones por un sistema de IA*

La definición de IA es muy ambigua y depende de diferentes autores y disciplinas. De forma genérica se ha definido como “la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano”⁸. La Comisión Europea la ha definido como aquella que “se aplica a los sistemas que manifiestan un comportamiento inteligente, pues son capaces de analizar su entorno y pasar a la acción –con cierto grado de autonomía– con el fin de alcanzar obje-

⁶ CERRILLO i MARTÍNEZ, Agustí (2019): “El impacto de la inteligencia artificial en el derecho administrativo ¿nuevos conceptos para nuevas realidades técnicas?”, *Revista General de Derecho Administrativo*, nº 50

⁷ PONCE SOLÉ, J. (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico”. *Revista General de Derecho Administrativo*, nº 50.

⁸ ROUHIAINEN, L. (2018): *Inteligencia artificial. 101 cosas que debes saber hoy sobre nuestro futuro*, Planeta, Barcelona, p.17

tivos específicos”⁹. Entre ellos incluye tanto programas informáticos (asistentes de voz, programas de análisis de imágenes, motores de búsqueda, sistemas de reconocimiento facial y de voz) como dispositivos de hardware (robots, automóviles autónomos, drones o aplicaciones del internet de las cosas (IoT)). Su materialización puede ser muy diversa, aunque a grandes rasgos implica técnicas que intentan simular el comportamiento humano (analítica predictiva, lenguaje natural, aprendizaje automatizado (*machine learning*), redes neuronales o *deep learning*). Su concreción, desde un punto de vista jurídico, también es compleja¹⁰.

Teniendo en cuenta esta concepción de la IA como sistema autónomo, su implantación no está aún demasiado extendida y existe distancia entre el potencial del *big data* y de la IA y su implantación en la Administración¹¹. Si nos ceñimos a su impacto sobre el empleo público y sobre las funciones que desempeñan los empleados públicos deberíamos tener en cuenta dos planos. El primero afecta a las tareas que pueden desempeñar estos sistemas. En este sentido nos remitimos a lo ya expuesto sobre la automatización y la supresión de tareas y puestos. El segundo caso es más problemático, en tanto que implica la toma de decisiones de forma autónoma por sistemas de IA, lo que puede afectar no solo al ejercicio de potestades públicas, sino también a la toma de decisiones con un margen de discrecionalidad.

1.3.2.- ¿Hay una reserva de intervención de los empleados públicos?

Como norma, las potestades administrativas deben ser ejercidas por funcionarios de acuerdo con el art. 9.2 TREBEP, cuestión ésta controvertida, que plantea problemas a la luz de la extensión de la laboralización en el sector público, y produce algunas situaciones in-

⁹ COMISIÓN EUROPEA (2018): *Comunicación de la Comisión al Parlamento Europeo, al Consejo Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones “Inteligencia artificial para Europa”*. COM(2018) 237 final

¹⁰ SCHUETT, J. (2019): “Defining the scope of AI regulations”, *Computers and Society*, arXiv:1909.01095 [cs.CY] <https://doi.org/10.48550/arXiv.1909.01095>

¹¹ DESOUZA, K. C.; JACOB, B. (2017): “Big Data in the Public Sector: Lessons for Practitioners and Scholars”, *Administration & Society*, nº 49(7): 18

cluso paradójicas, como el hecho de que algunas potestades puedan ser externalizadas, desapareciendo la garantía de su ejercicio por un funcionario¹². Aun así se considera que el ejercicio de potestades administrativas en las que existe un margen de discrecionalidad debe corresponder a un funcionario público, aunque hay autores que defienden cierto margen para la utilización de algoritmos¹³ o la posibilidad de su uso en el caso de la discrecionalidad técnica¹⁴. Un ejemplo en el ámbito laboral es la posible utilización de algoritmos en materia de empleo o seguridad social –que en muchos casos no implica discrecionalidad–, pero su aplicación puede ser problemática en otras materias como prevención de riesgos laborales, donde hay un margen de interpretación a la hora de determinar una actuación administrativa.

Lo que es más probable, dentro de nuestro ordenamiento, es que estos sistemas de IA puedan desarrollar tareas *cuasi decisoriales* –una suerte de *asesores robotizados* (*robo-advisors*)¹⁵– pero que la decisión corresponda en todo caso a un funcionario (o empleado) público. Desde el punto de vista del ejercicio de potestades públicas estos sistemas pueden desarrollar actividades técnicas o auxiliares¹⁶, aunque deslindar éstas de las actividades en las que se está produciendo una decisión pública es una tarea ciertamente compleja.

¹² GAMERO CASADO, E. (2021): “El ejercicio de potestades administrativas en el empleo público”, en GAMERO CASADO, E. (dir.): *La Potestad Administrativa: concepto y alcance práctico de un criterio clave para la aplicación del Derecho administrativo*. Valencia: Tirant lo Blanch.

RAMÍO, C. (2019): “El impacto de la inteligencia artificial y de la robótica en el empleo público”, *GIGAPP Estudios Working Papers*, nº 5(98-110), p. 487-488

¹³ HUERGO LORA, A. (2020): “Una aproximación a los algoritmos desde el derecho administrativo”, en HUERGO LORA, A. (ed.): *La regulación de los algoritmos*, Cizur Menor, Aranzadi-Thomson-Reuters.

¹⁴ MARTÍN DELGADO, I. (2009): “Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada”, *Revista de administración pública*, nº 180, p. 353-386

¹⁵ ENGIN, Z.; TRELEAVEN, P. (2018). Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies. *The Computer Journal*, xy082.

¹⁶ BERNING PRIETO, A. (2021): “Delimitación entre potestades administrativas y actividades técnicas o auxiliares”, en GAMERO CASADO, E. (dir.): *La Potestad Administrativa: concepto y alcance práctico de un criterio clave para la aplicación del Derecho administrativo*. Valencia: Tirant lo Blanch.

Habría que reflexionar en todo caso sobre qué margen de decisión cabe si el sistema de IA ofrece una solución teóricamente *óptima*, debiendo motivar la separación del criterio propuesto. También el posible cuestionamiento por los interesados de una decisión *humana* contraria a la decisión supuestamente racional y objetiva que propone el sistema. En este sentido será necesario que los empleados públicos que utilicen estos sistemas como fundamentación de sus decisiones conozcan el funcionamiento de los algoritmos y sepan reconocer cuando éstos no están funcionando correctamente o están actuando de forma sesgada:

“Public organisations might start considering general and elementary knowledge on AI as a basic knowledge for all their employees, or at least for all the employees that deal with a system that uses AI techniques. Algorithms are biased and this might lead to wrong suggestions – if not decisions – that require correcting when they occur. Hence public employees should have a general understanding of this concept to be able to act properly in a new routine based on the information given by an algorithm. They should be able to challenge the system, understand in which situation the system does not work properly, and be able to act accordingly.”¹⁷

En un sentido similar Cerrillo propone la supervisión del uso de la IA en el ejercicio de las potestades discrecionales y, entre otros elementos afirma sobre su uso que

La supervisión humana puede realizarse por las personas que adoptan decisiones sobre la base de la propuesta elaborada por un algoritmo. Estaríamos ante un modelo de decisión con intervención humana que consistiría en que las Administraciones públicas utilizan la inteligencia artificial como apoyo o ayuda a los decisores públicos. En estos casos es necesario asegurar que la persona dispone de otros datos que puedan facilitarle elementos para cuestionar la propuesta realizada por el algoritmo. De no ser así, su intervención podría reducirse a ratificar la decisión propuesta. Es necesario que se garanticen las circunstancias que permitan a la persona no solo adoptar substancialmente la decisión sino también identificar los errores o los sesgos que pueda contener la decisión propuesta por el algoritmo¹⁸.

¹⁷ JOINT RESEARCH CENTRE (2022): *AI Watch European Landscape on the Use of Artificial Intelligence by the Public Sector*. JRC Science for Policy Report, p. 56

¹⁸ CERRILLO i MARTÍNEZ, A. (2020): “¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?”, *European Review of Digital*

A día de hoy, por tanto, estos sistemas de IA pueden desarrollar actividades regladas –y previamente definidas por *humanos*- o auxiliares –o de apoyo a la toma de decisiones. No se trata, sin embargo, de un tema cerrado, puesto que el ejercicio de potestades administrativas *solo* por los funcionarios, de acuerdo con el art. 9.2 TREBEP, plantea dificultades actualmente, situación que se tensionará más en el futuro próximo. Habría que considerar si esa concepción restrictiva debe evolucionar de forma que cualquier empleado público pueda ejercer dichas potestades ya que como hemos visto incluso sujetos privados, de forma indirecta, pueden hacerlo.

2.- ¿HACIA UNA NUEVA SEGMENTACIÓN DE LOS EMPLEADOS PÚBLICOS?

La dualización del régimen de los empleados públicos se ha consolidado en las últimas décadas, a pesar de los intentos de generar –especialmente a través del Estatuto Básico- de unas bases comunes para todos los empleados públicos. El impacto en el empleo público a causa de la irrupción tecnológica en la actividad pública crea una necesidad de nuevos perfiles, y afecta también a los empleados públicos existentes. Este desajuste entre necesidades y recursos existentes ha provocado el protagonismo generalizado de las empresas privadas, que suplen la incapacidad tecnológica de las Administraciones para diseñar y ejecutar proyectos de automatización o sistemas basados en datos. El contexto del empleo público actual –envejecimiento, procesos de estabilización- y la necesidad de adaptación tecnológica nos llevan a considerar un doble régimen como alternativa pragmática. A corto plazo no nos encontramos en una división entre *empleados públicos robots* y humanos¹⁹, sino entre *viejos* y *nuevos* empleados públicos.

Administration & Law – Erdal, Vol. 1 (1-2), p. 33

¹⁹ RAMÍO, C. (2019): “El impacto de la inteligencia artificial y de la robótica en el empleo público”, *op. cit.*, p. 416 y ss.

2.1.- *Un impacto a diferentes niveles*

La creciente datificación de la actividad pública tiene un efecto evidente en el empleo público: muchas de las tareas que hasta hoy han llevado a cabo los empleados públicos ahora pueden automatizarse, reducirse o transformarse por el uso de los datos y la tecnología. La primera cuestión –no jurídica- es si ello implica un redimensionamiento de las plantillas o una redefinición de los puestos de trabajo. Esta situación también implica una distinción entre los nuevos empleados y la necesaria adaptación de los existentes.

2.1.1.- Redefinición de puestos y tareas

Las estimaciones de destrucción de empleo a causa de la automatización y la IA difieren en su alcance, pero no en su inevitabilidad. Frey y Osborne, en su ya clásico estudio, hablan de un 47% de los empleos manuales o rutinarios en el mercado de trabajo estadounidense en riesgo de automatización²⁰. Otros autores ofrecen pronósticos sobre ámbitos concretos del sector público que son susceptibles de ser robotizados, como el de transportes, correos, personal administrativo, personal de Administración de justicia, etc²¹. No faltan tampoco visiones más optimistas, que defienden que se generarán nuevas ocupaciones relacionadas con la tecnología.

No hay duda que los factores que provocan la destrucción o transformación del mercado de trabajo también se dan en el sector público. La diferencia es que la relación de empleo de los trabajadores públicos no está en riesgo –en principio- toda vez que el TREBEP establece la inamovilidad en la condición de funcionario de carrera (art. 1.3 e) TREBEP), como un derecho inherente a su condición (art. 14 a) TREBEP). Inamovilidad que a la práctica se traslada también a los laborales indefinidos. El sistema de empleo público ha generado

²⁰ FREY, C. B.; OSBORNE, M. A. (2017): “The future of employment: how susceptible are jobs to computerisation?”, *Technological forecasting and social change*, 114.

²¹ RAMIÓ, C.; SALVADOR, M. (2018): *La Nueva Gestión del Empleo Público. Recursos Humanos e Innovación de la Administración*. Barcelona: Tibidabo Ediciones; RAMIÓ, C. (2019): “El impacto de la inteligencia artificial y de la robótica en el empleo público”, *op.cit.*, p. 402 y ss.

incluso que a la práctica un elevado número de temporales –funcionarios y laborales- lo sean más allá de lo admisible legalmente, como es bien sabido.

En este contexto hablar de destrucción del empleo público por la automatización de procesos es poco más que ilusorio. Si asumimos que no va a haber una reducción sustancial de los efectivos –más allá de jubilaciones-, habrá que pensar en que los empleados públicos lleven a cabo tareas diferentes. De hecho, el mayor impacto de la automatización en el empleo público no tiene que ver con la supresión de efectivos, sino con la supresión o modificación de puestos de trabajo. No necesariamente destrucción de empleo público, pero sí redefinición de tareas.

Las Administraciones deberían por tanto redefinir y redimensionar algunos puestos de trabajo. Ello debería afectar a todas las tareas automatizadas o que vayan a serlo en un futuro próximo, y en última instancia a la supresión de los puestos de trabajo asociados. Esta tarea debería llevarse a cabo a través de la modificación de la Relación de Puestos de Trabajo correspondiente –o instrumento equivalente-, cuestión sobre la que volveremos más adelante. Los empleados que ocupan las plazas afectadas por dichos puestos de trabajo deberían ser reorientados a otras tareas o puestos de trabajo, aportando más valor a la organización.

2.1.2.- Cambios necesarios en los empleados actuales

El sector público presenta algunos problemas estructurales que dificultan la adopción de soluciones basadas en los datos, como son la escasa formación, especialización y tecnificación de los empleados públicos. A estos problemas estructurales hemos de añadir otros que dificultan la transformación, como la cultura organizativa y la rigidez de la estructura burocrática.

La digitalización y automatización afecta a habilidades y tareas, por lo que será necesario un esfuerzo de adaptación, no solo en políticas de formación sino también de cambio organizativo. Tal y como se recuerda en el Plan Nacional de Competencias Digitales es fundamental el desarrollo de las competencias digitales de los empleados públicos, no solo en los puestos especializados sino en ge-

neral de todos ellos para que “[...] la Administración lleve a cabo una transformación interna (en procedimientos, relación con los administrados y cultura organizativa) [...]”²². Como se ha afirmado “la nueva identidad del empleado público trasciende las competencias de la administración pública tradicional para incorporar una visión global de la persona que realiza una función pública”²³. Se han puesto en marcha iniciativas que buscan definir las competencias y habilidades que debería tener el empleado público en un entorno digital, por ejemplo el proyecto de la Generalitat de Cataluña²⁴, que busca definir que competencias y habilidades debe tener el “empleado público digital” y ofrece instrumentos para su transformación.

2.1.3.- Nuevos empleados públicos

Como hemos visto, el impacto de los datos y la automatización deberían provocar una redefinición de puestos y tareas, y la adaptación de los empleados en activo. Pero son también necesarios nuevos efectivos, especializados y no necesariamente vinculados a la Administración con una relación estatutaria o laboral indefinida, sino condicionada a la ejecución de un proyecto o por un período determinado. Se trata de empleados públicos que responden a nuevos perfiles profesionales (STEM, analistas de datos, etc.). No nos referimos aquí a las competencias digitales que en mayor o menor medida afectan a todos los empleados públicos sino a los perfiles profesionales que deberían integrarse en la Administración.

²² GOBIERNO DE ESPAÑA (2021): *Plan Nacional de Competencias Digitales*, p. 34

²³ TRUJILLO SÁEZ, F.; ÁLVAREZ JIMÉNEZ, D. (2021): “Transformación digital de la administración pública: ¿Qué competencias necesitan los empleados públicos?”. *Gestión y Análisis de Políticas Públicas*, (27), p. 49–67.

²⁴ GENERALITAT DE CATALUNYA (2020): *Manual de transformación digital del empleado público*. Departament de Polítiques Digitals i Administració Pública. Accesible en [https://administraciodigital.gencat.cat/ca/teletreball/Guies-i-recursos/manuals/manual-de-transformacio-digital-de-lempleat-public/index.html#googtrans\(cales\)](https://administraciodigital.gencat.cat/ca/teletreball/Guies-i-recursos/manuals/manual-de-transformacio-digital-de-lempleat-public/index.html#googtrans(cales))

2.2.- *La contratación pública: ¿personal especializado o recurso al sector privado?*

El ejercicio por privados de funciones de autoridad existe desde hace tiempo²⁵, y responde a la incapacidad de la Administración de disponer internamente del conocimiento especializado en muchos ámbitos. Sujetos privados ejercen funciones de control, inspección o certificación, supliendo la falta de conocimiento experto en determinados ámbitos. Desde otro punto de vista, esta colaboración con particulares para ejercer funciones públicas o prestar servicios públicos puede producirse, entre otras formas, a través de las entidades colaboradoras de la Administración²⁶. Las razones para esta colaboración son la falta de medios de la Administración –materiales y personales-, la falta de conocimientos técnicos y la mejora de la calidad y de la eficacia en la gestión²⁷. Sin ánimo entrar aquí en la regulación de la cooperación público-privada, simplemente constatamos que, con sus dificultades y complejidad, el ejercicio de potestades administrativas por parte de sujetos privados está delimitada.

Cuestión diferente es el ejercicio de funciones técnicas o *auxiliares*, como hemos visto anteriormente. Estas tareas se desarrollan generalmente a través de la contratación pública. En lo que se refiere a la implantación de la Administración digital, -y especialmente en la automatización de procesos, soluciones tecnológicas basadas en datos o incluso sistemas de IA- las empresas privadas tienen un papel creciente en la Administración. Es muy habitual que empresas privadas colaboren con la Administración a través de contratos de consultoría o asistencia técnica. En este caso la colaboración es más opaca, en tanto que el ciudadano no sabe que los sistemas a través de los cuales se ejercen funciones públicas han sido diseñados y ejecutados por privados. Aquí surge una pregunta: ¿Si la Administración ha recurrido a grandes consultoras ante su falta de conocimiento experto, hasta qué

²⁵ CANALS AMETLLER, D.: *El ejercicio por particulares de funciones de autoridad. Control, inspección y certificación*. Granada: Comares, 2002.

²⁶ GALÁN GALÁN, A. (2019): *Entidades privadas colaboradoras de la Administración*. Valencia: Tirant lo Blanch.

²⁷ GALÁN GALÁN, A.; PRIETO ROMERO, C. (2009): “El ejercicio de funciones públicas por entidades privadas colaboradoras de la Administración”, *Anuario de Derecho Municipal* 2008, pp. 63-104

punto los privados han determinado la configuración de un proceso automatizado o un sistema de análisis de datos? ¿Hasta qué punto ha prevalecido el interés general? ¿Cuál es la posición de ese privado, que tiene acceso a las orientaciones de política pública, datos, etc.? ¿Todo ello está determinado en las condiciones de contratación?

La participación de privados en los procesos de automatización, toma de decisiones basadas en datos e IA debe ser instrumental, y por tanto no debe predeterminar decisiones que corresponden a la Administración. En definitiva, si bien es necesario contar con conocimiento y asistencia expertos externos, deben existir equipos internos que controlen el desarrollo y la ejecución de un proyecto, es decir, empleados públicos que participen en el diseño, implementación, supervisión y evaluación de algoritmos desde una perspectiva del interés público.

Por otro lado -desde un punto de vista estructural- las Administraciones deben reducir su dependencia externa fortaleciendo sus capacidades internas en ciencia de datos e IA a través de la atracción de expertos, construcción de redes de expertos o comunidades de práctica, o la colaboración con centros de investigación, entre otros²⁸. Como se ha afirmado, “Public organisations should ensure the presence of in-house knowledge on AI, for the – partial or complete – internal development of the solution, for the direction and adjustment of the system developed by external suppliers and/or for ensuring proper management of procurement activities”²⁹.

En definitiva las Administraciones deberían evitar la inercia en su colaboración con privados, en la que la dependencia o captura por parte de empresas -con sus propios intereses, legítimos pero que no responden necesariamente al interés general- es un riesgo real. El papel de las empresas tecnológicas y consultoras debe ser instrumental. Para ello es necesario que la Administración disponga de personal

²⁸ KUPI, M., JANKIN, S.; HAMMERSCHMID, G. (2022). *Data science and AI in government*. Hertie School. Recuperado de https://hertieschool-f4e6.kxcdn.com/fileadmin/2_Research/2_Research_directory/Research_Centres/Centre_for_Digital_Governance/2020-01_Documents/HS_Policy_Brief_English_Final_Version_Print.pdf

²⁹ JOINT RESEARCH CENTRE (2022): *AI Watch European Landscape on the Use of Artificial Intelligence by the Public Sector*, op. cit., p. 56

experto que pueda orientar esta colaboración y definir las actividades internamente y desde una perspectiva del interés público.

2.3.- *¿Dos clases de empleados? Si no es posible transformar a todos los empleados.... ¿Lo hacemos con una parte?*

Diferentes autores han puesto de manifiesto la necesidad de un relevo generacional en las plantillas públicas, ante su creciente envejecimiento. Estos enfoques han presupuesto un relevo de los efectivos por otros que respondan a otros perfiles, con un período en el que convivirán dos tipos de funcionarios de perfiles muy distintos y en el que hay que transmitir el *know how* a los nuevos empleados así como cambiar la cultura organizativa de forma paulatina³⁰. Esta es la teoría, aunque en la práctica los procesos de selección suelen incidir en los que de forma muy crítica se ha calificado como “procesos selectivos blandos”³¹ o simplemente como *aplantillamiento*³². La sustitución por empleados que respondan a los nuevos perfiles que demanda el sector público es limitada y condicionada por las prácticas de *consolidación* del personal temporal.

En este sentido, los procesos de estabilización derivados del Real Decreto-ley 14/2021, de 6 de julio, de medidas urgentes para la reducción de la temporalidad en el empleo público y la Ley 20/2021, de 28 de diciembre, de medidas urgentes para la reducción de la temporalidad en el empleo público no solo inciden en esta tendencia, sino que han introducido un elemento disruptor en los escasos cambios existentes. Los gestores de recursos humanos están actualmente centrados en los procesos de estabilización del personal temporal.

³⁰ PINEDA NEBOT, C. (2019): “Retos de futuro en la gestión de los recursos humanos públicos en España”, *Administração Pública e Gestão Social*, 11(4)

³¹ BOLTAÍNA BOSCH, X. (2018): “Los procesos selectivos “blandos” y sus efectos sobre la profesionalización del empleo público”, *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, nº. extraordinario 2.

³² JIMÉNEZ ASENSIO, R. (2018): “Función pública, selección e interinos: 10 puntos críticos [1]”, blog *La mirada institucional*, entrada de 21 de noviembre de 2018. Disponible en <https://rafaeljimenezasensio.com/2018/11/21/funcion-publica-seleccion-e-interinos-10-puntos-criticos-1/>

La situación actual hace que sea pragmático proponer un cambio a “dos velocidades”. Asumir que la mayor parte de los empleados públicos –incluidos los temporales incursos o no en los procesos de estabilización- no van a responder a los perfiles necesarios para la Administración digital -o lo van a hacer de forma parcial-, de forma que el nivel de exigencia sea menor. Por otra parte profundizar en la atracción de nuevos efectivos que sí respondan a esos nuevos perfiles profesionales, pero de acuerdo con nuevas reglas y con incentivos diferentes. Ramió lo ha expresado con cierto desánimo ante la parálisis actual:

“Mi propuesta sería la siguiente: vamos a continuar todos tal y como estamos (empleados públicos y unidades de gestión del personal). Plan- teemos un nuevo modelo solo de cara al futuro (los nuevos empleados públicos que se vayan a incorporar a partir de ahora), con reglas de juego totalmente distintas y que nos permita crear unidades de gestión de perso- nal nuevas y *ad hoc* para gestionar estos recién incorporados empleados públicos. Me parece que empezar desde cero con el nuevo modelo de función pública y duplicar las unidades de gestión de personal puede ser técnicamente más sencillo que apostar por un cambio en esta materia. Las nuevas unidades de gestión de personal encargadas de administrar el nuevo modelo pueden comenzar a trabajar sin capturas conceptuales, sin presiones e inercias históricas y de manera creativa e innovadora. Podrían iniciar su andadura sin complejos y de manera totalmente nueva. Empe- zarían por ser pequeñas unidades de personal frente a las grandes y con- solidadas que gestionarían el anterior no-modelo. El tiempo y el natural relevo intergeneracional se encargarían de hacer el resto del trabajo: poco a poco, las nuevas unidades de personal irían ocupando la centralidad del sistema y las viejas irían languideciendo hasta su total desaparición.”³³

La propuesta no deja de ser una hipérbole, presenta grandes di- ficultades y las teóricas nuevas unidades seguirían estando influen- ciadas –si no sometidas- a las “capturas conceptuales”, presiones e inercias de su entorno. Sin embargo la propuesta tiene el valor de intentar, como afirma el autor, “romper la actual dinámica de mera supervivencia reactiva por una dinámica de renovación” y estrategias proactivas que miren hacia el futuro³⁴. No es posible crear una *doble*

³³ RAMIÓ, C. (2022).: *Burocracia inteligente. Guía para transformar la Adminis- tración pública*. Madrid: Catarata. p. 103

³⁴ *Ibid.*, p. 104.

función pública, pero sí ofrecer incentivos a la transformación a través de las retribuciones y la promoción.

En este sentido es exigible a todo empleado público el cumplimiento de una alfabetización digital mínima, independientemente del perfil y tareas que desempeña. La formación, como veremos más adelante, es básica. No solo formación técnica, para el desempeño de sus funciones en un entorno digital, sino también de cambio en la cultura organizativa. El grueso del personal existente debería por lo tanto adquirir las competencias necesarias para el puesto de trabajo que desempeña.

Por otra parte el personal especializado debería estar sometido a reglas de selección específicas en base a su experiencia o especialización, habilidades (*skills*) y competencias, etc. Su relación con la Administración no debería ser necesariamente permanente y podría incorporar elementos –como la flexibilidad, trabajo por objetivos– más propios del sector tecnológico pero incorporando la visión del interés público. El marco regulador de la función pública permite introducir mejoras retributivas (por ejemplo incremento sustancial en el complemento específico o en el de productividad) u otras relacionadas con la movilidad o la promoción³⁵. Diseñar un marco atractivo para estos perfiles profesionalizados no es solo una cuestión de selección, sino también de retención³⁶.

3.- CUATRO PRESUPUESTOS NECESARIOS PARA LA ADAPTACIÓN DEL EMPLEO PÚBLICO

Aunque hay una necesidad de cambio de los recursos humanos para la automatización o la IA, y así lo confirman numerosos estudios al respecto, las políticas públicas en general no marcan una hoja de ruta clara en este sentido. A modo de ejemplo, la Estrategia nacional

³⁵ El uso de los datos en la gestión de recursos humanos ya se ha destacado como un campo en el que es posible introducir diferencias en el personal respecto a promoción o retribuciones variables. ROGGE, N.; AGASISTI, T.; DE WITTE, K. (2017): “Big data and the measurement of public organizations’ performance and efficiency: The state-of-the-art”, *Public Policy and Administration*, nº 32(4).

³⁶ CIVIL SERVICE (2016). *Civil Service Workforce Plan 2016 – 2020*

de inteligencia artificial (2020) dice poco sobre los recursos humanos necesarios en el sector público, sino solo recomendaciones genéricas de fomento del uso de la IA en la Administración (acceso a la función pública, la formación de los empleados públicos, la definición de puestos de trabajo, etc. (Medida 23)). Sin perjuicio de diferentes estudios sobre transformación digital existentes destacamos aquí, a modo de conclusión, cuatro “pasos” que parece lógico contemplar a la hora de transformar la Administración. En primer lugar, es necesario planificar que recursos humanos van a ser necesarios a corto, medio y largo plazo –lo que exige como paso previo identificar que actividades o servicios van a ser digitalizados, y si se va a tratar de una automatización, del uso de datos para mejora de los procesos, el *big data* o incluso la IA-. En segundo lugar valorar si los sistemas existentes de selección son adecuados y se adaptan a los perfiles que se quiere incorporar. En tercer lugar, adoptar medidas para transformar los recursos humanos existentes a las tareas que van a tener que llevar a cabo. Finalmente, incorporar la gestión a través de los datos también a los recursos humanos.

3.1.- Planificación

Existe poca planificación de los recursos humanos del sector público, a pesar de que el TREBEP dispone que ésta “[...] tendrá como objetivo contribuir a la consecución de la eficacia en la prestación de los servicios y de la eficiencia en la utilización de los recursos económicos disponibles mediante la dimensión adecuada de sus efectivos, su mejor distribución, formación, promoción profesional y movilidad” (art. 69.1 TREBEP). El TREBEP establece que las Administraciones públicas *podrán* aprobar Planes para la ordenación de sus recursos humanos (PORH), que pueden incluir –entre otras cuestiones- el “a) Análisis de las disponibilidades y necesidades de personal, tanto desde el punto de vista del número de efectivos, como del de los perfiles profesionales o niveles de cualificación de los mismos.”, y “b) Previsiones sobre los sistemas de organización del trabajo y modificaciones de estructuras de puestos de trabajo.” Las previsiones del TREBEP son especialmente pertinentes en el contexto de la adaptación de las plantillas a la transformación digital. Aun así no se desprende del texto su obligatoriedad, y la realidad nos muestra que las Administraciones no la suelen tener en consideración.

Las Administraciones se han centrado en la gestión ordinaria del personal³⁷ o en la “administración del personal”³⁸. Lejos queda la necesaria planificación y la gestión estratégica de los recursos humanos³⁹. La planificación en este sentido debe ser un ejercicio de visualizar la Administración a corto, medio y largo plazo, analizar los efectivos existentes y los necesarios para afrontar las actividades y servicios en el futuro. Más en el caso de la transformación digital –y eventualmente la implantación de sistemas automatizados o basados en los datos o la IA-, que exige redefinir los puestos de trabajo y la incorporación de nuevos perfiles.

Desafortunadamente, como ya hemos avanzado, no se planifica sino que se gestiona el día a día y se prevén necesidades a corto plazo. A los problemas estructurales de envejecimiento y relevo intergeneracional se añade actualmente la estabilización de una parte importante del empleo público, lo que absorbe la actividad de las unidades de gestión de recursos humanos. A este escenario de parálisis podemos unir las habituales presiones –de sindicatos y los propios empleados- para estabilizar los efectivos existentes, que en gran medida no responden a los perfiles que hemos enunciado anteriormente.

Siendo conscientes de las dificultades del contexto actual, no debemos dejar de insistir en la planificación, que es clave para diseñar la Administración a medio y largo plazo, y evitar su “obsolescencia”⁴⁰. No está en juego la “desaparición” de la Administración, sino su capacidad para intervenir en la sociedad y servir a sus fines de defensa del interés público. La progresiva penetración de los intereses priva-

³⁷ ARROYO YANES, L. M. (2008). “La ordenación de la actividad profesional en el Estatuto Básico del Empleado Público: planificación de recursos humanos y estructuración del empleo público”, en: LINDE PANIAGUA, E. (dir.): *El Estatuto Básico del Empleado Público y su desarrollo por el Estado y las Comunidades Autónomas*. Madrid: Colex.

³⁸ PRATS i CATALÀ, J. (1987). Política, gestión y administración de recursos humanos en las Administraciones Públicas. *Documentación Administrativa*, nº 210-211: 124

³⁹ CUENCA CERVERA, J. J. (2018). “Instrumentos de planificación de recursos humanos y selección: ¿Cambio de paradigma?”, *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, nº Extra 2

⁴⁰ JIMÉNEZ ASENSIO, R. (2018b). “Repensar” la selección de empleados públicos. Momento actual y retos de futuro, *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, Nº. Extra 2.: 24

dos en la actividad pública es una advertencia en este sentido. Una Administración fuerte, con personal cualificado, es una garantía para no estar condicionados a los intereses privados.

3.2.- Selección

Las críticas al sistema de selección de los empleados públicos son unánimes entre los diferentes autores⁴¹. El sistema actual es inadecuado en muchos sentidos. En el contexto de una Administración digital, simplemente no selecciona los mejores candidatos. El art. 61.2 TREBEP sin embargo dispone que en los procesos selectivos debe procurarse “la conexión entre el tipo de pruebas a superar y la adecuación al desempeño de las tareas de los puestos de trabajo”, incluyendo pruebas prácticas, en las que se comprueben los conocimientos y la capacidad analítica de los aspirantes a través de la realización de ejercicios que demuestren la posesión de habilidades y destrezas. El texto básico ya ofrece un marco adecuado, pero la realidad es que la inmensa mayoría de las pruebas selectivas se basan en pruebas test o similares que apenas miden las competencias y habilidades de los futuros empleados públicos.

A la falta de adecuación de los sistemas de selección podemos añadir otros problemas en el caso de los candidatos con perfiles profesionales, que como hemos dicho son básicos para una Administración que pretenda fundamentar su actuación en la automatización, los datos y la IA. Sin ánimo de exhaustividad podemos mencionar algunos elementos que dificultan –o directamente impiden– el acceso de estos perfiles: procesos de selección lentos, pruebas memorísticas, poca o nula valoración de la experiencia profesional en el sector privado o el peso excesivo de los méritos en fase de concurso, en los que se valora especialmente el desempeño en el mismo puesto al que

⁴¹ Entre otros vid. ARESES VIDAL, X. et. Al (2017). *Nuevos tiempos para la función pública. Propuestas para atraer y desarrollar el talento en la Administración General del Estado*. Madrid: INAP; FONDEVILA ANTOLÍN, J. (2018). *Problemas y soluciones al Empleo Público actual. Una valoración a los 10 años de la aprobación del EBEP*. Granada: CEMCI: 250, 253; FERNÁNDEZ RAMOS, S. (2017). “El acceso al empleo público”, en CASTILLO BLANCO, F. A. (dir.): *Remodelando el empleo público y la organización administrativa para garantizar el Estado del bienestar*. Madrid: Iustel: 171).

se está optando. Son necesarios nuevos sistemas de selección ágiles y que permitan la incorporación de empleados con las habilidades y competencias necesarias para adaptarse a situaciones cambiantes. Habría que explorar nuevos sistemas, que incluyan la captación de “talento” y hacer más atractivo el empleo público para este tipo de profesionales, incluyendo nuevas formas de relación laboral, como el trabajo por proyectos o la incorporación temporal de empleados del sector privado.

3.3.- *Transformación*

Teniendo en cuenta la selección de nuevos perfiles, es necesario un esfuerzo de formación de los recursos humanos existentes. Debe abordarse la formación de los empleados de acuerdo con las nuevas necesidades: formación continua como necesidad y como deber. La referencias a la formación en el TREBEP son vagas, como la obligación de los empleados públicos de mantener “actualizada su formación y cualificación” (art. 54.8 TREBEP), además de otras disposiciones menores (arts. 69 y 73 TREBEP). Habría que incidir en la especialización –incluida la formación externa–, la cooperación con otros entes públicos o mecanismos colaborativos como las comunidades de práctica.

Los empleados públicos deberán asumir que sus funciones serán cambiantes, adaptables a las diferentes situaciones y necesidades que la Administración vaya a afrontar en el futuro. Ello implica el aprendizaje a lo largo de la vida (*lifelong learning*), no solo materialmente sino como una actitud hacia el aprendizaje, la mejora y la innovación. En este sentido se dirigen nuevas propuestas que intentan especializar al empleado a través de nuevas competencias, habilidades y conocimientos (*upskilling*), además del reciclaje a través también de nuevas competencias, habilidades y conocimientos aplicables al puesto de trabajo (*reskilling*). Especialización y reciclaje deberían por tanto formar parte de la cultura organizativa y de los *deberes* del nuevo empleado público. Hay que, en definitiva, cambiar la concepción de puesto de trabajo –y de los empleados que los desempeñan–: más flexible, polivalente y adaptable, basado en competencias (técnicas y digitales, habilidades).

3.4.- Los datos en la gestión de los recursos humanos

El uso de los datos en la planificación, selección y gestión de los recursos humanos es limitado en el sector público, y generalmente se circunscribe a la gestión ordinaria. Las aplicaciones utilizadas, en general son contratadas a empresas privadas. Sin embargo, las potencialidades de los datos en la planificación, selección y gestión son muy grandes. Sin ánimo de exhaustividad⁴², a continuación enunciamos algunas áreas de potencial desarrollo:

Conocer el funcionamiento real de los efectivos, en conexión con los servicios públicos, los procesos y procedimientos administrativos y la organización en su conjunto. Nada impide en este sentido que sistemas basados en los datos detecten necesidades e ineficiencias y recomienden soluciones.

Automatizar buena parte de las tareas estrictamente de gestión, detectando a qué empleados afecta, para redistribuir tareas y puestos, como ya hemos visto anteriormente.

Detectar necesidades de formación o de nuevos efectivos, utilizando los datos sobre personal, así como los derivados de la actividad administrativa y la prestación de servicios públicos.

Evaluar empleados, puestos o unidades administrativas. Disponer de grandes cantidades de datos podría permitir la evaluación –y retribución variable si se considera- tanto individual de los empleados públicos.

Existen diversas fuentes de información contenidas en los instrumentos de ordenación (registro de personal, plantilla, catálogo, relación de puestos de trabajo), además de en aplicaciones de gestión del personal. Uno de ellos –la relación de puestos de trabajo- es un instrumento obligatorio, que debe contener, “al menos, la denominación de los puestos, los grupos de clasificación profesional, los cuerpos o escalas, en su caso, a que estén adscritos, los sistemas de provisión y las retribuciones complementarias.” (art. 74 TREBEP).

⁴² Sobre el uso de los datos en la gestión de los recursos humanos del sector público *vid.* GALINDO CALDÉS, R. (2019): “Big data e inteligencia artificial en la gestión de los recursos humanos del sector público”, *Revista Catalana de Dret Públic*, nº 58, pp. 49-63

Las relaciones de puestos de trabajo deberían ser el instrumento técnico para la organización efectiva de los recursos humanos de la Administración, de carácter dinámico⁴³ y “y no estático o pétreo”⁴⁴. Es, además, el instrumento esencial que conecta las necesidades de la organización con el estatuto del personal a su servicio⁴⁵, y desde un punto de vista más práctico dota a la entidad de clarificación organizativa, se produce el análisis de duplicidades y redundancias, y se identifica qué valor aporta cada puesto⁴⁶. Como ya hemos defendido anteriormente a partir de la relación de puestos de trabajo puede construirse un sistema objetivo y basado en datos⁴⁷. No solo porque la información contenida es datificable (tareas, habilidades, retribuciones complementarias, organigrama, etc.), sino porque se construye a través de un proceso objetivo, en tanto que requiere del análisis, descripción, valoración y clasificación del puesto o grupos de puestos de trabajo.

La relación de puestos de trabajo parece sobre el papel una solución idónea no solo para la gestión de los recursos humanos de una Administración, sino para disponer de información precisa para llevar a cabo una política de personal en la que se parta de una planificación fundamentada en datos y puedan identificarse necesidades de selección o de formación. Sin embargo se trata de un instrumento que presenta algunas dificultades en cuanto a su funcionamiento real. La principal es que su aprobación y/o modificación supone a menudo un proceso ineficaz, costoso y rígido, sometido también a presiones de sindicatos y empleados –no en vano es objeto de negociación sindical, de acuerdo con el art. 37 TREBEP, con unos límites poco precisos-. La relación de puestos de trabajo debería ser el instrumento de ordenación idóneo como cuadro de mando basado en datos, que fundamentara la política de recursos humanos, pero incluso muchos entes ni

⁴³ STS 902/2014, de 5 febrero de 2014, Sala de lo Contencioso

⁴⁴ SJCA 1.273/2018, 5 de Noviembre de 2018, de Logroño, FJ 2º

⁴⁵ MAURI MAJÒS, J. (2003): “Las relaciones de puestos de trabajo y la gestión de la función pública local”, *Cuadernos de Derecho Local*, nº 2: 43

⁴⁶ FERREIRO SEOANE, F. J. (2016). “La RPT como un instrumento de organización de las administraciones locales”, *REALA*, nueva época nº 5.: 79

⁴⁷ GALINDO CALDÉS, R. (2020): “Automatización, inteligencia artificial y empleados públicos”, PEGUERA POCH, M.; CERRILLO MARTÍNEZ, A. (dirs.): *Inteligencia artificial*, Cizur Menor: Aranzadi, pp. 93-112.

siquiera la han aprobado, aun siendo obligatorio. Falta pues voluntad de utilizarlas, especialmente como un instrumento estratégico de gestión de los recursos humanos.

BIBLIOGRAFÍA CITADA

- ARESES VIDAL, X. et. Al (2017). *Nuevos tiempos para la función pública. Propuestas para atraer y desarrollar el talento en la Administración General del Estado*. Madrid: INAP
- ARROYO YANES, L. M. (2008). “La ordenación de la actividad profesional en el Estatuto Básico del Empleado Público: planificación de recursos humanos y estructuración del empleo público”, en: LINDE PANIAGUA, E. (dir.): *El Estatuto Básico del Empleado Público y su desarrollo por el Estado y las Comunidades Autónomas*. Madrid: Colex.
- BERNING PRIETO, A. (2021): “Delimitación entre potestades administrativas y actividades técnicas o auxiliares”, en GAMERO CASADO, E. (dir.): *La Potestad Administrativa: concepto y alcance práctico de un criterio clave para la aplicación del Derecho administrativo*. Valencia: Tirant lo Blanch.
- BOLTAÍNA BOSCH, X. (2018): “Los procesos selectivos “blandos” y sus efectos sobre la profesionalización del empleo público”, *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, nº. extraordinario 2.
- CANALS AMETLLER, D.: *El ejercicio por particulares de funciones de autoridad. Control, inspección y certificación*. Granada: Comares, 2002.
- CERRILLO i MARTÍNEZ, A. (2021): “Robots, asistentes virtuales y automatización de las administraciones públicas”, *Revista Galega de Administración Pública* (REGAP), nº 61
- CERRILLO i MARTÍNEZ, A. (2020): “¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?”, *European Review of Digital Administration & Law* – Erdal, nº 1 (1-2)
- CERRILLO i MARTÍNEZ, Agustí (2019): “El impacto de la inteligencia artificial en el derecho administrativo ¿nuevos conceptos para nuevas realidades técnicas?”, *Revista General de Derecho Administrativo*, nº 50
- CIVIL SERVICE (2016). *Civil Service Workforce Plan 2016 – 2020*
- COMISIÓN EUROPEA (2018): *Comunicación de la Comisión al Parlamento Europeo, al Consejo Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones “Inteligencia artificial para Europa”*. COM(2018) 237 final
- CUENCA CERVERA, J. J. (2018). “Instrumentos de planificación de recursos humanos y selección: ¿Cambio de paradigma?”, *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, nº Extra 2

- DESOUZA, K. C.; JACOB, B. (2017): "Big Data in the Public Sector: Lessons for Practitioners and Scholars", *Administration & Society*, nº 49(7)
- DÍAZ GONZÁLEZ, G. M. (2020): "Algoritmos y actuación policial: la policía predictiva", en HUERGO LORA, A. (ed.): *La regulación de los algoritmos*, Cizur Menor, Aranzadi-Thomson-Reuters.
- ENGIN, Z.; TRELEAVEN, P. (2018). Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies. *The Computer Journal*, xy082.
- FERNÁNDEZ RAMOS, S. (2017). "El acceso al empleo público", en CASTILLO BLANCO, F. A. (dir.): *Remodelando el empleo público y la organización administrativa para garantizar el Estado del bienestar*. Madrid: Iustel
- FERREIRO SEOANE, F. J. (2016). "La RPT como un instrumento de organización de las administraciones locales", *REALA*, nueva época nº 5
- FONDEVILA ANTOLÍN, J. (2018). *Problemas y soluciones al Empleo Público actual. Una valoración a los 10 años de la aprobación del EBEP*. Granada: CEMCI
- FREY, C. B.; OSBORNE, M. A. (2017): "The future of employment: how susceptible are jobs to computerisation?", *Technological forecasting and social change*, 114.
- GALÁN GALÁN, A. (2019): *Entidades privadas colaboradoras de la Administración*. Valencia: Tirant lo Blanch.
- GALÁN GALÁN, A.; PRIETO ROMERO, C. (2009): "El ejercicio de funciones públicas por entidades privadas colaboradoras de la Administración", *Anuario de Derecho Municipal 2008*, pp. 63-104
- GALINDO CALDÉS, R. (2020): "Automatización, inteligencia artificial y empleados públicos", PEGUERA POCH, M.; CERRILLO MARTÍNEZ, A. (dirs.): *Inteligencia artificial*, Cizur Menor: Aranzadi, pp. 93-112.
- GALINDO CALDÉS, R. (2019): "Big data e inteligencia artificial en la gestión de los recursos humanos del sector público", *RCDP*, nº 58, pp. 49-63
- GAMERO CASADO, E. (2021): "El ejercicio de potestades administrativas en el empleo público", en GAMERO CASADO, E. (dir.): *La Potestad Administrativa: concepto y alcance práctico de un criterio clave para la aplicación del Derecho administrativo*. Valencia: Tirant lo Blanch.
- GENERALITAT DE CATALUNYA (2020): *Manual de transformació digital del empleado público*. Departament de Polítiques Digitals i Administració Pública. Accesible en [https://administracioidigital.gencat.cat/ca/teletreball/Guies-i-recursos/manuals/manual-de-transformacio-digital-de-lempleat-public/index.html#googtrans\(cales\)](https://administracioidigital.gencat.cat/ca/teletreball/Guies-i-recursos/manuals/manual-de-transformacio-digital-de-lempleat-public/index.html#googtrans(cales))
- GOBIERNO DE ESPAÑA (2021): *Plan Nacional de Competencias Digitales*
- HUERGO LORA, A. (2020): "Una aproximación a los algoritmos desde el derecho administrativo", en HUERGO LORA, A. (ed.): *La regulación de los algoritmos*, Cizur Menor, Aranzadi-Thomson-Reuters.

- JIMÉNEZ ASENSIO, R. (2018): “Función pública, selección e interinos: 10 puntos críticos [1]”, blog *La mirada institucional*, entrada de 21 de noviembre de 2018. Disponible en <https://rafaeljimenezasensio.com/2018/11/21/funcion-publica-seleccion-e-interinos-10-puntos-criticos-1/>
- JIMÉNEZ ASENSIO, R. (2018b). “Repensar” la selección de empleados públicos. Momento actual y retos de futuro, *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, N°. Extra 2
- JOINT RESEARCH CENTRE (2022): *AI Watch European Landscape on the Use of Artificial Intelligence by the Public Sector*. JRC Science for Policy Report
- KUPI, M., JANKIN, S.; HAMMERSCHMID, G. (2022). *Data science and AI in government*. Hertie School. Recuperado de https://hertieschool-f4e6.kxcdn.com/fileadmin/2_Research/2_Research_directory/Research_Centres/Centre_for_Digital_Governance/2020-01_Documents/HS_Policy_Brief_English_Final_Version_Print.pdf
- MARTÍN DELGADO, I. (2009): “Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada”, *Revista de administración pública*, n° 180
- MAURI MAJÒS, J. (2003): “Las relaciones de puestos de trabajo y la gestión de la función pública local”, *Cuadernos de Derecho Local*, n° 2
- PINEDA NEBOT, C. (2019): “Retos de futuro en la gestión de los recursos humanos públicos en España”, *Administração Pública e Gestão Social*, n° 11(4)
- PONCE SOLÉ, J. (2019): “Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico”. *Revista General de Derecho Administrativo*, n° 50.
- PRATS i CATALÀ, J. (1987). Política, gestión y administración de recursos humanos en las Administraciones Públicas. *Documentación Administrativa*, n° 210-211
- RAMIÓ, C. (2022).: *Burocracia inteligente. Guía para transformar la Administración pública*. Madrid: Catarata.
- RAMIÓ, C. (2019): “El impacto de la inteligencia artificial y de la robótica en el empleo público”, *GIGAPP Estudios Working Papers*, n°5 (98-110)
- RAMIÓ, C.; SALVADOR, M. (2018): *La Nueva Gestión del Empleo Público. Recursos Humanos e Innovación de la Administración*. Barcelona: Tibidabo Ediciones
- ROGGE, N.; AGASISTI, T.; DE WITTE, K. (2017): “Big data and the measurement of public organizations’ performance and efficiency: The state-of-the-art”, *Public Policy and Administration*, n° 32(4).
- ROUHIAINEN, L. (2018): *Inteligencia artificial. 101 cosas que debes saber hoy sobre nuestro futuro*, Planeta, Barcelona.

- SCHUETT, J. (2019): “Defining the scope of AI regulations”, *Computers and Society*, arXiv:1909.01095 [cs.CY] <https://doi.org/10.48550/arXiv.1909.01095>
- TODOLÍ SIGNES, A. (2020): “Retos legales del uso del big data en la selección de sujetos a investigar por la Inspección de Trabajo y de la Seguridad Social”, *Revista Galega de Administración Pública* (REGAP), n° 59, pp. 79-102
- TRUJILLO SÁEZ, F.; ÁLVAREZ JIMÉNEZ, D. (2021): “Transformación digital de la administración pública: ¿Qué competencias necesitan los empleados públicos?”. *Gestión y Análisis de Políticas Públicas*, (27), p. 49-67

CUARTA PARTE:
ORGANIZACIÓN Y GOBERNANZA

Capítulo XIII

Organización y gobernanza de la inteligencia artificial: marco general

JUAN CARLOS HERNÁNDEZ PEÑA

Profesor Contratado Doctor

Universidad de Navarra

Resumen: En este trabajo se aborda el marco general de gobernanza y organización de la IA a partir de RIA. En primer lugar, se hace referencia a la estructura normativa y los enfoques que subyacen en el diseño de la Unión Europea. Posteriormente se repasan algunos principios de especial relevancia: neutralidad, innovación, precaución y competencia desde el diseño. Finalmente, se abordan las estructuras de gobernanza multinivel, destacando su reparto competencial y algunas fórmulas de cooperación.

Abstract: This contribution addresses the general governance and organizational framework of IA following RIA. First, a reference is provided with regard to the normative structure and approaches underlying the design of the European Union IA Framework. Then, some principles of particular relevance are examined: neutrality, innovation, precaution and competition by design. We will finally address the multilevel governance structures, emphasizing the allocation of powers and some cooperative mechanisms.

Palabras-clave: Inteligencia artificial; Gobernanza; Principios.

Key-words: Artificial Intelligence; Governance; Principles.

SUMARIO. - 1. EL MARCO DE LA GOBERNANZA DE LA IA: UN TRINOMIO NORMATIVO BAJO UN ENFOQUE BASADO EN RIESGOS Y DERECHOS.- 2. ALGUNOS PRINCIPIOS QUE CONDICIONAN LA ACTUACIÓN DE LOS PODERES PÚBLICOS Y EL MODELO DE GOBERNANZA.- 2.1. Principio de neutralidad tecnológica.- 2.1. Principio de innovación.- 2.3 Principio de precaución.- 2.4. Principio de competencia por diseño.- 3. ESTRUCTURAS DE GOBERNANZA DE LA IA.- 3.1 El Comité Europeo de Inteligencia Artificial. Composición, competencias y funciones.- 3.2 Autoridades nacionales. Designación y competencias.- 3.3 Organismos notificados y organismos de evaluación de la conformidad.- BIBLIOGRAFIA CITADA.

1.- EL MARCO DE LA GOBERNANZA DE LA IA: UN TRINOMIO NORMATIVO BAJO UN ENFOQUE BASADO EN RIESGOS Y DERECHOS

A efectos de considerar el alcance de la gobernanza en materia de inteligencia artificial, es ineludible hacer referencia al modelo que se ha ido construyendo a nivel comunitario. La Unión ha venido pergeñando un marco que se instrumentaliza mediante un trinomio normativo (normas éticas, normas técnicas y normas de *hard law*) que estructura la intervención en diferentes momentos y con distinto alcance, pero que deriva en un intenso proceso de fertilización cruzada y de fuerte dirección pública.

Las normas éticas, de alto nivel, actúan a modo de precondition para la aprobación de normas técnicas y el marco jurídico de *hard law*, informando también la actuación del sector público. Las normas de estandarización, que podríamos ubicar en un segundo círculo, contribuyen a optimizar las directrices éticas, así como la normativa jurídica europea, además de facilitar el desarrollo y circulación a nivel global de las aplicaciones de IA. Las normas jurídicas, aún de alcance limitado, estructuran -entre otros aspectos- la ineludible protección de los derechos fundamentales, la salud o la seguridad de las personas. La cascada normativa se inicia con RIA, aunque en ejercicio de sus competencias los Estados miembros también están adoptando algunas medidas de naturaleza normativa¹.

Avanzamos que este capítulo se centrará fundamentalmente en la propuesta comunitaria, que recoge el inacabado modelo de gobernanza multinivel de la Unión, así como en una serie de principios que condicionan y dirigen la actuación de los *stakeholders*, incluyendo los poderes públicos bien sea que actúen como proveedores, importadores o usuarios de sistemas de IA, bien ejerciendo sus competencias

¹ Así, por ejemplo, el art. 23 de la reciente Ley 15/2022, de 12 de julio, integral para la igualdad de trato y no discriminación, aborda mediante normas de carácter finalista cuestiones relacionadas con la transparencia, la minimización de sesgos o la discriminación mediante decisiones automatizadas basadas en IA. También de art. 13 del RAFESP incorporó una referencia a la actuación administrativa automatizada, aunque interesa destacar que no existe equivalencia entre automatización y decisiones o procedimientos basados en IA.

y potestades de supervisión. El alcance y valoración de las normas éticas han sido tratadas suficiente por la doctrina², y poco interesa añadir. Lo mismo puede decirse respecto a la estandarización, normas técnicas o el “nuevo enfoque” armonizador comunitario, sobre el que pivota la intervención de la Unión en esta materia³.

Nótese que la propuesta de reglamento se valorará en otro capítulo de esta obra, al que remitimos. Únicamente destacamos que RIA tiene un objeto ciertamente limitado⁴. No persigue una regulación agotadora y sistemática de esta tecnología, y que evita regular cuestiones de relevancia como la responsabilidad, la propiedad intelectual o mecanismos específicos de tutela de derechos de los afectados⁵. Aspectos que cobran especial relieve ante supuestos de aplicación de modelos de IA por AAPP⁶.

Ahora bien, antes de pasar a analizar los principios y las estructuras de gobernanza, haremos algunas precisiones adicionales. En primer lugar, la intervención a nivel de la UE atiende a la necesidad de no fragmentar excesivamente el mercado único digital, así como su

² Así, entre otros, MORENO REBATO, M. (2021): *Inteligencia artificial (Umbrales éticos, Derecho y administraciones públicas)*, Thomson Reuters Aranzadi, Cizur Menor.

³ Sin ánimo de ser exhaustivo, puede verse ÁLVAREZ GARCÍA, V. (2010): *Derecho de la regulación económica. VIII. Industria*, Iustel, Madrid; ÁLVAREZ GARCÍA, V. (2020): *Las normas técnicas armonizadas (una peculiar fuente del Derecho europeo)*, Iustel, Madrid; DARNACULLETA i GARDELLA, M. M. (2004): *Autorregulación y Derecho público: La autorregulación regulada*, Marcial Pons, Madrid.

⁴ Esto es, establecer un marco armonizado que regule la introducción en el mercado, puesta en servicio y utilización de sistemas de inteligencia artificial en la Unión, incluyendo normas y procedimientos sobre control y vigilancia del mercado; prohíbe determinadas prácticas, fija un cúmulo de requisitos horizontales y obligaciones vinculadas a los sistemas que califica de alto riesgo; y, establece algunas obligaciones transversales para sistemas que interactúen con humanos.

⁵ COTINO HUESO, L. *et al.* (2021): “Un análisis crítico constructivo de la Propuesta de Reglamento de la Unión Europea por el que se establecen normas armonizadas sobre la Inteligencia Artificial (Artificial Intelligence Act)”, *Diario La Ley*.

⁶ Destaca la importancia de reconstruir el sistema de garantías jurídicas para este caso, VALERO TORRIJOS, J. (2019): Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración, *RCDP*, N.º 58, pp. 86 y ss.

competitividad en el marco de la economía global de los datos⁷. Por tanto, la dirección que ejerzan los poderes públicos nacionales ha de tener en consideración este objetivo.

Por otra parte, se promueve un ecosistema basado en derechos y en la intervención quirúrgica atendiendo a los riesgos. Esto se traduce en diferenciar entre distintas aplicaciones y someterlas a regímenes jurídicos, estándares y técnicas de intervención diferenciadas atendiendo a la capacidad de afectación de la libertad personal o los derechos fundamentales. En unos casos, se exigiría una intervención especialmente intensa, mientras que en otros pueden acogerse a fórmulas de autorregulación, códigos de conducta e incluso etiquetados voluntarios, informados por normas técnicas y estándares éticos aprobados por la UE. De esta forma persigue un marco sujeto a una intervención proporcionada, que otorgue seguridad jurídica bajo un enfoque humano-céntrico y sujeto a control humano.

Con esto se pretende que la IA comunitaria sea ética y fiable, garantizando -como bien señaló la doctrina- un ecosistema de confianza⁸, que refleje la apuesta por los insoslayables valores europeos y los derechos fundamentales. En este sentido, y aunque ya se hayan detectado deficiencias en el diseño previsto por la Unión⁹, no se admite transar derechos de los ciudadanos, afectar su seguridad o poner en riesgo su salud, en pos de la competitividad¹⁰. La función directiva de los poderes públicos debe dirigirse en todo caso a crear un entorno

⁷ No se trata de una argumentación especulativa de escaso impacto. Basta contrastar las estrategias nacionales de varios países miembros aprobadas antes de RIA, para constatar que ya se recogían distintas aproximaciones y respuestas para conjurar los riesgos de la IA, derivando en falta de seguridad jurídica para los stakeholders comunitarios.

⁸ Gamero Casado, E. (2021): “El enfoque europeo de inteligencia artificial”, *Revista de Derecho Administrativo – CDA*, N.º 20, 2021.

⁹ Así, la insuficiencia de la regulación vigente sobre seguridad de los productos frente a una tecnología que inherentemente conlleva su modificación a lo largo de su ciclo de vida; problemas relacionados con decisiones algorítmicas; o la compleja cuestión de los algoritmos de caja negra y el déficit de transparencia que puede afectar el ejercicio de derechos fundamentales. Sobre estas cuestiones se ocupan otros trabajos recogidos en esta obra a los que remitimos.

¹⁰ Interesa destacar, en línea con esto, que una de las razones que motivaron a los Estados miembros a adoptar una acción coordinada en esta materia, fue precisamente garantizar un elevado nivel de protección y seguridad de los ciudadanos

coherente y ajustado a estos objetivos, ya sea mediante regulación ex-ante o bien mediando intervenciones ex-post, así como técnicas idóneas y proporcionadas para mitigar los riesgos inadmisibles a los derechos fundamentales, si bien recordando que esto no exige condenar el fomento de las inversiones ni la innovación.

Finalmente, aunque se adopta un enfoque horizontal mediante un instrumento jurídico que introduce un mayor grado de homogeneidad, no se persigue una armonización total y exhaustiva. De hecho, se garantiza ciertos espacios competenciales para los Estados miembros, sobre los que daremos cuenta con la extensión del caso en el último apartado de este trabajo.

2.- ALGUNOS PRINCIPIOS QUE CONDICIONAN LA ACTUACIÓN DE LOS PODERES PÚBLICOS Y EL MODELO DE GOBERNANZA

La depuración de principios consustanciales a la relación Derecho-tecnología representa una alternativa razonable y proporcionada, para construir modelos de gobernanza resilientes. Del marco político-jurídico de la IA -aún en proceso de despliegue-, interpretado sistemáticamente atendiendo al acervo comunitario y los valores comunes de los Estados miembros, se destilan algunos principios a los que interesa prestar atención¹¹.

Nótese que sólo haremos referencia a unos pocos principios. La razón no es sólo la habitual limitación de este tipo de trabajos, sino que pretendemos evitar introducirnos en cuestiones abordadas en otros capítulos de esta obra. Por tanto, únicamente daremos cuenta de los principios de neutralidad tecnológica, innovación, precaución y competencia por defecto. Cuestiones relacionadas con una IA humanocéntrica, o el diseño por defecto de políticas y programas normativos

y de sus derechos fundamentales. Véase en este sentido la evaluación de impacto de RIA, p. 32.

¹¹ Sobre estos principios ya nos hemos ocupado en HERNÁNDEZ PEÑA, J.C. (2022): El marco jurídico de la inteligencia artificial. Principios, procedimientos y estructuras de gobernanza, Thomson-Reuters Aranzadi, Cizur Menor, pp. 101 y ss.

respetuosos de la privacidad, la protección de datos o la seguridad, se remiten a las colaboraciones correspondientes.

2.1.- Principio de neutralidad tecnológica

El principio de neutralidad tecnológica está asentado en la Unión Europea desde 1997, acuñándose en el sector de las comunicaciones electrónicas¹². Tiene estrecha conexión con la promoción de la equivalencia tecnológica, y la adopción de un marco regulatorio común -si es técnica y jurídicamente posible- entre los entornos online y offline, siguiendo los lineamientos de la Conferencia Ministerial de Bonn¹³.

Es de especial importancia en la construcción de marcos normativos vinculados a las TIC, dado que actúa a modo de barrera frente a la posible discriminación de unas tecnologías frente a otras o de políticas desproporcionadas que impidan su desarrollo. Por reflejarlo con un ejemplo, piénsese en una definición de inteligencia artificial que haga referencia al uso de computación tradicional excluyendo de manera voluntaria o inadvertida nuevas fórmulas prometedoras, como la computación neuromórfica. Una definición que favorezca una tecnología existente generaría una ventaja competitiva para empresas incumbentes de ese sector excluyendo injustificadamente a otras, reduciendo la competencia y las posibles opciones de los consumidores. El principio exige, por tanto, que el diseño normativo, así como las políticas o actuaciones de las AAPP, no establezcan preferencias por una tecnología, salvo que exista una razón de suficiente entidad para restringirla, sometida siempre al principio de proporcionalidad y con

¹² Como recuerda VAN DER HAAR, I. (2007): “Technological Neutrality. What Does It Entail?”, TILEC Discussion Paper DP 2007-009, p. 1, el principio se apareció por vez primera en el marco de la Comunicación sobre la revisión de las comunicaciones de 1999. Sobre la incorporación de este principio en el Derecho comunitario, véase GUTIÉRREZ ALONSO, J. J. (2009): “El elemento tecnológico: el principio de neutralidad y sus implicaciones en la regulación”, en Muñoz Machado, S. y Esteve Pardo, J. (Dirs.), *Derecho de la regulación económica. I. Fundamentos e instituciones de la regulación*, Iustel, Madrid, pp. 692 y ss.

¹³ Conferencia ministerial europea: Global information networks: Realising the potential, Bonn, 6 -8 de julio de 1997, apdo. 22. No obstante, ya se había promovido su adopción en Estados Unidos una década antes en el marco de la regulación federal de comunicaciones electrónicas. REED, C. (2007): “Taking Sides on Technology Neutrality”, *SCRIPT-ed*, Vol. 4, N.º 3, p. 264.

escrupuloso respeto del contenido esencial de los derechos fundamentales afectados. Con esto se busca, como ha señalado la doctrina, que el mercado tecnológico sea suficientemente abierto para equilibrar las posibilidades entre competidores¹⁴.

Ahora bien, interesa destacar que el principio no proscribe la regulación dirigida específicamente a una tecnología. Mantener una posición absolutista en este sentido supondría excluir cualquier iniciativa normativa o política pública dirigida a mitigar o excluir los riesgos de la IA. Requerirá, eso sí, una justificación de suficiente calado si puede condicionar su desarrollo. Esta justificación puede encontrarse en el impacto negativo e inasumible sobre derechos fundamentales. También puede derivar de la necesidad de conjurar riesgos sociales, políticos o económicos que han de ser valorados por aquellos que ostentan legitimidad democrática. En todo caso, la virtualidad del principio se manifestará en la necesidad de aquilatar y explicar adecuadamente los objetivos que persigue la regulación o intervención de los poderes públicos.

2.2.- Principio de innovación

El principio de innovación no tiene recogimiento directo en el Derecho originario ni en normativa secundaria de la UE. Tampoco formaba parte de los derechos nacionales de los Estados miembros, ni proviene de la Comisión o sus agencias. Fue propuesto en 2013 por *European Risk Forum*, organización que representa los intereses de empresas innovadoras, que han mantenido una posición combativa respecto a la aplicación del principio de precaución y su impacto negativo sobre la autorización y puesta en el mercado de productos¹⁵. No obstante, el Consejo Europeo lo asumió y promovió su aplicación en el marco de los paquetes de mejora de la regulación¹⁶.

¹⁴ GUTIÉRREZ ALONSO, J. J. (2009), *op.cit.*, p. 692.

¹⁵ En este sentido, GARNETT, K., *et al.* (2018): “Towards an innovation principle: An industry trump or shortening the odds on environmental protection?”, *Law, Innovation and Technology*, Vol. 10, N.º 1, pp. 1–2.

¹⁶ Véase, Conclusiones del Consejo de la Unión Europea: Better regulation to strengthen competitiveness, de 26 de mayo de 2016, apartado 2º, donde fomenta la aplicación del principio de innovación al desarrollar o revisar marcos norma-

En nuestro caso, la declaración sobre cooperación en materia de inteligencia artificial promueve un enfoque basado en este principio, al igual que la Comunicación “Inteligencia artificial para Europa”¹⁷. En este sentido, aunque el marco normativo promovido por la UE sea un sello distintivo para generar fiabilidad y confianza, no olvida el objetivo de alcanzar una posición global de liderazgo¹⁸, lo que exige fomentar la inversión público-privada. Por tanto, la creación de un ecosistema de confianza, no sólo se relaciona con la construcción de un marco basado en los derechos fundamentales y valores de la Unión. También se identifica con la seguridad jurídica que dimana del propio Estado de Derecho. Esto tiene como correlato generar un entorno inversor fiable.

Desde la perspectiva de aplicación, el principio establece mandatos de valoración y optimización entre posibles fuerzas contrapuestas en mercados altamente innovadores afectados por riesgos sociales o ambientales de consideración. En este sentido, se dirige a adoptar e implementar regulación y políticas públicas que fomenten la innovación manteniendo los estándares de protección de los ciudadanos, la salud o el medio ambiente. Cómo puede concluirse, el principio de innovación puede entrar en disputa con el principio de precaución, recogido por los Tratados y abundantemente por la normativa secundaria, así como los Derechos internos. Aunque sea imposible abordar esta problemática con detalle en este trabajo, la perspectiva que se propugna desde la UE no supone claudicar respecto a la contención o reducción de riesgos sociales o ambientales derivados de la aplicación del principio de precaución, sino que exige atemperar posiciones maximalistas que imposibiliten la innovación y los efectos, también positivos, para la sociedad. Lo dicho es importante, y debe destacarse suficientemente que el principio de innovación no supone *per se* una apuesta por la desregulación, sino por su complementariedad y pleno ajuste con los derechos fundamentales¹⁹.

tivos de todas las políticas comunitarias, haciendo además un llamado a los Estados miembros a evaluar y determinar su potencial.

¹⁷ Véase en la Comunicación, COM(2018) 237 final, la nota al pie 56.

¹⁸ Comunicación: Completar un Mercado Único Digital de confianza y para todos, COM(2018) 320 final, de 15 de mayo de 2018, p. 14.

¹⁹ En este sentido, EPSC (2016): “Towards an Innovation Principle Endorsed by Better Regulation”, EPSC Strategic Notes, Comisión Europea, N.º 14, 30 de junio, pp. 39-40.

2.3.- Principio de precaución

El principio de precaución representa un postulado casi universal de la sociedad post-industrial, categorizada como sociedad del riesgo²⁰. Adquiere pleno sentido en supuestos en los que la incertidumbre es la nota dominante, por lo que no es de extrañar que se haya recurrido a él ante la tangibilización de riesgos derivados del uso de la inteligencia artificial²¹. El principio ha sido estudiado con intensidad por la doctrina más autorizada a la que remitimos²². Dadas las limitaciones de este trabajo, nos centraremos únicamente en realizar una ajustada valoración sobre su aplicación en el caso de la IA.

Son comunes los planteamientos que presentan como un juego de suma cero la pugna entre precaución e innovación tecnológica²³. Algunos plantean que una admisión sin cortapisas del principio de precaución socavaría los beneficios que la IA brindaría a la sociedad²⁴, justificado en la similitud entre algunos riesgos de esta tecnología y los que derivan del resto de las TIC, ya asumidos socialmente²⁵. Desde

²⁰ Al respecto, BECK, U. (2006): *La sociedad del riesgo. Hacia una nueva modernidad*, Paidós, Barcelona, pp. 75 y ss.

²¹ Así, el Parlamento Europeo al considerar los principios que han de regular la robótica y la inteligencia artificial con fines civiles. Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica [2015/2103(INL)]. También lo ha hecho doctrina: COTINO HUESO, L., (2019): “Riesgos e impactos del big data, la inteligencia artificial y la robótica. Enfoques, modelos y principios de la respuesta del Derecho”, *RGDA*, N.º 50, pp. 27-28.

²² Entre otros trabajos, ESTEVE PARDO, J. (2021), *Principios de Derecho regulatorio. Servicios económicos de interés general y regulación de riesgos*, Marcial Pons, Madrid, pp. 39-49, 165 y ss.

²³ Se ha pronunciado por aplicar con intensidad el principio de precaución en nuestro ámbito, entre otros, MANTELERO, A. (2020): “Chapter III. Analysis of international legally binding instruments. Final report”, Council of Europe Study DGI, Towards regulation of AI systems, Consejo de Europa, p. 72. En contra, por su carácter restrictivo CASTRO, D. y McLAUGHLIN, M. (2019): *Ten Ways the Precautionary Principle Undermines Progress in Artificial Intelligence*, ITIF.

²⁴ Así, CASTRO, D. y McLAUGHLIN, M. (2019), *op.cit.*, pp. 1-35.

²⁵ En este sentido, el CESP Task Force, ha sostenido que la mayoría de las soluciones de IA no introducen riesgos éticos o regulatorios significativamente diferentes a los que han derivado de otras tecnologías de la información. RENDA, A. (2019): *Artificial Intelligence. Ethics, governance and policy challenges*, Center for European Policy Studies (CESP), Bruselas, p. 74.

nuestra perspectiva es aconsejable huir de posiciones maximalistas. Es razonable alinearse con una interpretación equilibrada que permita ajustar la aplicación del principio, reservando espacio para la innovación. Entendemos que este es el sentido de la doctrina española que se ha ocupado del tema²⁶, incluso de aquellos que proponen moratorias de aplicaciones específicas o vinculándolas a garantías de intervención humana²⁷. También es el que defiende el Parlamento Europeo²⁸, y es el enfoque adoptado en ámbitos como la protección de datos, donde tratamientos especialmente invasivos se someten a análisis de riesgos.

Ahora bien, este enfoque -entendemos- permite justificar moratorias para valorar el desarrollo de determinadas aplicaciones, pero no persigue *per se* establecer prohibiciones si cabe recurrir a medidas menos restrictivas que protejan derechos y bienes fundamentales, en aplicación del principio de proporcionalidad. Por tanto, se trataría de hacer valoraciones casuísticas en atención a las distintas aplicaciones, activando -de ser necesario- mecanismos preventivos ante riesgos inaceptables. Los procedimientos que recoge RAI, canaliza en cierta medida las competencias con las que cuentan las administraciones nacionales en esta materia.

2.4.- Principio de competencia por diseño

El impacto de la IA sobre el funcionamiento competitivo de los mercados no se separa de lo señalado en otros apartados. Así, la economía de datos alimentada por IA puede generar interesantes ganancias competitivas para los consumidores²⁹. Las ventajas son variadas.

²⁶ COTINO HUESO, L., (2019), *op.cit.*, p. 27; CERRILLO i MARTÍNEZ, A. (2019): “El impacto de la inteligencia artificial en el Derecho administrativo ¿Nuevos conceptos para nuevas realidades técnicas?”, *RGDA*, N.º 50, pp. 11-12.

²⁷ En este sentido, se pronuncia PONCE SOLÉ, J. (2019A): “Inteligencia artificial, Derecho administrativo y reserva de humanidad: Algoritmos y procedimiento administrativo debido tecnológico”, *RGDA*, N.º 50, p. 32; CERRILLO i MARTÍNEZ, A. (2020): “¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?”, *European Review of Digital Administration & Law – Erdal*, Vol. 1, N.º 1-2, p. 25-26.

²⁸ Véase Resolución de 12 de febrero de 2019, “A comprehensive European industrial policy on artificial intelligence and Robotics”, apdo. 19.

²⁹ BANICEVIC, A., *et al.* (2018): “Algorithms: Challenges and opportunities for antitrust compliance”, *ABA Section of Antitrust Law’s Compliance and Ethics*

Puede aportar eficiencia ante cambios en las condiciones del mercado, permitiendo ajustar oferta y demanda. Desde la perspectiva del consumidor o usuario, se beneficiarían de la reducción de costes derivados de la aplicación de IA a procesos productivos, e incluso del ajuste de bienes y servicios a sus preferencias, incrementando su bienestar desde el punto de vista competitivo. Finalmente, puede permitir adoptar mecanismos para incrementar el poder de consumidores y usuarios, por ejemplo, agregando la demanda y creando plataformas de compra basadas en modelos de IA³⁰.

En contraposición, los riesgos y retos para el Derecho de la competencia aconsejan renovar el arsenal regulatorio para reducir impactos negativos. El poder y posición de dominio que ostentan grandes tecnológicas cuestiona los cimientos del ordenamiento *Antitrust*. En este sentido, la poco sospechosa OCDE, señaló que la experiencia acumulada debe al menos alertar respecto a la mayor facilidad de riesgos colusorios cuando participan algoritmos³¹. A la misma conclusión ha llegado la doctrina que se ha ocupado del tema³², identificando también riesgos de discriminación de precios³³.

Los algoritmos de IA pueden ayudar a explotar fallos en mercados, facilitando su distorsión y manipulación. Estos efectos anticompetitivos pueden ser resultado de actuaciones voluntarias dirigidas -por ejemplo- a monitorizar la actividad de los competidores³⁴, lo que facilitaría conductas colusorias tácitas. Por otra parte, los

Committee. Fall 2018, special report, p. 7.

³⁰ PETIT, N. (2018): “Emerging competition, innovation, and market structure issues around algorithms, artificial intelligence, and predictive analytics”, *Submission to the FTC hearings on competition and consumer protection in the 21st century*, p. 6.

³¹ OECD (2017): *Algorithms and collusion. Competition Policy in the Digital Age*, OECD Publishing, Paris, p. 24.

³² Así, entre otros, BANICEVIC, A., (2018), *op.cit.*, p. 9; EZRACHI, A. y STUCKE, M.E. (2016), *Virtual Competition. The promise and perils of the algorithm-driven economy*, Harvard University Press, Cambridge, pp. 35 y ss.

³³ Ezrachi, A. y Stucke, M.E. (2016), *op.cit.*, pp. 83 y ss.

³⁴ Respecto a los supuestos de colusión, EZRACHI, A. y STUCKE, M. E. (2020): “Sustainable and Unchallenged Algorithmic Tacit Collusion”, *Northwestern Journal of Technology & Intellectual Property*, Vol. 17, N.º 2, p. 224-230, detallan las condiciones necesarias para que puedan desarrollarse y mantenerse estas prácticas anticompetitivas.

riesgos anticompetitivos también podrían derivarse de supuestos de *hub-and-spoke* alimentados por algoritmos de IA sin mediar comunicación explícita³⁵.

No sólo preocupa el ejercicio voluntario de esas posiciones de dominio, cuestión que puede encajarse con cierta coherencia dentro de la respuesta destilada por la UE³⁶ y otras jurisdicciones a lo largo de los años³⁷, sino supuestos donde la cadena de responsabilidad empieza a difuminarse dificultando la labor de supervisores y reguladores. En este sentido, los algoritmos de *machine learning* y *deep learning* introducen nuevos problemas. La posibilidad de que se adapten a las condiciones competitivas del mercado, pueden arrojar conductas anticompetitivas sin que hayan sido programados explícitamente para ello. La misma OCDE ha destacado que es posible que los algoritmos aprovechen su capacidad de aprendizaje y ensayo/error, para eventualmente alcanzar equilibrios cooperativos³⁸. La Comisión Europea en el *Inception Impact Assessment* del *New Competition Tool*, comparte la misma preocupación³⁹. La problemática se acentúa si se recurre a algoritmos de caja negra y algunos autores han establecido que los de *backpropagation* pueden incluso concluir que, para alcanzar de manera más eficiente su propósito, deben eliminar a sus competidores⁴⁰.

Por ello, en la conferencia internacional sobre competencia de 2017, organizada por el *Bundeskartellamt*, la Comisaria Europea de Competencia Margrethe Vestager, al hilo de la -entonces reciente- aprobación del RGPD, señalaba que algunos de sus instrumentos podrían ser útiles para afrontar los riesgos de la economía digital. Apuntaba a las cláusulas desde el diseño, recomendando que las empresas asumieran un enfoque similar diseñando sus productos y servicios para dar cumplimiento por defecto a la normativa de competencia⁴¹,

³⁵ Así, EZRACHI, A. y STUCKE, M.E. (2016), *op.cit.* p. 47 y ss.

³⁶ Puede verse, por ejemplo, el asunto Eturas, STJUE C-74/14 de 21 de enero de 2016.

³⁷ Así, por ejemplo, en el Derecho antitrust norteamericano puede citarse el caso del cartel de los pósteres de 2015 (US v. David Topkins).

³⁸ OECD (2017), *op.cit.*, p. 33.

³⁹ Véase el apartado referente al contexto de dicha propuesta regulatoria.

⁴⁰ En este sentido, PETIT, N. (2018), *op.cit.*, p. 3.

⁴¹ Gráficamente señala la Comisaria: “Like a more honourable version of the computer HAL in the film *2001*, they need to respond to an offer of collusion by

evitando así la colusión o conductas anticompetitivas. La propuesta se recogió en el anexo al plan coordinado de IA de 2018, donde la Comisión incluyó la competencia por diseño como uno de los elementos medulares para construir el mercado digital único⁴². Mandato extensible, como es lógico, a la actuación de las AAPP, por ejemplo, cuando compitan en el mercado.

Esto, en definitiva, exige diseñar y aplicar instrumentos que promuevan el alineamiento de la economía digital con el Derecho de la competencia. En nuestro caso concreto, que los algoritmos y modelos de inteligencia artificial se diseñen orientados a evitar conductas anticompetitivas. Deben dedicarse esfuerzos no sólo a embeber en los diseños la normativa de competencia, sino también entender en profundidad cómo funcionan sus algoritmos para desactivar tempranamente posibles comportamientos anticompetitivos no esperados o derivados de decisiones algorítmicas, especialmente en los casos de IA no supervisadas y basadas en cajas negras. Por tanto, se trata de incorporar soluciones orientadas a promover la competencia ex-ante, y no meramente reactivas.

La adopción del principio, no obstante, no está exenta de dificultades⁴³. Algunos de estos problemas son comunes a otras aplicaciones sectoriales de IA, como es el caso de la opacidad y transparencia, así como la problemática inherente a la explicabilidad de algunos modelos. Otros son propios, como la naturaleza de las normas de competencia, muchas de carácter finalista que dejan en manos de las autoridades de competencia completar el programa normativo⁴⁴. Esto deriva en la necesidad de recurrir no sólo a procedimientos automatizados, sino a garantizar la supervisión humana, así como la permanente colaboración entre aplicadores y autoridades supervisoras.

Conviene realizar algunos comentarios finales. En primer lugar, las exigencias de la competencia por diseño, deben aquilatararse con el principio de innovación, así como las libertades económicas. En este

saying: I'm sorry, I'm afraid I can't do that".

⁴² Anexo al plan coordinado sobre IA, p. 20.

⁴³ Así, entre otros, PETIT, N. (2018), *op.cit.*, p. 3.

⁴⁴ Sobre esto, véase HERNÁNDEZ, J. C. (2011): *El régimen jurídico-administrativo de las concentraciones de empresas en el Derecho comunitario: Test sustantivo y control judicial*, Thomson-Reuters Aranzadi, Cizur Menor, p. 36 y ss.

sentido, la autoridad alemana de competencia ha recordado que las medidas regulatorias deben contenerse de forma tal que los estándares establecidos no bloqueen estrategias algorítmicas legítimas ni creen barreras de acceso al mercado debido al incremento de costes⁴⁵.

Por otra parte, no solo debe promoverse la competencia por diseño, sino además la competencia por defecto. Por tanto, plataformas, aplicaciones, modelos y algoritmos que utilicen IA deben establecer como configuración estándar actuaciones pro-competitivas, aun sin que medie la participación o se requiera actuación alguna por parte de los usuarios. Se trata pues, de una aproximación *ex-ante*, cuyo interés no sólo concierne a consumidores y reguladores, sino también a las empresas que de esta manera pueden evitar futuras sanciones.

4.- ESTRUCTURAS DE GOBERNANZA DE LA IA

Para implementar el marco comunitario se establece un sistema de gobernanza multinivel que incluye estructuras de cooperación entre la Comisión, los Estados miembros y una pléyade de instituciones u organizaciones a las que se les atribuyen funciones relacionadas con la IA. Se trata, no obstante, de una fórmula de coordinación integrada dentro de la administración compartida europea, donde autoridades y organismos de nueva creación, comités diversos y grupos de trabajo se interrelacionan mediante sistemas de intercambio de información y complejos procedimientos de comunicación, consenso y resolución compartidas⁴⁶.

Para reticular esta cooperación resulta central el recién creado Comité Europeo de Inteligencia Artificial. Por otra parte, la propuesta se armoniza con la regulación sectorial en materia de seguridad de los productos, por lo que también participan las estructuras y autoridades competentes en atención a la normativa interna: autoridades

⁴⁵ Al respecto, véase PICHT, P.G. y LODERER, G.T. (2019): “Framing algorithms: Competition law and (other) regulatory tools”, *World Competition*, Vol. 42, N.º 3, p. 415.

⁴⁶ SCHIMDT-ASSMANN, E. (2021): *La dogmática del Derecho administrativo*, Global Law Press, Sevilla, pp. 320-321.

nacionales de supervisión, autoridades de vigilancia del mercado y autoridades notificantes.

4.1.- El Comité Europeo de Inteligencia Artificial. Composición, competencias y funciones

Un eslabón fundamental del sistema de gobernanza de la IA es el Comité Europeo de Inteligencia Artificial. Está compuesto por un grupo de autoridades naturales: las autoridades nacionales de supervisión, representadas por su máximo representante o funcionario equivalente; el Supervisor Europeo de Protección de Datos; y, la Comisión. A ésta corresponde presidirlo, convocar las reuniones y elaborar la orden del día, así como brindar su estructura administrativa y apoyo técnico con el objetivo de que los análisis sean sólidos y se realicen sobre la base de pruebas objetivas. Aunque el articulado de RIA no lo recoge, se admite la creación de un grupo de expertos que brinde asesoramiento y conocimientos técnicos específicos. Igualmente, y en atención a las previsiones reglamentarias, el Comité puede establecer subgrupos de trabajo que se encarguen de cuestiones concretas.

Además de las autoridades y representantes señalados, podrán participar -previa invitación- otras autoridades nacionales si los temas a tratar les resultan de relevancia, así como expertos u observadores externos. En el caso de Estados miembros que hayan establecido espacios controlados de prueba, las autoridades competentes en atención al reparto competencial interno han de coordinarse y cooperar con el Comité, lo que de suyo lleva su incorporación.

También puede realizar intercambios con terceros interesados, incluyendo otras instituciones comunitarias, en la medida que se considere necesario para el desarrollo de las actividades encomendadas. En cualquier caso, su funcionamiento concreto se ajustará al reglamento interno que ha de aprobarse siguiendo los lineamientos recogidos por reglamento y en ejercicio de las potestades domésticas que se le reconocen.

El Comité es, fundamentalmente, un órgano de asesoramiento, asistencia y consulta técnica de la Comisión y las autoridades nacionales de supervisión. Tiene como fin último facilitar la aplicación armonizada y coherente del Reglamento, así como la cooperación.

Al aglutinar autoridades nacionales y comunitarias, está llamada a brindar orientaciones sobre problemas emergentes de la IA que afecten al mercado único. En este sentido, entendemos, representa el foro adecuado para determinar el riesgo admisible de los avances de estas tecnologías para la sociedad europea.

Por otra parte, la búsqueda de una aplicación armonizada, aunque respetando el margen competencial de los Estados miembros, exige la transmisión de información. Al tratarse además de un sector en permanente evolución, esa información ha de alcanzar no sólo a episodios de riesgo o incidentes, sino también cuestionamientos y perspectivas acerca de la valoración técnica de los riesgos de la IA, e incluso las diferentes sensibilidades sociales esperables ante la diversidad del espacio comunitario. De ahí que corresponda a este ente actuar como nodo que recopile y ponga a disposición de los Estados miembros las mejores prácticas y conocimientos técnicos, lo que contribuirá a una aplicación armonizada de la normativa.

Se le atribuye, además, en el marco de su actividad asesora emitir dictámenes, recomendaciones y directrices, así como informes sobre cuestiones relacionadas con RIA. Esta labor es especialmente relevante, y se atribuye de manera expresa respecto a los requisitos horizontales de los sistemas de alto riesgo, así como las especificaciones técnicas, normas armonizadas y, de ser el caso, especificaciones comunes.

El reglamento recoge normas de intensidad regulatoria reducida e incluso escasa. Por tanto, es esperable que corresponda al comité -al menos parcialmente- contribuir a completar el programa normativo sectorial. Por tanto, explícitamente se le atribuye la aprobación de normas de Soft Law, con el fin de contribuir a uniformar prácticas administrativas de los Estados miembros. En este sentido, se hace referencia a dos tipos de normas. Por una parte, las dirigidas a complementar las previsiones acerca del funcionamiento de espacios controlados de prueba para la IA, cuestión que exige que los Estados miembros cooperen intensamente con miras a que estos *sandboxes* funcionen de manera coordinada. En relación con esto, los Estados deben presentar al Comité y a la Comisión informes anuales sobre su funcionamiento, buenas prácticas, experiencias, recomendaciones, y aplicación de la normativa comunitaria relacionada con estos espacios de experimentación. Por otra parte, y debido a que buena parte

del régimen sancionador se deja en manos de los Estados miembros, corresponde al Comité aprobar las directrices para fijar dicho régimen, así como otras medidas necesarias para la aplicación efectiva del Reglamento en esta materia.

Finalmente, a lo largo del articulado se recogen otra serie de funciones y atribuciones del Comité relacionadas también con el asesoramiento y asistencia. Así, se le atribuye la propuesta de recomendaciones a la Comisión acerca de los informes anuales de los Estados miembros respecto a la idoneidad, capacidad financiera y recursos humanos de las autoridades nacionales competentes. Por otra parte, es competente para fomentar y facilitar, junto con la Comisión, la elaboración y adopción de códigos de conducta voluntarios dirigidos al cumplimiento de objetivos sociales.

4.2.- Autoridades nacionales. Designación y competencias

Junto a la Comisión, los Estados miembros desempeñarán un papel protagonista en la aplicación de marco comunitario, según se desprende de RIA. A éstos corresponde, según el caso, designar una o un grupo de autoridades nacionales, a las que se atribuye el ejercicio de competencias de supervisión y ejecución.

Se hace referencia a autoridades nacionales competentes y a éstas se atribuyen distintas funciones. No obstante, se deja en manos de los Estados miembros, en ejercicio de sus potestades domésticas, bien la designación de una única autoridad que agrupe dichas atribuciones, bien dos o más si lo consideran necesario por razones organizativas internas. En cualquier caso, ya se trate de una o varias autoridades nacionales, debe garantizarse que actúen con imparcialidad y objetividad. En línea con esto, cabe destacar, que no se exige que se trate de autoridades independientes como ocurre, por señalar un ejemplo, en materia de protección de datos. Puede tratarse, en nuestro caso, de alguna de las AAPP recogidas por la LRJSP, siempre que ajuste su actuación a los principios señalados.

En cambio, para garantizar su autonomía si exige que cuenten de manera permanente con recursos financieros y humanos. Respecto a estos últimos, la elección de sus miembros ha de orientarse por la idoneidad para el ejercicio de sus atribuciones con miras a asegurar

el acierto de sus decisiones. A tales efectos, debe tenerse en consideración la experiencia y conocimientos técnicos sobre tecnologías de inteligencia artificial y ciencia de datos, así como en cuestiones legales que permitan valorar los riesgos para los derechos fundamentales, la seguridad o la salud de los ciudadanos.

El cumplimiento de los extremos señalados, por otra parte, se somete a control de la Comisión. Por ello, los Estados miembros han de remitir un informe que se transmitirá al Comité Europeo de IA para debate y formulación de recomendaciones. Con esto se pretende fomentar una actuación armonizada, así como una coordinación incremental. Se recurre, así, a conocidas fórmulas de gobernanza por indicadores⁴⁷, que permiten la evaluación por parte de la Comisión, el Supervisor Europeo de Protección de Datos y las autoridades nacionales de supervisión, buscando el ajuste a las prescripciones comunitarias de manera preventiva mediante la presión y revisión por pares⁴⁸. Entendemos, pues, que se apuesta de manera primaria por fórmulas de control suave (*soft governance*) y técnicas relacionadas como el *namings and shaming*, conocidas en otros sectores donde existen afectaciones de derechos fundamentales, o en otros entornos tecnológicos de riesgo⁴⁹.

RIA diferencia entre distintas autoridades nacionales, con independencia de la decisión organizativa adoptada por los Estados miembros que, en todo caso, han de comunicar a la Comisión. Así, se recoge una categoría general, la autoridad nacional competente, que puede englobar otras tantas: autoridad nacional de supervisión, autoridad notificante y autoridad de vigilancia del mercado. Como se ha señalado, se admite la concentración de todas estas autoridades en una única, aunque desplegarán en todo caso las distintas funciones previstas. Cabe mencionar que esta fórmula no es novedosa en

⁴⁷ Acerca de la gobernanza por indicadores véase, entre otros, DAVIS, K., KINGSBURY, B. y MERRY, S. (2010): "Indicators as a Technology of Global Governance", *IILJ Working Paper* 2010/2.

⁴⁸ Sobre el *Peer Review* y el *Peer Pressure* como técnica de gobernanza, véase BALDWIN, R., *et al.* (2012): *Understanding Regulation. Theory, Strategy, and Practice*, segunda edición, Oxford University Press, Nueva York, p. 431.

⁴⁹ FINNEMORE, M. y HOLLIS, D. B. (2020): "Beyond Naming and Shaming: Accusations and International Law in Cybersecurity", *European Journal of International Law*, Vol. 31, N.º 3, pp. 969 y ss.

el ámbito de seguridad de los productos. Por el contrario, está en plena sintonía con la normativa sectorial comunitaria, aunque sujeta a variaciones en atención al contexto técnico o regulatorio. Así, y por destacar sólo un ejemplo, el Reglamento sobre drones⁵⁰, establece la obligación de designar una autoridad de vigilancia del mercado, así como autoridades notificantes. Como es evidente, cada Estado miembro hará la designación teniendo en consideración su reparto competencial interno.

Ahora bien, con independencia de la fórmula que elija cada Estado miembro, se recogen una serie de funciones y atribuciones generales. Atendido a la elección organizativa interna, las autoridades nacionales correspondientes estarán facultadas para brindar orientaciones y asesoramiento, esto es, dictar normas de *soft law* referidas a la aplicación de RIA, y especialmente dirigidas a proveedores a pequeña escala. Por otra parte, respecto a sistemas de alto riesgo, se les atribuye competencia para supervisar el cumplimiento de los requisitos horizontales exigidos por la normativa.

Nótese que la estructura institucional, no se agota con las típicas del ámbito de la calidad industrial, pues se establece el mandato de designar una autoridad nacional de supervisión, así como mecanismos de reparto competencial con autoridades sectoriales en materia de protección de datos. No obstante, estos entes se relacionan intensamente atendiendo a la infraestructura de seguridad de los productos creada por la normativa comunitaria, constituyéndose un triple círculo de protección para los ciudadanos.

Así, la aplicación y cumplimiento de la cascada regulatoria a la que hicimos referencia, que guía el diseño, modelado, puesta en funcionamiento y retirada de los sistemas de IA por parte de los proveedores (primer círculo), puede ser verificada por los organismos notificados mediante los procedimientos establecidos por la normativa comunitaria e interna. Por su parte, estos órganos notificados han de estar previamente acreditados por los organismos notificantes, a los que corresponde comprobar que cuentan con los requisitos técnicos

⁵⁰ Reglamento delegado (UE) 2019/945 de la Comisión, de 12 de marzo de 2019, sobre los sistemas de aeronaves no tripuladas y los operadores de terceros países de sistemas de aeronaves no tripuladas.

para realizar su labor, y garantizar que los productos se ajusten a los requerimientos de seguridad para la salud, el medio ambiente y los derechos fundamentales, antes de ponerse en el mercado (segundo círculo). Finalmente, una vez en el mercado, ha de garantizarse que los sistemas en funcionamiento no entrañen riesgos inadmisibles. A tales efectos, se les otorga competencias para supervisarlos *ex-post*, mediante procedimientos también recogidos por la normativa (tercer círculo). Todas estas funciones se atribuyen a la autoridad de vigilancia del mercado. Dicho esto, pasaremos a hacer una apretada referencia a las atribuciones previstas para cada una de las instituciones comentadas, aunque los organismos notificados serán abordados separadamente.

Pues bien, la autoridad nacional de supervisión se define como aquella designada para ejecutar y aplicar el Reglamento, representar al Estado miembro ante el Comité Europeo de IA, y coordinar todas aquellas atribuciones y funciones de competencia interna, amén de actuar como punto de contacto único con la Comisión. Por su parte, la autoridad notificante es la responsable de establecer y aplicar los procedimientos relacionados con la evaluación, designación y notificación de los organismos de evaluación de la conformidad, y hacerles seguimiento, es decir, la que asume las funciones más estrechamente relacionadas con los procedimientos de evaluación que abordaremos más adelante; mientras que a la autoridad de vigilancia del mercado, se atribuyen funciones y la adopción de medidas relacionadas con la aplicación del Reglamento 2019/1020 relativo a la vigilancia del mercado y la conformidad de los productos⁵¹. El espíritu de la norma es que se nombre una autoridad nacional de supervisión, que actúe como autoridad notificante y de vigilancia del mercado. No obstante, como se ha señalado se permite que, en ejercicio de potestades domésticas, mediando comunicación a la Comisión, los Estados miembros elijan la fórmula que consideren más adecuada.

Por su parte, a la autoridad nacional de supervisión, como se ha señalado, compete la supervisión y ejecución del Reglamento, además

⁵¹ Reglamento (UE) 2019/1020, de 20 de junio de 2019 relativo a la vigilancia del mercado y la conformidad de los productos. Sobre este Reglamento ya se ha pronunciado la doctrina. Entre otros, ÁLVAREZ GARCÍA, V. (2020): *op.cit.*, pp. 48 y 53.

de actuar -según prevén los considerandos- como punto de contacto oficial con el público, el resto de Estados miembros y la Comisión. En este sentido, ha de informar a la Comisión, de manera periódica, sobre las actividades relacionadas con la supervisión del mercado, y de manera peculiar sobre cualquier información de potencial interés a efectos de la aplicación del Derecho de la competencia. En este último supuesto, dicha información ha de comunicarse -de ser el caso- a las autoridades nacionales competentes en dicha materia. Respecto a los ciudadanos, es medular la atribución de servir como nodo de información y control del alcance de ésta, especialmente ante supuestos de riesgos para la seguridad, salud o derechos fundamentales de los ciudadanos. Es una cuestión cuyos límites, a nivel interno, han de acompañarse en atención a la nutrida jurisprudencia relacionada con el derecho a ser informado por la administración pública, así como por los mandatos recogidos por la LTBG. Se trata de una cuestión que, cabe esperar, generará ciertas fricciones y controversias a efectos de no afectar otros intereses y derechos también legítimamente protegidos⁵², tal como ya ha puesto en evidencia jurisprudencia reciente⁵³.

⁵² Cabe traer a colación aquí el ejemplo de la información relacionada con los medicamentos, sector cuya regulación también se basa en el riesgo y el fomento de la innovación. La delimitación del acceso a la información a los usuarios a cuestiones diversas más allá de los mandatos recogidos por la normativa sectorial, ha sido depurado por las resoluciones del Consejo de Transparencia y Buen Gobierno, así como por la jurisprudencia dictada al hilo de ésta. Se ha ocupado de esta cuestión GUICHOT, E. (2020): “Capítulo tercero. Límites”, en Guichot, E. y Barrero Rodríguez, C., *El derecho de acceso a la información pública*, Tirant lo blanch, Valencia, pp. 369 y ss. No obstante, se trata de una cuestión compleja que exige un elevado aquilatamiento, tal como expone CERRILLO i MARTÍNEZ, A. (2016): “Transparencia pública de las entidades privadas y confidencialidad de los datos empresariales”, en Canals Ametlers, D. (ed.) *Datos. Protección, transparencia y buena regulación*, Documentación Universitaria, Girona, pp. 87 y ss.

⁵³ Ejemplo de esto es la SJCCA n° 143/2021, de 30 de diciembre de 2021, denegando el acceso al código fuente de BOSCO, herramienta utilizada por el Gobierno y empresas eléctricas para determinar si usuarios vulnerables podrían ser acreedores del bono social. Aunque no está claro que la herramienta BOSCO utilizara algoritmos de IA, dado -precisamente- que no fue posible acceder al mencionado código fuente, sirve para poner en evidencia las tensiones entre derechos e intereses de sistemas de IA, especialmente cuando son utilizados por el sector público.

Respecto a la autoridad de vigilancia del mercado, en línea con la definición recogida por el Reglamento, se le atribuyen funciones de vigilancia y control de los sistemas en materia de seguridad, y se les otorga competencias para supervisar los sistemas de IA diseñados como componentes de seguridad de otros productos sujetos a la normativa “nuevo enfoque”. Como competencia específica se les designa responsables de la supervisión del sector Fintech, correspondiéndoles la vigilancia de las entidades financieras que introduzcan, pongan en servicio o utilicen sistemas de IA, con arreglo a legislación comunitaria sobre servicios financieros.

Los poderes atribuidos son intensos. Pueden solicitar acceso a datos de diverso tipo, incluyendo los utilizados para entrenar y validar modelos, valiéndose incluso de medios técnicos que permitan el acceso a distancia, como APIs e interfaces de programación. Por otra parte, previa solicitud y mediando justificación razonable, pueden acceder al código fuente de sistemas de alto riesgo, con la finalidad de evaluar el cumplimiento de los requisitos horizontales comentados. Como cabe concluir, estos poderes –en todo caso- han de ejercerse con estricta sujeción a deberes de confidencialidad y sigilo, puesto que pueden afectar derechos e interés comerciales, así como la protección de la propiedad intelectual de los proveedores y demás participantes afectados.

Finalmente, se les atribuye competencias vinculadas a algunos procedimientos de evaluación de la conformidad, así como respecto a los procedimientos y medidas post-comercialización de vigilancia del mercado. Así, les corresponde autorizar la introducción de sistemas de IA, de manera excepcional y por circunstancias tasadas, mediante la exención del procedimiento de evaluación de la conformidad; constatar, valorar y adoptar medidas correctoras en el marco de procedimientos de salvaguarda y de notificación de incidentes graves y fallos de funcionamiento.

Por su parte, las autoridades notificantes son las autoridades que designan los Estados miembros con el fin de establecer y aplicar los procedimientos necesarios para la evaluación, notificación y supervisión de los organismos de apoyo que comentaremos más abajo, esto es, los organismos notificados. Fundamentalmente se les atribuyen competencias relacionadas con la conformidad de los sistemas, dejan-

do a salvo lo mencionado respecto a las competencias de la autoridad de vigilancia del mercado. Así, estas autoridades se encargarán de establecer los procedimientos para la evaluación, designación y supervisión de los organismos de evaluación de la conformidad, es decir, de los organismos notificados. Aunque la competencia para regular estos procedimientos se cede a dichas autoridades, interesa notar que la propuesta recoge algunas coordenadas que guiaran necesariamente su aprobación: han de ser proporcionados, evitar cargas innecesarias a los proveedores en aplicación del principio de innovación, y tener en consideración circunstancias como el tamaño de las empresas y el sistema de IA específico. Se trata, no obstante, de mandatos ya recogidos por la normativa española⁵⁴ que somete los procedimientos administrativos al principio de proporcionalidad, que actúa como una suerte de principio-director reforzando la exigencia de simplificación⁵⁵ y limita la imposición de cargas administrativas en el ejercicio de sus competencias. Además, en línea con las tendencias y principios de buena regulación⁵⁶, somete toda esta cascada regulatoria a evaluación *ex-ante* y *ex-post* periódica con el fin de mejorar su eficacia y corregir desviaciones respecto a los fines perseguidos⁵⁷. En definitiva, los mandatos comunitarios encuentran ya un claro dique protector en la normativa interna.

⁵⁴ Recogidos en su momento por el capítulo I del Título primero de la Ley 2/2011, de 4 de marzo, de Economía Sostenible, y actualmente por el art. 129 de la LPAC.

⁵⁵ GAMERO CASADO, E. (2014): “La simplificación del procedimiento administrativo como categoría jurídica”, en Gamero Casado, E. (Coord.), *Simplificación del procedimiento administrativo y mejora de la regulación. Una metodología para la eficacia y el derecho a la buena administración*, Tirant lo blanch, Valencia, pp. 39.

⁵⁶ Sobre el significado y consecuencias prácticas de los principios de buena regulación, PONCE SOLÉ, J. (2019B): *Mejora de la regulación, lobbies y huella normativa: un análisis empírico de la informalidad*, Tirant lo blanch, Valencia, pp. 87-94; GAMERO CASADO, E. (2014), *op.cit.*, pp. 76 y ss.

⁵⁷ VAQUER CABALLERÍA, M. (2017): “Auge y problemas de la metarregulación: La iniciativa legislativa y la potestad reglamentaria en la Ley de Procedimiento Administrativo Común”, en Ponce Solé, J. y Cerrillo i Martínez, A. (coord.), *Innovación en el ámbito del buen gobierno: ciencias del comportamiento, transparencia y prevención de la corrupción*, INAP, Madrid, p. 179.

En el marco de la normativa de seguridad de productos, la designación de estas autoridades no se limita a autoridades de rango específico, y a nivel interno se ha venido recurriendo a distintos órganos superiores y directivos⁵⁸. Por otra parte, y en sana lógica, RIA permite aprovechar las estructuras institucionales creadas a partir de la aplicación de la normativa sectorial. No se exige que todas las competencias estén atribuidas a una misma autoridad. De hecho, a nivel interno es posible encontrar regulación sectorial sobre seguridad de los productos en la que se encomiende a diferentes autoridades el establecimiento de los procedimientos y la evaluación o supervisión de los organismos de evaluación de la conformidad, ya sea de mediante asignación de competencias⁵⁹ o delegación⁶⁰.

Cabría esperar que en el caso del RIA se acoja una solución similar, a pesar de la preferencia comunitaria por acumular funciones. Así, podría nombrarse como autoridad notificante a efectos de establecer los procedimientos, notificar a nivel comunitario los organismos notificados, así como a efectos de la exigencia de presentación de declaración responsable en atención a la normativa interna⁶¹, al Ministerio de Asuntos Económicos y Transformación digital, que podría delegar competencias en la SEDIA⁶². Para el resto de funciones, esto es, la evaluación y supervisión de los organismos notificados, podría recurrirse

⁵⁸ Véase el sistema de información Nando (*New Approach Notified and Designated Organisations*).

⁵⁹ Así, en el caso de los productos fertilizantes, de conformidad con el art. 3 del Real Decreto 568/2020, de 16 de junio, por el que se designa a la autoridad notificante, y se determinan otras obligaciones, en cumplimiento del Reglamento (UE) 2019/1009.

⁶⁰ Tal es el caso del anexo II del Real Decreto 989/2015, de 30 de octubre, por el que se aprueba el Reglamento de artículos pirotécnicos y cartuchería.

⁶¹ Cabe recordar que, además de la necesaria acreditación que comentaremos seguidamente, para que los organismos notificados (u organismos de control según la denominación interna), pueden ejercer sus funciones deben presentar tal autorización, a tenor de lo previsto por el artículo 43 del Real Decreto 2200/1995, de 28 de diciembre, por el que se aprueba el Reglamento de la Infraestructura para la Calidad y la Seguridad Industrial (RICSI).

⁶² A semejanza de la solución establecida respecto a la comercialización, puesta en uso y vigilancia del mercado de equipos radioeléctricos por el Real Decreto 188/2016, de 6 de mayo, por el que se aprueba el Reglamento por el que se establecen los requisitos para la comercialización, puesta en servicio y uso de equipos radioeléctricos, y se regula el procedimiento para la evaluación de la

a la designación de autoridades nacionales que ejerzan funciones de acreditación.

Atendiendo a la situación española, cabría atribuir esas competencias al organismo nacional de acreditación designado en aplicación del Reglamento 765/2008⁶³, la Entidad Nacional de Acreditación (ENAC)⁶⁴, entidad colaboradora de la administración pública⁶⁵, solución que resulta coherente con el mandato de designación única⁶⁶. Cabe recordar que, según el marco comunitario e interno de seguridad de los productos, la acreditación se dirige a verificar y reconocer formalmente la capacidad técnica de los organismos notificados, que se encargarán de evaluar la conformidad de los sistemas y modelos de IA con la normativa de seguridad⁶⁷. Por tanto, la entidad de acreditación vendría a ejercer, en último término, una función autorizadora, emitiendo un acto de naturaleza pública que se recoge además en un documento público (certificado de acreditación), aunque no se trate de un verdadero acto administrativo⁶⁸.

De la propuesta normativa pueden extraerse una serie de principios de actuación que han de regir la actuación de estas autoridades, y que no hacen más que recoger los mandatos establecidos por la

conformidad, la vigilancia del mercado y el régimen sancionador de los equipos de telecomunicación.

⁶³ Reglamento (CE) 765/2008 del Parlamento Europeo y del Consejo, de 9 de julio de 2008, por el que se establecen los requisitos de acreditación y vigilancia del mercado relativos a la comercialización de los productos.

⁶⁴ Sobre el régimen jurídico de este organismo, al que se otorgan potestades públicas, han de carecer de fines lucrativos y cumplir los requisitos previstos por el art. 8 del Reglamento (CE) 765/2008, véase ÁLVAREZ GARCÍA, V. (2010): *op.cit.*, pp. 305-309; Izquierdo Carrasco, M. (2000): *La seguridad de los productos industriales*, Marcial Pons, Madrid, pp. 509-511.

⁶⁵ CUETO PÉREZ, M. (2008): *Procedimiento administrativo, sujetos privados y funciones públicas*, Thomson Civitas, Cizur Menor, pp. 163 y ss.

⁶⁶ Esto, de conformidad con el artículo 4.1 del Reglamento (CE) 765/2008. Cabe destacar en este sentido, que este mandato ha generado tensiones respecto al reparto competencial interno con las comunidades autónomas que, según la doctrina, han de resolverse atendiendo a la primacía del Derecho comunitario. Véase en este sentido, ÁLVAREZ GARCÍA, V. (2015): “La recentralización de competencias en el ámbito de la designación de las entidades de acreditación de los verificadores medioambientales”, *RGDA*, N.º 40, pp. 38 y ss.

⁶⁷ ÁLVAREZ GARCÍA, V. (2020): *op.cit.*, p. 38.

⁶⁸ CUETO PÉREZ, M. (2008), *op.cit.*, pp. 166-167.

normativa general⁶⁹. Así, han de ser imparciales y objetivas, evitar conflictos de interés tanto respecto a las actividades como a los participantes en los procesos de evaluación, y garantizar la competencia personal y técnica al llevar a cabo sus funciones. Su régimen, como es menester, se completa con obligaciones de confidencialidad y sigilo acerca de la información a la que accedan; la prohibición de consecución de ánimo lucrativo; así como el principio de no competencia con los organismos notificados, lo que impide ofertar -por ejemplo- servicios de consultoría que compitan con dichos entes.

4.3.- Organismos notificados y organismos de evaluación de la conformidad

En línea con el sistema comunitario de seguridad de productos, RIA establece la participación de organismos de evaluación de la conformidad y organismos notificados. Como se avanzó, no se trata de figuras de novedosa creación, sino habitualmente de entes colaboradores de las AAPP sobre los que la doctrina ha dado buena cuenta⁷⁰. El Reglamento define a los primeros como aquellos organismos independientes que desempeñan actividades de evaluación de la conformidad, entre las que figuran la prueba, la certificación y la inspección⁷¹. Se trata de la misma definición que recoge el art. 2.16 del Reglamento de acreditación y vigilancia del mercado relativo a la comercialización de productos⁷². Por su parte, se define como Organismo notificado todo organismo de evaluación de conformidad designado con arreglo al Reglamento o por otra legislación armonizada. De esto se desprende, que los organismos

⁶⁹ Art. 4 del Reglamento (CE) 765/2008; artículo 17 de la Ley 21/1992, de 16 de julio, de Industria.

⁷⁰ Véanse, entre otros, los trabajos citados de ÁLVAREZ GARCÍA, ya citados.

⁷¹ La ley 21/1992, de 16 de julio, de Industria en el ámbito español hace referencia a organismos de control, que define como “aquellas personas físicas o jurídicas que teniendo capacidad de obrar y disponiendo de los medios técnicos, materiales y humanos e imparcialidad e independencia necesarias, pueden verificar el cumplimiento de las condiciones y requisitos de seguridad establecidos en los Reglamentos de Seguridad para los productos e instalaciones industriales”.

⁷² Reglamento (CE) 765/2008 del Parlamento Europeo y del Consejo de 9 de julio de 2008 por el que se establecen los requisitos de acreditación y vigilancia del mercado relativos a la comercialización de los productos.

de evaluación han de solicitar y seguir un procedimiento a efectos de poder verificar los sistemas de IA, demostrando previamente el cumplimiento de los requisitos recogidos por la normativa⁷³. Esto es, deben ser previamente acreditados, tal como se señaló previamente, y además ser habilitados mediando declaración responsable ante los organismos notificantes⁷⁴. Además, en cumplimiento de la normativa de normalización, es exigible el cumplimiento de las normas “UNE-EN ISO/IEC 17065:2012, sobre evaluación de la conformidad. Requisitos para organismos que certifican productos, procesos y servicios”, así como los previstos por la normativa interna⁷⁵.

La designación de los organismos notificados se comunicará a la Comisión y a los Estados miembros haciendo uso de las plataformas que a tales efectos se pongan en funcionamiento, dejando constancia de las actividades de evaluación de conformidad que podrán realizar y las tecnologías de IA respecto de las que acreditó capacidad para evaluar. A cada organismo notificado se le asignará un número de identificación por parte de la Comisión, que publicará un listado actualizado de éstos, así como las actividades para las que han sido notificados.

Corresponde a la Comisión garantizar la coordinación y cooperación de todos los organismos notificados, para lo que ha de poner en marcha un grupo sectorial de trabajo de ámbito comunitario. Ahora, debido a que es previsible que en cada Estado miembro se designen múltiples organismos notificados, deberán articularse fórmulas para estructurar la participación de éstos en el mecanismo de coordinación y cooperación comunitario. Cabe aventurar que corresponderá al Ministerio de Asuntos Económicos y Transformación digital designar, por un período de tiempo determinado, a un organismo notificado que represente a este colectivo ante el citado mecanismo.

El Reglamento es detallado respecto a la funciones y requisitos que han de cumplir los organismos notificados. También acerca de los procedimientos en caso de incumplimiento de los requisitos u obligaciones que les son exigibles. En cuanto a las funciones, la fundamental

⁷³ Recogidos en el art. 33 de RIA.

⁷⁴ Esto, en concordancia con el art. 15.4 de la Ley de industria.

⁷⁵ Véase, art. 41 del RICSI.

es la verificación de la conformidad de los sistemas de alto riesgo mediante el procedimiento de evaluación de conformidad, función que han de desempeñar teniendo en cuenta el tamaño de las empresas, el sector en que operan, así como la estructura y complejidad del sistema. En cuanto a los requisitos, se exige que cuenten con una estructura de entidad suficiente para desempeñar la función señalada, garantizando su independencia respecto a los proveedores de sistemas de IA y demás operadores con intereses económicos en el ámbito de su actividad, así como la confidencialidad, objetividad e imparcialidad. Dicha organización debe, en todo caso, asegurar la confianza respecto a las evaluaciones de conformidad, por lo que además se les exige integridad profesional y competencia técnica demostrada. En este sentido, han de contar con personal científico y técnico que acredite conocimientos en materia de IA, datos, ciencia de datos, así como de los requisitos horizontales exigidos por el Reglamento. Toda la documentación que acredite el cumplimiento de estos requisitos, así como la suscripción de un seguro de responsabilidad adecuada para dar cobertura a las funciones previstas, han de ponerse a disposición de la autoridad notificante en caso de que les sea solicitada.

Ahora bien, dado que es posible que algún organismo notificado se vea imposibilitado o no demuestre adecuado cumplimiento de los requisitos señalados, se faculta a las autoridades notificantes para investigar los incumplimientos, siguiendo para ello los lineamientos generales recogidos en la propuesta⁷⁶, y que pueden derivar en la restricción, suspensión o retirada de la notificación atendiendo a la gravedad del caso. En tales supuestos, se ha de informar a la Comisión y a los Estados miembros. Idéntica atribución se reconoce a la Comisión que, en caso de duda respecto al cumplimiento de los requisitos por un organismo notificado, puede iniciar un procedimiento a tenor de lo previsto en el artículo 5 del Reglamento (UE) n.º 182/2011. En estos supuestos las autoridades notificantes han de colaborar con la Comisión, transmitiendo toda la información que le sea solicitada respecto a dicho organismos y, en caso de comprobar el incumplimiento, adoptará una decisión motivada que solicitará al Estado miembro correspondiente la retirada de la notificación.

⁷⁶ Art. 36 de RIA.

Aunque nuestro trabajo se centra fundamentalmente en la organización y estructuras de gobernanza, cabe recordar que el modelo comunitario recoge distintos procedimientos dirigidos a asegurar la seguridad de los sistemas de IA, reduciendo a niveles aceptables los riesgos para la salud y derechos fundamentales de las personas. Estos procedimientos se desarrollan, bajo el principio de cooperación que vertebra el sistema europeo de administración compartida⁷⁷, donde los Estados miembros cuentan con diversas competencias, y pueden dar lugar a variopintas fórmulas decisorias o de intercambio de información, incluyendo supuestos de colaboración privada para el ejercicio de funciones públicas, y donde es habitual que se reserve a la Comisión facultades decisorias determinantes, así como la presencia de actos administrativos transnacionales⁷⁸.

BIBLIOGRAFÍA CITADA

- ÁLVAREZ GARCÍA, V. (1999): *La normalización industrial*, Tirant lo blanch, Valencia.
- (2010): *Derecho de la regulación económica. VIII. Industria*, Iustel, Madrid.
- (2015): “La recentralización de competencias en el ámbito de la designación de las entidades de acreditación de los verificadores medioambientales”, *RGDA*, N.º 40.
- (2020): *Las normas técnicas armonizadas (una peculiar fuente del Derecho europeo)*, Iustel, Madrid.
- BALDWIN, R., *et al.* (2012): *Understanding Regulation. Theory, Strategy, and Practice*, segunda edición, Oxford University Press, Nueva York.
- BANICEVIC, A., *et al.* (2018): “Algorithms: Challenges and opportunities for antitrust compliance”, ABA Section of Antitrust Law’s Compliance and Ethics Committee. Fall 2018, special report.
- BECK, U. (2006): *La sociedad del riesgo. Hacia una nueva modernidad*, Paidós, Barcelona.

⁷⁷ SCHMIDT-ASSMANN E. (2021), *op.cit.*, pp. 178-179.

⁷⁸ Esto es, actos administrativos de autoridades nacionales que llegan a desplegar sus efectos jurídicos en todo el territorio de la Unión. Interesa destacar que estos actos, a pesar del alcance de sus efectos se rigen por el Derecho nacional, por lo que se aplica la construcción jurídica interna. Al respecto, SCHMIDT-ASSMANN, E. (2021), *op.cit.*, pp. 180-181.

- CASTRO, D. y McLAUGHLIN, M. (2019): *Ten Ways the Precautionary Principle Undermines Progress in Artificial Intelligence*, ITIF.
- CERRILLO i MARTÍNEZ, A. (2016): “Transparencia pública de las entidades privadas y confidencialidad de los datos empresariales, en Canals Ametlers, D. (ed.) *Datos. Protección, transparencia y buena regulación*, Documentación Universitaria, Girona.
- (2019): “El impacto de la inteligencia artificial en el Derecho administrativo ¿Nuevos conceptos para nuevas realidades técnicas?”, *RGDA*, N.º 50.
- (2020): “¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?”, *European Review of Digital Administration & Law – Erdal*, Vol. 1, N.º 1-2.
- COTINO HUESO, L. *et al.* (2021): “Un análisis crítico constructivo de la Propuesta de Reglamento de la Unión Europea por el que se establecen normas armonizadas sobre la Inteligencia Artificial (Artificial Intelligence Act)”, *Diario La Ley*.
- (2019): “Riesgos e impactos del big data, la inteligencia artificial y la robótica. Enfoques, modelos y principios de la respuesta del Derecho”, *RGDA*, N.º 50.
- CUETO PÉREZ, M. (2008): *Procedimiento administrativo, sujetos privados y funciones públicas*, Thomson Civitas, Cizur Menor.
- DARNACULLETA i GARDELLA, M. M. (2004): *Autorregulación y Derecho público: La autorregulación regulada*, Marcial Pons, Madrid.
- DAVIS, K., KINGSBURY, B. y MERRY, S. (2010): “Indicators as a Technology of Global Governance”, IILJ Working Paper, 2010/2.
- EPSC (2016): “Towards an Innovation Principle Endorsed by Better Regulation”, EPSC Strategic Notes, Comisión Europea, N.º 14, 30 de junio.
- ESTEVE PARDO, J. (2021), *Principios de Derecho regulatorio. Servicios económicos de interés general y regulación de riesgos*, Marcial Pons, Madrid.
- European Group on Ethics in Science and New Technologies – EGE (2018), *Statement on Artificial Intelligence, Robotics and ‘Autonomous’ Systems*, Comisión Europea.
- EZRACHI, A. y STUCKE, M.E. (2016), *Virtual Competition. The promise and perils of the algorithm-driven economy*, Harvard University Press, Cambridge.
- (2020): “Sustainable and Unchallenged Algorithmic Tacit Collusion”, *Northwestern Journal of Technology & Intellectual Property*, Vol. 17, N.º 2.
- FINNEMORE, M. y HOLLIS, D. B. (2020): “Beyond Naming and Shaming: Accusations and International Law in Cybersecurity”, *European Journal of International Law*, Vol. 31, N.º 3.

- GAMERO CASADO, E. (2014): “La simplificación del procedimiento administrativo como categoría jurídica”, en Gamero Casado, E. (Coord.), *Simplificación del procedimiento administrativo y mejora de la regulación. Una metodología para la eficacia y el derecho a la buena administración*, Tirant lo blanch, Valencia.
- (2021): “El enfoque europeo de inteligencia artificial”, *Revista de Derecho Administrativo – CDA*, N.º 20, 2021.
- GARNETT, K., *et al.* (2018): “Towards an innovation principle: An industry trump or shortening the odds on environmental protection?”, *Law, Innovation and Technology*, Vol. 10, N.º 1.
- GUICHOT, E. (2020): “Capítulo tercero. Límites”, en Guichot, E. y Barrero Rodríguez, C., *El derecho de acceso a la información pública*, Tirant lo blanch, Valencia.
- GUTIÉRREZ ALONSO, J. J. (2009): “El elemento tecnológico: el principio de neutralidad y sus implicaciones en la regulación”, en Muñoz Machado, S. y Esteve Pardo, J. (Dirs.), *Derecho de la regulación económica. I. Fundamentos e instituciones de la regulación*, Iustel, Madrid.
- HERNÁNDEZ, J. C. (2011): *El régimen jurídico-administrativo de las concentraciones de empresas en el Derecho comunitario: Test sustantivo y control judicial*, Thomson-Reuters Aranzadi, Cizur Menor.
- HERNÁNDEZ PEÑA, J.C. (2022): *El marco jurídico de la inteligencia artificial. Principios, procedimientos y estructuras de gobernanza*, Thomson-Reuters Aranzadi, Cizur Menor.
- HILDEBRANDT, M. y KOOPS, B.-J. (2010): “The Challenges of Ambient Law and Legal Protection in the Profiling Era”, *The Modern Law Review*, Vol. 73, N.º 3.
- IZQUIERDO CARRASCO, M. (2000): *La seguridad de los productos industriales*, Marcial Pons, Madrid.
- MANTELERO, A. (2020): “Chapter III. Analysis of international legally binding instruments. Final report”, Council of Europe Study DGI, Towards regulation of AI systems, Consejo de Europa.
- MORENO REBATO, M. (2021): *Inteligencia artificial (Umbrales éticos, Derecho y administraciones públicas)*, Thomson Reuters Aranzadi, Cizur Menor.
- OECD (2017): *Algorithms and collusion. Competition Policy in the Digital Age*, OECD Publishing, Paris.
- PETTIT, N. (2018): “Emerging competition, innovation, and market structure issues around algorithms, artificial intelligence, and predictive analytics”, Submission to the FTC hearings on competition and consumer protection in the 21st century.
- PICHT, P.G. y LODERER, G.T. (2019): “Framing algorithms: Competition law and (other) regulatory tools”, *World Competition*, Vol. 42, N.º 3.

- PONCE SOLÉ, J. (2019A): “Inteligencia artificial, Derecho administrativo y reserva de humanidad: Algoritmos y procedimiento administrativo debido tecnológico”, *RGDA*, N.º 50.
- (2019B): *Mejora de la regulación, lobbies y huella normativa: un análisis empírico de la informalidad*, Tirant lo blanch, Valencia.
- REED, C. (2007): “Taking Sides on Technology Neutrality”, *SCRIPT-ed*, Vol. 4, N.º 3.
- REDA, A. (2019): *Artificial Intelligence. Ethics, governance and policy challenges*, Center for European Policy Studies (CESP), Bruselas.
- SCHIMDT-ASSMANN, E. (2021): *La dogmática del Derecho administrativo*, Global Law Press, Sevilla.
- VALERO TORRIJOS, J. (2019): Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración, *RCDP*, N.º 58.
- VAN DER HAAR, I. (2007): “Technological Neutrality. What Does It Entail?”, *TILEC Discussion Paper DP 2007-009*
- VAQUER CABALLERÍA, M. (2017): “Auge y problemas de la metarregulación: La iniciativa legislativa y la potestad reglamentaria en la Ley de Procedimiento Administrativo Común”, en Ponce Solé, J. y Cerrillo i Martínez, A. (coord.), *Innovación en el ámbito del buen gobierno: ciencias del comportamiento, transparencia y prevención de la corrupción*, INAP, Madrid.

Capítulo XIV

Organización y gobernanza de la inteligencia artificial en la Junta de Andalucía: la Estrategia Andaluza de IA y la Agencia Digital de Andalucía

MARÍA PÉREZ NARANJO

*Directora General de Estrategia Digital
Agencia Digital de Andalucía*

Resumen: La inteligencia artificial se ha convertido en un área de importancia estratégica para cualquier región, por su capacidad de ser vector de crecimiento económico y bienestar social. Andalucía es consciente de ello y cuenta con el liderazgo y la visión estratégica de una administración que se basa en la tecnología para mejorar la eficiencia de la prestación de servicios públicos, a la vez que impulsa la transformación digital de la región. Para afrontar los retos tecnológicos que se tienen por delante, se crea la Agencia Digital de Andalucía, una estructura unificada para dar respuesta a las necesidades de la era digital, a través de una hoja de ruta única alineada con las políticas europeas. La Agencia Digital de Andalucía trabaja ya en la puesta en marcha de un plan estratégico de inteligencia artificial para la región, junto a un grupo de expertos de reconocido prestigio en diversos ámbitos de la IA. El plan persigue generar sinergias entre los agentes clave del ecosistema andaluz: universidad, startups, centros de investigación, empresas tecnológicas, talento, inversión,...y una administración pública que tendrá un papel principal impulsando el ecosistema y mejorando las oportunidades de colaboración entre todos los agentes.

Abstract: AI has become an area of strategic importance for any region with the potential to be a key driver of economic development and social welfare. Andalusia is aware of this and has an administration that holds the leadership and has the strategic vision to improve public services through an intensive use of technology, while promoting the transformation of the region. To address the technological and innovation challenges ahead, the Digital Agency of Andalusia is created, a unified structure to respond to the needs of the digital era, through a unique digital roadmap aligned with European policies. The Digital Agency is already working on an strategic plan of AI in Andalusia, together with a group of renowned experts in different fields in AI. The plan seeks to generate synergies between the key agents of the andalusian ecosystem: universities, startups, ICT industry, SMEs, talent, investment, R&D centers,...and a public administration that has a main role fostering this ecosystem and enhancing collaboration opportunities.

Palabras-clave: Inteligencia artificial, administración pública, tecnología, Andalucía, Agencia Digital de Andalucía, estrategia, gobernanza, casos de uso.

Key-words: Artificial intelligence, public administration, technology, Andalusia, Digital Agency of Andalusia, strategy, governance, use cases.

SUMARIO.- 1. LA AGENCIA DIGITAL DE ANDALUCÍA.- 1.1.Contextualización.- 1.2. Un nuevo modelo de gestión TIC.- 1.3. Plan inicial de actuación.- 1.3.1. *Líneas y objetivos estratégicos*.- 1.3.2. *Cinco planes estratégicos*.- 2. CASOS DE USO DE INTELIGENCIA ARTIFICIAL EN LA JUNTA DE ANDALUCÍA.- 2.1. Automatización de procesos para la mejora en la gestión de procedimientos administrativos.- 2.2. Proyecto BioData.- 2.3. Robots conversacionales y asistentes virtuales.- 2.4. Otros casos de uso.- 3. ESTRATEGIA DE INTELIGENCIA ARTIFICIAL DE ANDALUCIA.- 3.1. La inteligencia artificial en Andalucía.- 3.2. Misión y Visión de la Estrategia.- 3.3. Comité de expertos.- 3.4. Metodología.- 3.5. Modelo de Gobernanza

1.- LA AGENCIA DIGITAL DE ANDALUCÍA

1.1.- Contextualización

Vivimos en un mundo tecnológico. No hay industria ni actividad diaria donde lo digital no esté presente o provocando grandes cambios. El crecimiento exponencial de Internet y la popularización de las tecnologías digitales sitúan a la ciudadanía, a las administraciones y a las empresas en el nuevo escenario de la digitalización. Se trata de una nueva realidad social y económica imparable, que no espera a nadie, ni tiene vuelta atrás, y que desde Andalucía se afronta como una oportunidad para generar riqueza y empleo, utilizando la tecnología como palanca para avanzar hacia un nuevo modelo productivo, más innovador y sostenible.

Es un momento histórico ilusionante a la vez que desafiante, en el que la renovada Estrategia Digital de la Unión Europea, alineada con el nuevo marco de financiación 2021-2027 y con los fondos de recuperación y resiliencia, presentan una oportunidad única para que Andalucía de un salto adelante en el tablero de la economía europea. Y para conseguirlo, es esencial contar con el firme liderazgo y una visión estratégica a largo plazo por parte de la Administración Pública que permita maximizar las capacidades y oportunidades que la digitalización y la innovación ofrecen al desarrollo de la economía regional.

En este contexto exigente y de profundos cambios, nace la Agencia Digital de Andalucía (ADA) con la firme ambición de hacer realidad la transformación digital de Andalucía.

1.2.- Un nuevo modelo de gestión TIC

En la Junta de Andalucía ha existido históricamente una descentralización y dispersión de las competencias y de los presupuestos de ámbito tecnológico entre los distintos organismos. Este modelo se ha revelado en los últimos tiempos como una fuente de ineficiencias, además de resultar poco flexible a la hora de cubrir las expectativas y necesidades de una ciudadanía ya inmersa en la era digital, que espera unos servicios públicos ágiles y una experiencia personalizada. La pandemia de la COVID-19 trajo consigo una demanda creciente de servicios tecnológicos por parte de los distintos departamentos de la Junta de Andalucía, que requerían de una puesta en marcha urgente y una racionalización de los costes ante la situación de crisis existente. Todo ello ha permitido evidenciar la necesidad de una visión y planificación común, y de una serie de cambios estructurales en la Administración Pública andaluza que permitan afrontar el desafío de hacer realidad la transformación digital de nuestra región.

De esta forma en 2020¹ el Gobierno de la Junta de Andalucía inicia un proceso de unificación y consolidación de competencias, cuyo hito más importante fue la creación de una agencia administrativa que integra todos los servicios de tecnologías de la información y las comunicaciones y gobierno abierto de la Administración de la Junta de Andalucía, horizontales y verticales, para dar respuesta a las necesidades específicas de las consejerías y agencias y que permita desarrollar las actuaciones de transformación digital en toda la organización, bajo criterios de racionalización y sostenibilidad económica.

Así, mediante la disposición adicional vigésima de la Ley 3/2020, de 28 de diciembre, del Presupuesto de la Comunidad Autónoma de Andalucía para el año 2021, se crea la Agencia Digital de Andalucía, como agencia administrativa de las previstas en el artículo 54.2.a) de la Ley 9/2007, de 22 de octubre, de la Administración de la Junta de Andalucía. La Agencia tiene personalidad jurídica pública diferenciada, plena capacidad jurídica y de obrar, patrimonio y tesorería propios, sin perjuicio del principio de unidad de caja establecido en el artículo 73.b) del Texto Refundido de la Ley General de la Hacienda

¹ Decreto del Presidente 3/2020, de 3 de septiembre, de la Vicepresidencia y sobre reestructuración de Consejerías.

Pública de la Junta de Andalucía, aprobado por Decreto Legislativo 1/2010, de 2 de marzo, así como autonomía de gestión para el cumplimiento de sus fines.

En 2022², la ADA adquirió también las competencias en materia de infraestructura de telecomunicaciones y economía digital, culminando así el proceso de centralización que se iniciara en 2020.

Con la creación y puesta en marcha de la ADA, la Junta de Andalucía avanza hacia una administración del siglo XXI, eficiente, abierta, transparente y sostenible, que contribuya a la mejora de calidad de vida de la ciudadanía y a la riqueza de la región. Se trata de una apuesta decidida por la digitalización de Andalucía, tal y como marca Europa: llevando la tecnología digital a las empresas y ciudadanos andaluces a través de una serie de inversiones estratégicas, y convirtiendo a la propia Administración Pública en motor de dicha transformación digital.

1.3.- Plan inicial de actuación

Por Orden de 3 de mayo de 2022 de la Consejería de Presidencia, Administración Pública e Interior³, se aprueba el plan inicial de actuación de la ADA. Se trata de una ambiciosa estrategia de transformación digital, cuyo fin último es mejorar la calidad de vida de la ciudadanía y contribuir a la reactivación económica a través de la digitalización de la sociedad y la administración andaluza.

Dicho plan inicial de actuación establece un marco temporal de cuatro años (2021-2024) y parte de un presupuesto inicial de 153M€ para el 2021, y más de 300M€/año a partir del 2022 y 2023, previendo una absorción progresiva de recursos financieros de las partidas presupuestarias TIC de organismos bajo el ámbito de la agencia⁴, hasta la total centralización del presupuesto TIC de la Junta de Andalucía.

² Decreto 152/2022, de 9 de agosto, por el que se establece la estructura orgánica de la Consejería de la Presidencia, Interior, Diálogo Social y Simplificación Administrativa.

³ Actual Consejería de Presidencia, Interior, Diálogo Social y Simplificación Administrativa.

⁴ Sin concretar.

Para la consecución de sus fines, la agencia cuenta con un capital humano de más de 1000 profesionales TIC procedentes de las distintas Consejerías, agencias administrativas y agencias de régimen especial.

1.3.1.- Líneas y objetivos estratégicos

El plan inicial de actuación actualmente abarca cinco grandes líneas estratégicas: sociedad digital; administración digital; ciberseguridad, infraestructuras y puesto de trabajo digital; digitalización de los servicios públicos sectoriales; y economía digital. A través de cada una de estas líneas de trabajo, se persigue alcanzar una serie de objetivos estratégicos para el posicionamiento de la Junta de Andalucía como administración de referencia, y como motor de la transformación digital de la región.

SOCIEDAD DIGITAL

- Ofrecer servicios más centrados en la ciudadanía
- Mejorar la relación digital con la ciudadanía mediante el uso de una plataforma unificada omnicanal.
- Maximizar el uso de los servicios puestos a disposición de la ciudadanía, superando las barreras existentes.
- Capacitar a la población andaluza en competencias digitales, disminuyendo la brecha digital.
- Mejorar los mecanismos necesarios para difundir información pública y permitir la participación de la ciudadanía en las políticas públicas.

ADMINISTRACIÓN DIGITAL

- Agilizar los procesos de negocio aplicando tecnologías de automatización inteligente.
- Mejorar la toma de decisiones a partir de la analítica avanzada del dato.
- Digitalizar de forma eficiente los servicios ofertados, consolidando la transformación de la Junta de Andalucía.

- Crear una cultura de innovación tecnológica centrada en el usuario para desarrollar ideas, retos, proyectos y programas innovadores.
- Impulsar la implantación y uso de la Administración electrónica.

CIBERSEGURIDAD, INFRAESTRUCTURAS Y PUESTO DE TRABAJO DIGITAL

- Aumentar la madurez en materia de ciberseguridad de la Junta de Andalucía y prestando servicio a las administraciones públicas de la Comunidad que así lo requieran, ofreciendo servicios públicos digitales seguros y resilientes.
- Ofrecer soluciones de comunicaciones innovadoras que incrementen la productividad de los empleados públicos y permitan a la Administración relacionarse con la ciudadanía.
- Consolidación de las infraestructuras TIC de la Junta, favoreciendo su evolución a modelos en “Cloud”.
- Implantar el puesto de trabajo digital.

DIGITALIZACIÓN DE LOS SERVICIOS PÚBLICOS SECTORIALES

- Transformar servicios con las oportunidades que brindan las tecnologías disruptivas.
- Utilización de los sistemas y plataformas corporativas y horizontales.
- Buscar sinergias entre los actuales servicios digitales y nuevas necesidades.
- Homogenizar y converger las diferentes soluciones tecnológicas, así como las metodologías y funciones utilizadas.
- Mejorar los tiempos, minimizar costes, tiempo y esfuerzos para la puesta en marcha de soluciones tecnológicas.

ECONOMÍA DIGITAL

- Aportar los recursos adecuados para la transformación digital en las Pymes.
- Capacitar a las empresas andaluzas en competencias digitales.

- Impulsar la igualdad de género en el sector TIC.
- Aumentar la competitividad de las empresas TIC andaluzas.
- Potenciar el emprendimiento digital en Andalucía.

1.3.2.- Siete planes estratégicos

El plan inicial de actuación se aborda a través del diseño, ejecución y evaluación de siete planes estratégicos, cada uno de los cuales está focalizado en un ámbito tecnológico identificado como prioritario para la transformación digital de la región.

En primer lugar, la Estrategia Andaluza de Ciberseguridad 2022-2025, cuya formulación fue aprobada por acuerdo de Consejo de Gobierno de 3 de agosto de 2021⁵, como instrumento general de planificación del Gobierno andaluz para las políticas de ciberseguridad, con el fin de proteger a la sociedad andaluza en su conjunto ante el cibercrimen; reforzar las capacidades de prevención, detección y respuesta a incidentes en la Administración de la Junta de Andalucía a través de servicios avanzados de ciberseguridad; cooperar y colaborar en aras a extender la capacidad de protección del conjunto de entidades del sector público andaluza, ayuntamientos, diputaciones y otros entes públicos; y además, aprovechar el alto potencial de crecimiento del sector de la ciberseguridad en Andalucía para contribuir a la creación de riqueza y de empleo en la región.

En segundo lugar, el Plan de Capacitación Digital de Andalucía 2022-2025, cuya formulación fue aprobada por Consejo de Gobierno de 28 de septiembre de 2021⁶, como instrumento general de planificación para las políticas tendentes a la transformación digital de la sociedad andaluza, estructurándose sus fines en torno a las prioridades establecidas a nivel europeo y estatal: competencias digitales para la ciudadanía; competencias digitales para el empleo; competencias digitales para profesionales TIC; competencias digitales en educación; y competencias digitales en salud.

⁵ <https://juntadeandalucia.es/boja/2021/151/1>

⁶ <https://juntadeandalucia.es/boja/2021/190/1>

En tercer lugar, la Estrategia de los Servicios para una Administración Digital centrada en el ciudadano, cuya formulación fue aprobada por Consejo de Gobierno de 20 de septiembre de 2022⁷, que potenciará y reforzará la atención personal a la ciudadanía, avanzando hacia unos servicios públicos ágiles, accesibles, proactivos y útiles, centrados en las necesidades de los usuarios, y que permitan una experiencia de calidad en su interacción con la Administración Pública andaluza.

En cuarto lugar, la Estrategia Cloud, que pone el foco en la importancia de las infraestructuras tecnológicas necesarias para dar soporte al proceso de transformación digital de la Administración, estableciendo la hoja de ruta de la Junta de Andalucía en su viaje a la nube, con el objetivo de habilitar y democratizar la innovación sostenible y continuada, así como el acceso a las tecnologías de la información más avanzadas y una gestión estandarizada y mucho más eficiente y eficaz de la infraestructura.

En quinto lugar, Estrategia Andaluza para el impulso del sector audiovisual y de contenidos digitales, que nace para apoyar al sector en su adaptación al nuevo contexto digital y tendrá en cuenta tanto a los contenidos audiovisuales tradicionales como a las nuevas fórmulas de creación y consumo de contenidos y a los formatos emergentes, como es el caso de los entornos digitales inmersivos. La formulación de esta estrategia fue aprobada en Consejo de Gobierno el 16 de Noviembre de 2022⁸.

En sexto lugar, la Agencia tiene previsto abordar una Estrategia de impulso a la economía digital en su vertiente más empresarial. Se trata de una hoja de ruta para apoyar a nuestras pequeñas y medianas empresas en el proceso de adaptación a los cambios que impone el nuevo entorno digital, como medio para asegurar su competitividad y crecimiento económico a largo plazo. Será pondrá especial foco en acercar soluciones tecnológicas en áreas clave (como la Inteligencia Artificial, el Big Data, el 5G o la Ciberseguridad) a sectores productivos que presentan un nivel bajo y medio de digitalización y fomentar la instauración de una cultura de la innovación en las empresas que contribuya al desarrollo y potenciación de nuevos modelos de nego-

⁷ <https://juntadeandalucia.es/boja/2022/184/1>

⁸ <https://juntadeandalucia.es/boja/2022/223/3>

cio. Esta estrategia abordará también el impulso del emprendimiento de base tecnológica, como vector que contribuye al cambio de modelo productivo, dota de robustez a nuestra economía, atrae y retiene talento y empresas de alto potencial de crecimiento y puede situar a Andalucía en la vanguardia tecnológica y empresarial.

Y en último lugar, la Estrategia andaluza de Inteligencia Artificial, cuya formulación fue aprobada por Consejo de Gobierno de 7 de febrero de 2022⁹, como instrumento general de planificación para las políticas de inteligencia artificial. Esta Estrategia se abordará en detalle en siguientes epígrafes.

2.- ESTRATEGIA DE INTELIGENCIA ARTIFICIAL DE ANDALUCÍA

La integración de las tecnologías emergentes, y en particular de la inteligencia artificial, en el sector público tiene un gran potencial para promover administraciones más ágiles, abiertas e innovadoras. Estas tecnologías permiten aumentar la eficiencia y la calidad de los servicios públicos, además de mejorar la calidad de vida de la ciudadanía por su capacidad para lograr impactos sociales, medioambientales y económicos.

Por este motivo, muchas administraciones públicas han comenzado ya a avanzar en la adopción de IA, a pesar de los desafíos que aún plantea en cuanto a cuestiones éticas y de confianza frente a la ciudadanía.

Concretamente en el caso de la administración regional andaluza, hasta el momento se han llevado a cabo diversas experiencias de éxito que suponen un aprendizaje útil para abordar la definición de la futura estrategia regional de inteligencia artificial. Asimismo, la Junta de Andalucía cuenta actualmente con una varios proyectos en definición y desarrollo enfocados a facilitar el trabajo de los empleados públicos, agilizar trámites, permitir una administración proactiva y mejorar la toma de decisiones a través del análisis de los datos disponibles. Todo ello sitúa a la Junta de Andalucía en una

⁹ <https://juntadeandalucia.es/boja/2022/28/2>

situación favorable para convertirse en administración de referencia en la adopción de IA, dentro del panorama nacional.

La inteligencia artificial está transformando la medicina, el transporte, la industria, el mundo del ocio, y la economía en su conjunto. En los últimos años hemos asistido a una verdadera materialización de sus capacidades, gracias al incremento en la capacidad de computación y de almacenamiento y al cada vez mayor número de conjuntos de datos para entrenamiento de la IA disponibles. Este rápido desarrollo tecnológico ha permitido acercar múltiples aplicaciones a ciudadanía, administraciones públicas y empresas. Pese a sus limitaciones actuales, resulta evidente que la IA tendrá un enorme impacto en la forma de vivir y de trabajar de las personas en las próximas décadas. Por este motivo, la IA ha trascendido del plano conceptual y tecnológico, al estratégico y político, y se sitúa entre las prioridades de numerosos países que actualmente están realizando fuertes inversiones en este ámbito.

La relevancia de la IA se ha puesto de manifiesto también en el contexto europeo, y la UE, a partir de la puesta en marcha de la estrategia europea de IA en abril de 2018¹⁰, la publicación en febrero de 2020 del Libro Blanco de la Comisión sobre la IA¹¹, y la propuesta de Ley de Inteligencia Artificial¹² está poniendo en marcha un paquete de medidas para promover el desarrollo de la IA en Europa y afrontar los posibles riesgos para la seguridad y los derechos fundamentales.

Igualmente, esta relevancia se ve también reflejada a nivel nacional en la Estrategia Nacional de Inteligencia Artificial (ENIA) para el período 2021-2023, cuyo objetivo es generar un entorno de confianza respecto al desarrollo de una IA inclusiva, sostenible, focalizada en el ciudadano y que permita posicionar a España entre los países más punteros en la investigación y el uso de una IA confiable al servicio del desarrollo económico y social.

¹⁰ Comunicación de la Comisión Europea Inteligencia Artificial para Europa COM(2018) 237.

¹¹ Libro Blanco sobre la inteligencia artificial de la Comisión Europea, febrero 2020, COM (2020) 65.

¹² <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>

Andalucía no es ajena a este contexto. La economía andaluza podría verse muy beneficiada de forma directa con la introducción de sistemas de IA en sus sectores estratégicos, y de forma indirecta si desarrollara capacidades en este ámbito que sean exportables al resto del mundo. A todo esto, se une además el potencial de esta tecnología para mejorar la prestación de servicios públicos, pudiendo conseguir servicios más eficientes, equitativos y personalizados. Por otro lado, aunque las oportunidades y beneficios que brinda son evidentes, su desarrollo y adopción también entraña múltiples riesgos asociados al mal uso de los datos, a la vulneración del derecho a la privacidad, o a cuestiones éticas sobre la toma de decisiones que afecten a la discriminación de grupos e individuos. Todas estas cuestiones deben por tanto ser consideradas, con rigor y equilibrio, a la hora de marcar las políticas y acciones públicas de impulso de la IA en Andalucía.

La Junta de Andalucía, consciente de la importancia económica, política y social de la IA para el futuro de la región, y de su capacidad para resolver los grandes retos actuales y venideros, desea poner en marcha una acción política que impulse el posicionamiento de Andalucía en este ámbito. Así, con fecha 7 de febrero de 2022, el Consejo de Gobierno aprueba la formulación de una Estrategia Andaluza de Inteligencia Artificial y encarga a la actual Consejería de Presidencia, Interior, Diálogo Social y Simplificación Administrativa, a través de la Agencia Digital de Andalucía, a su elaboración e impulso, habilitándola para adoptar cuantos actos sean necesarios para la ejecución de dicho compromiso.

En el acuerdo de Consejo de Gobierno de formulación de la Estrategia Andaluza de IA, ésta se conforma como el instrumento general de planificación para las políticas de inteligencia artificial, y se establecen como fines los siguientes:

- Promover la soberanía del conocimiento en IA andaluza para que la región tenga capacidades competitivas y propias en materia de IA en investigación, desarrollo e innovación, y las infraestructuras necesarias para su implantación.
- Fomentar la implantación de la IA en las empresas de la región, de todos los sectores, la creación y escalado de empresas tecnológicas andaluzas en el área de IA, y promocionar Andalucía

como un ecosistema atractivo para atraer sedes de empresas y profesionales del ámbito de la IA a nivel mundial.

- Utilizar la IA andaluza como palanca de cambio en la administración pública, que permita mejorar la competitividad en la gestión de recursos, la toma de decisión, y la agilización y mejora de servicios a los ciudadanos. Esto incluye la coordinación de capacitación del personal, el desarrollo de infraestructuras (datos y computación), la colaboración con centros de investigación/empresas, y la definición de servicios públicos empoderados por la IA. Se definirán una serie de sectores estratégicos, de acuerdo con las prioridades que se establezcan.
- El diseño y puesta en marcha de estructuras de referencia para la coordinación de las administraciones públicas, el tejido productivo y del sistema I+D+i en este ámbito, difundiendo conocimiento y buenas prácticas.

2.1.- La inteligencia artificial en Andalucía

Para llevar a cabo una estrategia de IA, Andalucía no parte de cero. Un reciente estudio¹³ llevado a cabo por la Consejería de Transformación Económica, Industrial, Conocimiento y Universidades de la Junta de Andalucía, que cuenta con la colaboración de la Corporación Tecnológica de Andalucía y el Instituto Andaluz de Investigación en Data Science and Computational (DaSCI), revela que existen al menos 189 entidades con capacidades en materia de IA, es decir, que ya se encuentran desarrollando productos o servicios basados en IA o con investigaciones publicadas y en curso sobre las capacidades de estas tecnologías.

De estas 189 entidades se han identificado 110 de ámbito empresarial, de las cuales son 73% pymes y tienen una participación femenina del 27%. En cuanto a entidades formativas y diseminadoras de conocimiento se han identificado 56 entre el total de 189, con especial foco en las universidades. Destaca la existencia de un centro tecnoló-

¹³ “Identificación de las capacidades andaluzas en el ámbito de la Inteligencia Artificial” de junio de 2021.

gico, 3 fundaciones y 3 grupos de investigación. El resto lo conforman 49 grupos de investigación PAIDI (Plan Andaluz de Investigación, Desarrollo e Innovación). Por último, según este informe Andalucía dispone también de una red de 23 son infraestructuras científicas y/o tecnológicas relevantes, compuesta por laboratorios, centros de supercomputación, Digital Innovation Hubs, centros tecnológicos y de investigación y 10 centros universitarios entre los que destaca la Universidad de Granada, que actualmente es centro de referencia internacional en inteligencia artificial. Además, existen en Andalucía dos fundaciones especializadas en IA: FIAL (Fundación de la inteligencia artificial legal) y la Fundación «AI Granada Research & Innovation».

Con respecto a las áreas de especialidad en IA, destacan disciplinas como análisis predictivo, aprendizaje automático y ciencia de datos. También se encuentran entidades con capacidades y productos relevantes en otras áreas más singulares, como procesamiento del lenguaje natural, visión artificial, automatización o robótica.

Las tecnologías relacionadas con IA se aplican a prácticamente todos los sectores productivos, si bien en Andalucía se catalogan como áreas de trabajo con mayor oportunidad aquellas en las que ya existen productos o servicios disponibles por parte de las empresas de este mapa. Estas áreas son el sector TIC propiamente dicho, seguido de biotecnología y salud, industria y procesos productivos, agroalimentación, administración pública, logística y transporte, aeroespacial o retail.

Dentro del mapa de capacidades andaluzas, destacan Sevilla y Málaga como las provincias donde se concentran la mayoría de las entidades, y Granada, en lo concerniente al ámbito académico e investigador. Precisamente son estas capacidades lo que hacen que Granada se postule como sede de la Agencia Nacional de Supervisión y Control de Algoritmos, habiendo sido su candidatura avalada por el Consejo de Gobierno de la Junta de Andalucía¹⁴.

Por tanto, la hoja de ruta de la Junta de Andalucía en materia de IA debe avanzar sobre las capacidades ya existentes, articulando a los agentes del entorno, potenciando las fortalezas con las que cuenta la región, y fijando criterios claros que hagan su desarrollo, uso y explotación seguro y confiable.

¹⁴ <https://juntadeandalucia.es/boja/2022/112/79>

2.2.- Misión y visión de la Estrategia

La visión de la Estrategia andaluza de Inteligencia Artificial es *“mejorar la calidad de vida de la ciudadanía y el posicionamiento del tejido productivo andaluz a través del aprovechamiento del potencial de la IA, situando a Andalucía entre las regiones líderes en economía digital”*. Para ello, el gobierno andaluz impulsará un ecosistema que fomente el desarrollo e implantación de la inteligencia artificial a través de la atracción de talento y el llamamiento a la inversión público y privada.

La misión de la Estrategia andaluza de Inteligencia Artificial es *“impulsar el desarrollo y aplicación de la IA en Andalucía, articulando a los agentes clave de su entorno, potenciando nuestras capacidades competitivas, generando las infraestructuras y recursos necesarios para su implantación, y fijando criterios que hagan su desarrollo, uso y explotación seguro, confiable y ético”*. Para dar respuesta a los retos derivados de la nueva economía digital, Andalucía apuesta por la incorporación de tecnologías disruptivas entre las que destaca la IA como palanca principal para mejorar la calidad de los servicios públicos percibidos por la ciudadanía y la maximización de la competitividad, eficiencia y eficacia de sus empresas.

2.3.- Comité de expertos

Con el fin de asesorar a la Junta de Andalucía en el diseño de sus políticas de impulso de la inteligencia artificial en la región, se desea contar con la visión experta de un conjunto de profesionales de reconocido prestigio, con conocimiento o experiencia en la materia, y con vinculación a Andalucía.

Así, el 6 de junio de 2022, se constituye oficialmente en Granada un comité de personas expertas cuya misión será prestar asesoramiento a la Junta de Andalucía en la formulación de la Estrategia Andaluza de IA, aportando su valiosa visión, exógena a la administración ejecutora de las actuaciones, bajo los principios de objetividad, transparencia, equilibrio territorial y social y uso seguro y ético de la IA, al objeto de adecuar las políticas públicas a las necesidades reales del tejido productivo y de la sociedad andaluza en su conjunto, maximizando las oportunidades para la región andaluza.

El comité de expertos de la Estrategia andaluza de IA está conformado por 26 personas, en su mayoría profesionales procedentes del ámbito empresarial y académico, que han sido identificadas y seleccionadas en base a su excelencia en el ámbito de la IA desde sus diversas vertientes (tecnología, investigación, innovación, regulación, ética).

Los miembros del comité de expertos son los siguientes:

Emilia Gómez – Comisión Europea

M^a Teresa Martín – Universidad de Jaén

Alicia Troncoso – Universidad Pablo de Olavide

Enrique Herrera-Viedma – Universidad de Granada

Antonio J. Fernández – Universidad de Málaga

Jose F. Aldana – Universidad de Málaga

Eduardo Gamero – Universidad Pablo Olavide

Gabriel Vestri – Universidad de Cádiz

Xavier Valero – Dekra Digital

Pilar Manchón – Google Research

David Carmona – Microsoft

Juan M. Melero – Ericsson

María Mora – Fujitsu

Juan Bernabé- Moreno – IBM

Leticia Gómez – Minsait

Víctor Gaspar – GMV

Luis Buezo – HP

Javier Porras – Unicaja

Juan Ignacio Arcos – BigML

Daniel Urbina – Amazon

Jose Manuel López – Sopra STERIA

Pablo Enciso – Predictiva

Ricardo Arjona – EC2CE

Francisco Alarcos – Universidad Loyola

Lola Carranza – Montero Aramburu

2.4.- Metodología

El proceso de formulación de la Estrategia Andaluza de IA se divide en cuatro fases, cada una de las cuales comprende la ejecución de una serie de tareas. Estas son:

FASE 0. LANZAMIENTO DEL PROYECTO

Se trata de una fase preparatoria dentro del proceso de formulación cuyo objetivo es establecer aspectos organizativos, como la delimitación del alcance, la planificación, los mecanismos de control y seguimiento, las fuentes de información, la gestión documental, la identificación de participantes y equipos de trabajo, los medios de participación, etc.

A la finalización de esta fase, se cuenta con un sistema de gobernanza ya establecido, habiéndose identificado los agentes clave, grupos de trabajo y responsabilidades.

Otro de los productos importantes de esta fase es la delimitación de la misión y visión de la estrategia, que es el motivo fundamental de la acción pública.

FASE I. DIAGNÓSTICO INICIAL

Para la elaboración del plan estratégico resulta imprescindible llevar a cabo un análisis de la realidad que nos permita realizar un diagnóstico de los problemas, necesidades y retos que podremos abordar a través de la estrategia.

Para elaborar el diagnóstico de situación de la IA en Andalucía, se han llevado a cabo las siguientes actividades:

- Análisis del contexto internacional y nacional de estrategias IA
- Análisis del contexto IA en Andalucía
- Análisis de fuentes de financiación
- Visión externa (comité de expertos)
- Visión interna (comité técnico)
- Visión sectorial (a través de encuestas a entidades de referencia)

Como resultado de este trabajo se ha obtenido un DAFO interno (principales debilidades, amenazas, fortalezas y oportunidades detectadas para la adopción de IA dentro de Junta de Andalucía) y un DAFO externo (principales debilidades, amenazas, fortalezas y oportunidades detectadas para impulsar la IA en la región). Estos análisis DAFO permitirán la reflexión y obtención de conclusiones para marcar las prioridades de la acción política.

FASE II. DESARROLLO DE LA ESTRATEGIA

A partir del diagnóstico de situación, se establecerán los objetivos, los ejes estratégicos, y las actuaciones del plan. Entre estas últimas, se contará con propuestas de casos de uso a abordar dentro de la Junta de Andalucía en los próximos años, así como medidas específicas orientadas a fortalecer e impulsar el ecosistema regional de IA. Para el diseño de estas actuaciones se contará con la participación activa del comité interno así como con el asesoramiento del comité de personas expertas. Por último, se diseñará un modelo de seguimiento y evaluación de la Estrategia.

Con toda esta información la oficina técnica redactará un primer borrador de Estrategia Andaluza de IA.

FASE III. TRAMITACIÓN Y APROBACIÓN DE LA ESTRATEGIA

Una vez elaborado el borrador de Estrategia, éste será sometido al trámite de información pública, por un periodo no inferior a quince días, previo anuncio en el Boletín Oficial de la Junta de Andalucía, en la sección de transparencia del Portal de la Junta de Andalucía y en la web de la Consejería de Presidencia, Interior, Diálogo Social y Simplificación Administrativa, siguiendo los cauces dispuestos en la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas. Asimismo, se recabarán los informe preceptivos exigidos por la normativa general de aplicación.

Finalmente, la ADA propondrá a la persona titular de la Consejería de Presidencia, Interior, Diálogo Social y Simplificación Administrativa el texto de la Estrategia, que elevará al Consejo de Gobierno para su aprobación definitiva.

Una vez aprobado el documento, se llevarán a cabo acciones de comunicación para la difusión de la Estrategia andaluza de IA.

2.5.- Modelo de gobernanza

La gobernanza conforma uno de los principios rectores en la formulación de todo plan estratégico y está presente en las distintas etapas del diseño, ejecución y evaluación de una Estrategia.

El nivel superior dentro del modelo de gobernanza articulado para el diseño de la Estrategia andaluza de IA es el Consejo de Gobierno de la Junta de Andalucía. Se trata del órgano que acuerda la formulación de la Estrategia y establece la base del proceso de elaboración del mismo: fines, contenido, metodología y organismos participantes. Asimismo, será el Consejo de Gobierno el que apruebe el documento final.

En un segundo nivel del modelo de gobernanza, se sitúa la Agencia Digital de Andalucía, que es el órgano responsable de impulsar y elaborar el borrador de Estrategia, ocupando un rol de liderazgo y de toma de decisiones de todo el proceso.

El Comité Técnico y el Comité de Expertos son los dos siguientes elementos dentro del sistema de gobernanza de la Estrategia andaluza de IA. Por una parte, el Comité Técnico está formado por los centros directivos competentes en cualquiera de los ámbitos sectoriales en los que la IA tiene afección. Los miembros del Comité Técnico participarán activamente en la elaboración del borrador de Estrategia, aportando su conocimiento y actuando como interlocutores dentro de su Consejería para presentar la propuesta de acción sectorial/competencial en el impulso de la IA tanto en la prestación de servicio público, como en el fortalecimiento de cualquiera de los pilares del ecosistema andaluz de IA. Por otra parte, el Comité de Expertos está integrado por una representación de profesionales destacados del ámbito de la IA, perteneciente a entidades externas a la administración de la Junta de Andalucía. Su misión es la de asesorar a la Junta de Andalucía en el proceso de formulación, aportando enfoques y visión experta, siempre bajo los principios de objetividad, transparencia, equilibrio territorial y social, y uso seguro y ético de la IA.

Por último, como un elemento clave dentro del modelo de gobernanza de la Estrategia, se sitúa la sociedad civil y empresarial que, a través del trámite de consulta pública, podrá participar en el proceso de elaboración de la estrategia previo a su elevación a Consejo de Gobierno.

Capítulo XV

Organización y gobernanza de la Inteligencia Artificial en las entidades locales

M. ASCENSIÓN MORO CORDERO

*Responsable del Departamento de Gobierno Abierto
Ayuntamiento de Sant Feliu de Llobregat*

Resumen: Es inaplazable que las Administraciones públicas y, en particular, las entidades locales aborden el proceso de cambio profundo y estructural que se necesita para transitar hacia una administración inteligente, sostenible, abierta, inclusiva y social, que se adapte a las demandas y nuevas formas de relación con la ciudadanía, en un contexto de revolución digital en el que las tecnologías disruptivas, en especial, la inteligencia artificial, presenta retos y desafíos aún inimaginables. El trabajo analiza, desde una perspectiva holística, el impacto que puede tener la incorporación de esta tecnología en las entidades locales y cuáles son los elementos esenciales que se deben tener en cuenta a la hora de plantear una estrategia de gobernanza integral y multidimensional de esta transformación.

Abstract: It is unplanted that public administrations and, in particular, local entities address the process of deep and structural change that is needed to travel towards an intelligent, sustainable, open, inclusive and social administration, which adapts to the demands and new forms of Relationship with citizenship, in a context of digital revolution in which disruptive technologies, especially artificial intelligence, presents still unimaginable challenges and challenges. The work analyzes, from a holistic perspective, the impact that the incorporation of this technology can have in local entities and what are the essential elements that must be taken into account when proposing an integral and multidimensional governance strategy of this transformation .

Palabras-clave: Gobernanza de datos, transformación digital, inteligencia artificial, administración digital, transparencia, protección de datos, ética

Key-words: Data governance, digital transformation, artificial intelligence, digital administration, transparency, data protection, ethics

SUMARIO.- 1. INTRODUCCIÓN.- 2. EL RETO DE LA TRANSFORMACIÓN DIGITAL DE LA ADMINISTRACIÓN LOCAL ¿UN PROCESO *AD INFINITUM*?- 3. LA NECESARIA GOBERNANZA MULTINIVEL Y MULTIDIMENSIONAL.- 3.1. Un modelo de gestión inteligente.- 3.2. Un modelo de gobernanza de los datos.- 3.3. Una metodología.- 4. POSIBLES USOS DE LA INTELIGENCIA ARTIFICIAL Y DESAFÍOS.- 5. CONCLUSIONES.- BIBLIOGRAFÍA

1.- INTRODUCCIÓN

Sin duda a estas alturas es una obviedad señalar que las Administraciones públicas, y en especial las entidades locales, están (o deberían estar) inmersas en un proceso de cambio profundo y estructural con la idea de transitar hacia una administración inteligente, sostenible, abierta y social, capaz de estar a la altura y dar respuesta a los múltiples desafíos actuales y a las nuevas demandas de la ciudadanía y de las empresas. En particular, a la agenda tradicional de la gestión pública se han sumado ahora retos en materia de sostenibilidad, resiliencia, de lucha contra las desigualdades de todo tipo, adaptaciones a cambios acelerados y constantes...sin olvidar el enorme impacto que supone la revolución digital y la nueva economía de los datos.

Los efectos que pueden tener los acelerados y previsibles cambios tecnológicos asociados a esta revolución digital han sido extensamente analizados y contamos con estudios¹ realizados por prestigiosos expertos de diversos países que destacan la necesidad de abordar cambios profundos en el sistema educativo, económico, cultural y político, y en la propia concepción de las formas de trabajo tradicionales. Entre las tendencias y factores de cambio que se destacan, los más relevantes a modo de resumen serían:

- A nivel político, desafección y crisis del sistema; innovación social aplicada a los desafíos de las sociedades...
- A nivel medioambiental, el cambio climático; las limitaciones de los recursos naturales; la necesidad de replantear el modelo energético y la movilidad sostenible...
- A nivel económico, la transición a una economía circular; los cambios en el mercado de trabajo; el *crowdsourcing*, la economía compartida y *peer-to-peer*...
- A nivel tecnológico, los avances en programación y *big data*, el impacto del Internet de las cosas (IoT), la robótica, la Inteligencia Artificial (IA) y el aprendizaje autónomo; las redes sociales; el *cloud*...

¹ https://www.fundaciontelefonica.com/artes_cultura/publicaciones-listado/pagina-item-publicaciones/itempubli/636/

- Y a nivel social, sociedades cada vez más envejecidas; los cambios en la estructura familiar; la creciente desigualdad entre ricos-pobres, así como en el acceso a la educación y a la sanidad; sin olvidar los riesgos que supone el volumen cada vez mayor de información personal que se difunde a escala mundial.

Las crisis a las que nos estamos enfrentando en este primer cuarto de siglo han puesto sobre la mesa algunos efectos no deseados de la globalización y, por otro lado, han reforzado algunos de los retos sociales ya enunciados y que requieren un planteamiento riguroso y efectivo, más allá del eterno diagnóstico permanente:

- La necesidad de avanzar en la digitalización y de reforzar las capacidades tecnológicas y organizativas de las Administraciones públicas, para poder impulsar el relanzamiento social y económico, en particular, de los municipios y su inestimable contribución a la regeneración a escala internacional.
- La relevancia de la innovación.
- Abordar la función pública profesional cualificada, nuevos perfiles y competencias, nuevas formas de trabajo en movilidad y ubicuas, así como la planificación y gobernanza del inevitable relevo generacional.
- La necesidad de reducir la brecha digital, en sus múltiples dimensiones.
- La falta de recursos para la prestación de servicios públicos y, en particular, la salud, y la urgencia de garantizar a los sectores de población más vulnerables el acceso a los servicios básicos y de asegurar un nivel de vida digno.
- Y, en definitiva, la importancia de construir comunidades más resilientes.

En este sentido, la organización mundial de ciudades y gobiernos locales unidos², en su decálogo para la postpandemia³, hace hincapié en los sacrificios que ha comportado la pandemia de la COVID 19, y

² <https://www.uclg.org/es/organizacion/sobre-nosotros>

³ https://www.uclg.org/sites/default/files/decalogo_covid19.pdf

las consecuencias para las generaciones futuras, lo que hace necesario dotarse de planes estratégicos sólidos para poder avanzar hacia un futuro sostenible.

Por ello, los principios de la Agenda mundial 2030 y los Objetivos de Desarrollo Sostenible (ODS) siguen siendo, ahora más que nunca, una referencia absolutamente imprescindible para enmarcar las medidas de transformación que se están implementando a nivel local. Por lo que se refiere a la digitalización, el contexto actual caracterizado por el avance continuado y exponencial de las tecnologías disruptivas (5G, computación cuántica, Inteligencia Artificial (IA), Internet de las Cosas (IoT) y de los Servicios (IoS), el uso móvil de los datos y la disponibilidad de potentes sistemas integrados, sensores y actuadores a precios razonables), está reconfigurando nuestra vida cotidiana y comporta cambios profundos en la economía y el entorno de trabajo.

En concreto, como retos clave para las Administraciones públicas en este ámbito cabe resaltar:

- La digitalización de los puestos de trabajo y la aparición de nuevos perfiles profesionales
- La escasez de personal especializado
- Los datos masivos (*Big Data*), los sistemas que gestionan grandes conjuntos de datos y las dificultades que comportan la captura, el almacenamiento, el procesamiento, la búsqueda, visualización, compartición y simulación
- Los retos que implican la colaboración segura entre robots y humanos
- La ciberseguridad y la resiliencia de un sistema para continuar funcionando incluso en caso de quiebras y ataques
- La protección de derechos fundamentales y libertades públicas, que implica una nueva cultura de gestión de los datos y de su control por parte de la ciudadanía, con la incorporación de tecnologías emergentes en proyectos que ya se están desarrollando a nivel europeo como la identidad europea soberana, que permite a los usuarios crear y controlar su propia identidad
- La posibilidad de crear nuevos servicios gracias a la tecnología 5G, imposibles de ofrecer con la tecnología 4G.

Lo cierto es que actualmente la situación a nivel europeo de prestación de servicios digitales está muy por encima de lo que podríamos imaginar. De hecho, el informe anual (*eGovernment Benchmark 2022*⁴) del que se dota la Comisión Europea para conocer el nivel de disponibilidad *online* de servicios de administración pública en Europa, ha reflejado esta aceleración que ha supuesto la crisis pandémica, destacando que más del 80% de los servicios públicos europeos ya se encuentran en formato digital. Si bien, el propio informe recoge también tres recomendaciones clave relacionadas con el despliegue de la administración digital:

- Priorizar el diseño centrado en las personas para garantizar que los servicios de administración electrónica sean inclusivos para las personas usuarias en toda su diversidad, incluidas aquellas con pocas habilidades digitales o aquellas que viven con discapacidades.
- Racionalizar la prestación de servicios de administración electrónica para que las personas usuarias puedan acceder a todos los servicios relacionados con su vida a través de ventanillas únicas.
- Agilizar la interoperabilidad entre las diferentes capas y niveles de gobierno para brindar una experiencia más coherente y menos engorrosa.

El informe refleja, a su vez, una realidad que hace tiempo los académicos y operadores jurídicos y técnicos venimos demandando en base a un diagnóstico casi podríamos decir generalizado:

Los servicios públicos digitales de las diferentes Administraciones públicas no están en el mismo nivel de madurez, lo que dificulta en gran medida la prestación de servicios públicos de calidad.

La necesidad de conseguir una verdadera gobernanza multinivel y multidimensional, con la incorporación de los múltiples agentes afectados: “*La implantación de soluciones interoperables en múltiples entidades y niveles gubernamentales creará una experiencia de*

⁴ <https://digital-strategy.ec.europa.eu/en/library/egovernment-benchmark-2022>

administración electrónica más coherente para los distintos usuarios y proveedores de servicios de los países europeos”.

La importancia que, cada vez más, cobra la experiencia del usuario y su incorporación real en el diseño y mejora de los procesos.

A todo ello hay que sumar el aumento exponencial del volumen de datos que se producen. De hecho se estima que para 2025 se alcanzará la cifra anual de 175 zettabytes (siendo un zettabyte el equivalente a mil trillones de bytes). Esto explica claramente el valor de los datos como motor de desarrollo económico y social, así como el creciente interés de las empresas, centros de investigación y Administraciones públicas por extraer el máximo rendimiento en el tratamiento y la utilización de estos datos con distintas finalidades. En parte, este ha sido uno de los motivos por lo que las tecnologías emergentes, y en especial, la Inteligencia Artificial (IA), está cobrando protagonismo en lo que tiene que ver con la innovación en la administración pública.

2.- EL RETO DE LA TRANSFORMACIÓN DIGITAL DE LA ADMINISTRACIÓN LOCAL ¿UN PROCESO AD INFINITUM?

La administración local inició su proceso de transformación interno bajo los albueros iniciales del despliegue de la denominada administración electrónica, que ha provocado, entre otras muchas cuestiones, una heterogeneidad de infraestructuras, modelos, metodologías y niveles de implementación, generalizando así una manifiesta asincronía de los servicios públicos digitales implementados, en función del nivel de madurez de cada organización.

Lo cierto es que el amplísimo marco normativo actual al que se tienen que enfrentar las entidades locales plantea un escenario, como poco, de esquizofrenia organizativa ya que, entre otras cuestiones, el ámbito material está muy fraccionado; las leyes y reglamentos de desarrollo no abordan los aspectos prácticos a la hora de implementar esta transformación y su carácter, en algunos casos, marcadamente técnico, obliga a establecer, necesariamente, metodologías de trabajo multidisciplinarias que chocan frontalmente con formas de trabajo tradicionales; sin olvidar la enorme brecha aún vigente en cuanto a perfiles profesionales capaces de abordar este cambio de paradigma.

Ello ha provocado, entre otras múltiples consecuencias, cierta frustración entre los propios usuarios internos y entre la ciudadanía, que no perciben que la introducción de la tecnología en los procesos tenga los resultados esperados.

De hecho, hace ya bastante tiempo que existe la idea compartida de que la transformación digital de las Administraciones públicas es un proceso multidimensional, y que no es suficiente incorporar la tecnología en la actividad municipal, sino que hace falta articular y gestionar, en paralelo, el necesario cambio cultural, organizativo y jurídico para que se produzca esta transformación real.

En este sentido se manifestó la Comisión Europea en el año 2003⁵ [COM (2003) 567 final] indicando que la administración electrónica es *“el uso de las TIC en las Administraciones públicas, combinado con cambios organizativos y nuevas aptitudes, con la finalidad de mejorar los servicios públicos y los procesos democráticos y reforzar el apoyo a las políticas públicas”*, es decir, que no es suficiente con incorporar los medios electrónicos en la actividad de las Administraciones públicas, también es necesario un cambio de cultura administrativa impulsado por un liderazgo claro y decidido, y acompañado de un marco regulador.

La tozuda realidad es que, efectivamente, el desarrollo de la administración electrónica no se ha visto reflejado en una innovación de las entidades locales que, en general, se han limitado a aplicar la tecnología a su alcance a los procesos tal y como se venían realizando en el modelo tradicional, en lugar de utilizar esta tecnología para transformar estos procesos generando realmente valor público para la sociedad, además de mejorar internamente en eficiencia y eficacia. Las Administraciones públicas, con carácter general, se han limitado a “electronificar” procesos y prestarlos a distancia, pero sin que ello suponga cambios disruptivos en la forma de prestación de los servicios públicos.

Por otro lado, en este proceso de transformación profunda se debe tener presente no únicamente el amplio marco normativo vigente, sino también los retos y objetivos que se derivan de estrategias y

⁵ <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2003:0567:FIN:ES:PDF>

agendas globales que están intentando orientar el sector público europeo hacia un crecimiento inteligente, sostenible e integrador, atendiendo a las nuevas demandas sociales inmersas en esta revolución digital, y teniendo en cuenta la convergencia actual de tecnologías emergentes que impactan de lleno en un nuevo modelo de administración digital basada en datos.

En este sentido, un punto de inflexión fue sin duda el Plan de Acción de Administración Electrónica de la Unión Europea para el período 2016-2020⁶ que establecía los principios que, como mínimo, se debían tener presentes a la hora de implementar servicios públicos digitales: digital por defecto, el principio de una sola vez, la inclusión y la accesibilidad, la apertura y transparencia, la interoperabilidad transfronteriza de forma predeterminada, la confianza y la seguridad.

Posteriormente, la Comisión Europea publicó la Comunicación “Configurar el futuro digital de Europa” [COM (2020) 67 final, de 19 de febrero de 2020]⁷, que establece la estrategia digital europea en el marco de una economía basada en datos, y que se desarrolla en tres grandes áreas:

- Una tecnología que sitúa en el centro a las personas;
- Una economía digital justa y competitiva;
- Una sociedad abierta, democrática y sostenible.

En este momento lo que se persigue es impulsar la disponibilidad, la calidad y el uso de los datos en beneficio de la sociedad, aplicando tecnologías emergentes como la Inteligencia Artificial, pero garantizando los derechos y libertades de las personas. Por ello, en paralelo, se han ido adoptando normas relacionadas con la protección de datos personales (Reglamento (UE) 2016/679); con la libre circulación de datos no personales (Reglamento (UE) 2018/1807); la Ciberseguridad (Reglamento (UE) 2019/881); o la Directiva sobre Datos Abiertos (Directiva (UE) 2019/1024), entre otras.

⁶ <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52016DC0179&from=ES>

⁷ <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020DC0067&from=ES>

A su vez, actualmente la estrategia digital europea se está materializando en documentos que, sin duda, tienen una enorme importancia para el despliegue de la administración digital, como por ejemplo el Libro Blanco sobre la inteligencia artificial⁸, que establece las bases para el futuro desarrollo de un marco legislativo en materia de aplicabilidad de la Inteligencia Artificial que genere confianza en la ciudadanía y, al mismo tiempo, impulso para su desarrollo por parte de las empresas. Con posterioridad, en relación con esta cuestión se ha presentado la propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión [COM(2021)0206]⁹.

Otro de los documentos que derivan de la estrategia digital europea es la Estrategia Europea de Datos (Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones de 19 de febrero de 2020 [COM(2020) 66 final]), donde se anuncia un marco legislativo para la gobernanza de los datos con el objetivo de crear un espacio común europeo de datos “como un mercado interior de datos en el que estos puedan utilizarse independientemente de su ubicación física de almacenamiento en la Unión de conformidad con el Derecho aplicable, lo que, entre otras cosas, podría resultar fundamental para el rápido desarrollo de las tecnologías de inteligencia artificial”¹⁰.

⁸ https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf

⁹ Cabe destacar el proyecto piloto que el Gobierno de España y la Comisión Europea presentaron el pasado 27 de junio de 2022 para poner en marcha el primer *sandbox* (banco de pruebas) regulatorio de la Unión Europea sobre Inteligencia Artificial. La experiencia pretende documentar las obligaciones que los proveedores de sistemas de IA deben cumplir e implementar, así como el método para hacer un control y un seguimiento que sirva de modelo para las autoridades nacionales de supervisión, tal y como se establecerá en el futuro reglamento. En esta prueba piloto participan tanto autoridades competentes como empresas desarrolladoras de IA, con el objetivo de definir de forma conjunta buenas prácticas que sirvan de base para la implementación de la futura regulación europea de Inteligencia Artificial (*Artificial Intelligence Act*), y garantizar que se puede aplicar en el tiempo previsto.

¹⁰ Considerando 2 Reglamento de Gobernanza de Datos.

A nivel estatal, la hoja de ruta que marca la transformación digital del país es la agenda España Digital 2026¹¹, que tiene como objetivo aprovechar las nuevas tecnologías para lograr un crecimiento económico basado en la sostenibilidad, el empleo de calidad, el aumento de la productividad y que contribuya a la cohesión social y territorial, aportando prosperidad y bienestar a la ciudadanía. La estrategia, que está sincronizada con el Plan de Recuperación, Transformación y Resiliencia de la economía española, presentado en octubre de 2020, persigue, entre otras finalidades, acelerar una transición digital humanista en España y, en lo que nos ocupa, contempla medidas agrupadas en diez ejes estratégicos y dos transversales:

1. Conectividad digital
2. Impulso a la tecnología 5G
3. Ciberseguridad
4. Economía del dato e Inteligencia Artificial
5. Transformación digital del sector público
6. Transformación digital de las pymes y emprendimiento digital
7. España, Hub Audiovisual
8. Transformación digital sectorial y sostenible
9. Competencias digitales
10. Derechos digitales
11. PERTE: Proyectos Estratégicos para la Recuperación y Transformación Económica
12. RETECH: Redes Territoriales de Especialización Tecnológica

En desarrollo de este marco estratégico nacional, se han publicado los siguientes planes específicos:

- Plan para la Conectividad y las Infraestructuras Digitales¹²
- Estrategia de Impulso de la Tecnología 5G¹³

¹¹ <https://espanadigital.gob.es/>

¹² https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/ficheros/201202_Plan_para_la_Conectividad.pdf

¹³ https://portal.mineco.gob.es/RecursosNoticia/mineco/prensa/noticias/2020/201201_np_impulso5G.pdf

- Estrategia Nacional de Inteligencia Artificial¹⁴
- Plan de Digitalización de Pymes¹⁵
- Plan de Digitalización de las Administraciones públicas¹⁶
- Plan Nacional de Competencias Digitales¹⁷

En particular, el Plan de Digitalización de las Administraciones públicas 2021-2025, define la Estrategia en materia de Administración Digital y Servicios Públicos Digitales, contempla como objetivos la implementación de servicios digitales, accesibles, eficientes, seguros y fiables; el establecimiento de políticas públicas basadas en datos, la modernización de la gestión de datos y la democratización del acceso a las tecnologías emergentes.

Parece pues que todo confluye en un salto cualitativo del anterior concepto de administración electrónica en un nuevo concepto de administración digital, que tenga como pilares el despliegue la gestión inteligente (*SmartCity*) y la incorporación de las tecnologías disruptivas con el doble objetivo de mejorar y fortalecer la relación entre las Administraciones públicas y la ciudadanía, pero a la vez, aportar valor público a la sociedad.

Pero para que no vuelvan a repetirse los errores detectados en el despliegue de la administración electrónica, y que las entidades locales, todavía transitando hacia lo digital competencialmente hablando, logren sobrevivir a esta inundación tecnológica y normativa, se hace indispensable que se doten de un instrumento propio, alineado por supuesto con las estrategias de carácter global, en especial en lo que tiene que ver con la consecución de los Objetivos de Desarrollo

¹⁴ https://portal.mineco.gob.es/RecursosNoticia/mineco/prensa/noticias/2020/201202_np_eniav.pdf

¹⁵ https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/ficheros/210127_plan_digitalizacion_pymes.pdf

¹⁶ https://administracionelectronica.gob.es/pae_Home/pae_Estrategias/Plan_Digitalizacion_AAPP.html?urlMagnolia=/pae_Home/pae_Estrategias/Estrategia-TIC/Plan-Digitalizacion-AAPP.html

¹⁷ https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/ficheros/210127_plan_nacional_de_competencias_digitales.pdf

Sostenible¹⁸, que aglutine la multitud de dimensiones que conforman esta transformación, y que necesariamente deben ir avanzando de forma conjunta. Este ejercicio es fundamental para garantizar, por un lado, que la organización es consciente de su grado de madurez (para ello resulta muy aconsejable realizar algún tipo de auditoría previa para que, de forma honesta, se ponga de manifiesto el estado actual de desarrollo de todas esas dimensiones que conforman este proceso de transformación), y por otro lado, poder diseñar una hoja de ruta realista y evidenciar, así, el conjunto de instrumentos, perfiles, recursos e infraestructuras necesarias.

3.- LA NECESARIA GOBERNANZA MULTINIVEL Y MULTIDIMENSIONAL

Si, como hemos visto en los apartados anteriores y defienden numerosos académicos, la conclusión a la que se llega después de todos estos años es que la transformación integral en la que están inmersas las Administraciones públicas tiene un carácter marcadamente poliédrico, donde ahora además se incorporan tecnologías emergentes que, sin duda, van a propiciar un salto exponencial en este cambio de paradigma, parece sensato que las entidades locales contemplen de forma holística su organización y gobernanza, identificando las principales líneas de actuación para afrontar todos estos nuevos retos y conseguir una mayor calidad de los servicios públicos y eficiencia en su gestión, la satisfacción de la ciudadanía y de las empresas en su relación con la administración, contribuyendo, así, a potenciar el crecimiento econó-

¹⁸ La Nueva Agenda Urbana aprobada en la Conferencia de las Naciones Unidas sobre la Vivienda y el Desarrollo Urbano Sostenible (Hábitat III), y refrendada por la Asamblea General de las Naciones Unidas 2016, configura un conjunto de retos que deben abordar los municipios en el siglo XXI para proveer bienestar, igualdad, desarrollo y seguridad a las personas y, de este modo, contribuir activamente a la consecución de los Objetivos de Desarrollo Sostenible previstos en la Agenda 2030. A nivel estatal en 2019 se aprobó la Agenda Urbana Española, configurada como un instrumento de planificación estratégica e integrada para los municipios. Esta agenda cuenta con una metodología específica que las entidades locales pueden adoptar de forma voluntaria para elaborar su propio plan de acción. Actualmente, constituye una de las líneas subvencionables en el marco del Plan de Recuperación, Transformación y Resiliencia aprobado por el Estado.

mico, la mejora de la calidad de vida de la sociedad, la reducción de desigualdades y cohesión social, e impulsando la transición ecológica y digital.

Las entidades locales deben, pues, dotarse de una estrategia propia en materia de transformación digital, con una visión a medio y largo plazo, que contemple como mínimo los siguientes aspectos:

- La necesidad de transformar la estructura productiva de la organización y facilitar el cambio cultural. Es preciso revisar el actual modelo organizativo para incrementar el buen rendimiento del funcionamiento interno. Entran aquí aspectos relacionados con la selección de nuevos perfiles profesionales; el relevo generacional; la capacitación y formación continua; nuevas formas de trabajo más relacionales y transversales; la movilidad y ubicuidad de los puestos de trabajo; la dirección pública profesional, entre otros aspectos.
- Desplegar un nuevo marco relacional con la ciudadanía, que incluya la atención ciudadana integral y omnicanal, que mejore la experiencia de la ciudadanía y las empresas, caracterizado por la agilidad, la proactividad, la proximidad y la capacidad de adaptación a las necesidades presentes y futuras.

En este sentido, cabe destacar el trabajo que se está realizando desde la Dirección general de servicios digitales y experiencia ciudadana de la Generalitat de Catalunya por desplegar un nuevo modelo basado en la escucha activa de las necesidades, expectativas, satisfacción y comportamiento de la ciudadanía, que incluye acompañarles para favorecer su acceso a los servicios públicos digitales; una atención integrada, garantizando los procesos de extremo a extremo; y una comunicación directa y con un lenguaje fácil y accesible.¹⁹ Para ello están desplegando “espacios

¹⁹ <https://www.salesforce.com/es/company/news-press/press-releases/2022/04/220426/> Las principales conclusiones de este estudio es que pese a que el 54% de los encuestados ha utilizado la administración digital, un 71% considera que el servicio de la administración pública digital no es de calidad, y un 35% afirma que los servicios digitales ofrecidos por la Administración pública son demasiado complicados. Por el contrario, la ciudadanía en general está dispuesta a compartir más datos con las Administraciones públicas (esencial para avanzar en la proactivi-

de experiencia ciudadana²⁰”, para testear con la ciudadanía los servicios públicos antes de su puesta en funcionamiento y mejorar los existentes, verificando aspectos como la accesibilidad, la usabilidad, el lenguaje fácil, los diferentes pasos que se incluyen en los trámites puestos a disposición en la sede electrónica, etc.

- El desarrollo de una administración digital abierta, inteligente, inclusiva y social, para conseguir una mayor eficiencia en los procesos internos pero también dotando a las ciudades de infraestructuras, procesos y servicios inteligentes.
- La implementación de un modelo de gobierno abierto que incorpore la transparencia, la gestión responsable y la ética en la actividad pública, impulsando el acceso y la reutilización de la información, mediante políticas que contemplen la cultura de la rendición de cuentas, la participación y la colaboración con la ciudadanía en el nuevo diseño, redefinición y mejora de los servicios públicos y en las iniciativas de *Smart City*.
- La creación de una oferta innovadora de servicios para la ciudadanía y eficiente en su gestión, a los que se pueda acceder en cualquier momento y lugar, en condiciones de igualdad; servicios inclusivos y personalizados, adaptándose de manera ágil a las demandas presentes y futuras de la ciudadanía, en las condiciones adecuadas de confianza y seguridad. La buena gobernanza requiere además que la prestación de estos servicios incorpore los indicadores necesarios que permitan evaluar los compromisos adoptados y rendir cuentas de forma permanente.

Lo ideal es que esta estrategia se materialice en instrumentos ejecutivos, tipo agendas digitales, como marcos de referencia para el desarrollo digital, sostenible y resiliente del municipio, que marquen la hoja de ruta, metodologías, ámbitos de actuación, objetivos estratégicos y operativos, acciones concretas medibles y auditables por parte de todos los agentes implicados (personal de la administración, ciudadanía, empresas,

dad y la personalización de los servicios públicos) y el 77% confía en la seguridad de los servicios públicos digitales y en la gestión y privacidad de sus datos personales.

²⁰

<https://web.gencat.cat/ca/temes/administracio/espais-dexperiencia-ciudadana/>

entidades...). De hecho, existen ejemplos interesantes a nivel local²¹ y propuestas tanto académicas²² como por parte de entidades supramunicipales. En este último caso cabe destacar el trabajo que está realizando el Consorcio Localret²³ del que forman parte más de 850 ayuntamientos de Catalunya, y que, mediante un proceso participativo, ha elaborado una agenda digital de los municipios de Catalunya, que aspira a tener una triple función: ser un instrumento aglutinador y facilitador para la concreción de estrategias, planes y proyectos de transformación digital comunes; proporcionar a las administraciones catalanas una visión holística de los elementos que intervienen en la transformación digital de las ciudades, así como poner en común las necesidades y dificultades para el despliegue digital de los municipios catalanes.

En todo caso, tomando como referencia la Agenda Digital recientemente aprobada por el Ayuntamiento de Sant Feliu de Llobregat²⁴, los ejes que, como mínimo, se deberían incorporar como estratégicos son:

- Eje 1: Administración digital. Avanzar en la transformación integral del Ayuntamiento, impulsando una administración digital abierta, inteligente, inclusiva y social; un nuevo marco de relación con la ciudadanía; el cambio cultural y la transformación de la estructura productiva y organizativa; la gobernanza de los datos y la reutilización de la información
- Eje 2: Infraestructuras digitales. Garantizar la conectividad digital y dotar todo el territorio de cobertura de redes de nueva generación, facilitando el desarrollo de tecnologías emergentes

²¹ Ayuntamientos de Sevilla, Las Palmas de Gran Canaria, Sant Feliu de Llobregat, A Coruña...

²² La obra colectiva “La transformación digital de la Administración local”, coordinada por el profesor Agustí Cerrillo i Martínez, y publicada por la Fundación Democracia y Gobierno Local, incorpora una propuesta de agenda de la transformación digital de la Administración local (<https://gobiernolocal.laminadigital.es/la-transformacion-digital-de-la-administracion-local/>)

²³ <https://www.localret.cat/agendadigital/>

²⁴ <https://www.santfeliu.cat/documents/4f5378cc-d80d-444a-9bce-2cbeb046aaff/SDEEXPEDX.pdf> La agenda digital del Ayuntamiento de Sant Feliu de Llobregat se aprobó conjuntamente con el Plan de Acción Local de la Agenda Urbana y constituyen dos instrumentos estratégicos con horizonte 2030 conectados entre sí y alineados con los Objetivos de Desarrollo Sostenible (ODS) de Naciones Unidas.

(5G, IoT, IoS, Blockchain, IA, etc.); disponer de las infraestructuras y soluciones tecnológicas necesarias para la digitalización de la administración y la provisión de servicios inteligentes; impulsar la computación y el almacenamiento en la nube, la ciberseguridad y la mejora de la resiliencia.

- Eje 3: Sociedad digital. Favorecer el apoderamiento digital de la ciudadanía para construir una sociedad digital más justa, cohesionada, inclusiva y participativa, a partir de la capacitación y la formación de la ciudadanía en competencias digitales; utilizar las tecnologías para crear ocupación y favorecer la igualdad y la inclusión social, garantizando los derechos y libertades de la ciudadanía, promoviendo una gobernanza distribuida y generando sinergias para impulsar la inteligencia colectiva.
- Eje 4: Economía digital. Promover la transformación sostenible de las empresas del municipio, el apoderamiento digital de los trabajadores, impulsando un ecosistema urbano de innovación abierto, que fomente la igualdad de género en el ámbito empresarial TIC y la sostenibilidad medioambiental de las empresas.
- Eje 5: Territorio digital y sostenible. Utilizar las tecnologías digitales para la consecución de los objetivos de resiliencia, transición digital y prevención y reducción de los impactos del cambio climático y otros riesgos mediante la transición energética, la movilidad sostenible, la gestión inteligente de los recursos, la optimización de la gestión de los residuos, la mejora de la seguridad, etc.

Más concretamente, el despliegue de la administración digital y la incorporación de tecnologías emergentes, en especial de la Inteligencia Artificial, en la actividad administrativa municipal requerirá, además, definir el modelo de gestión de la entidad local, el modelo de gobernanza de los datos propio de esa institución, así como la metodología y procesos para llevarlo a cabo.

3.1.- Un modelo de gestión inteligente

Es evidente que la creciente incursión de tecnologías emergentes en la actividad administrativa, permite más y mejores capacidades de captación, procesamiento y almacenamiento de datos, que implica, a

su vez, que las Administraciones públicas se doten de las herramientas e infraestructuras necesarias para refundar sus propios cimientos y sus propias expectativas en el ámbito de la prestación de servicios y la generación de valor público.

Cabe recordar que junto a la normativa, agendas y estrategias antes mencionadas, en materia de despliegue de la administración digital, a ello hay que sumar los procesos de apertura mediante la incorporación de un modelo de gobierno abierto que se sustenta en la transparencia, participación y colaboración con la ciudadanía, que está provocando una verdadera revolución interna, cuestionando la arquitectura de los sistemas de información corporativos y las herramientas de gestión y tramitación de procedimientos, más pensadas para una administración tradicionalmente en papel, muy alejada aún de esta nueva cultura de gestión de los datos (tanto los generados por las propias Administraciones públicas como por las personas y empresas a través de sensores o de las actividades que se desarrollan en el municipio) que debe basarse en la calidad y en la transparencia y acceso a la información de forma inmediata, garantizando los derechos y libertades públicas de la ciudadanía.

Es, por tanto, necesario transitar de un modelo tradicional “docu-céntrico” a un modelo “datacéntrico” que permita evolucionar hacia una organización administrativa inteligente, estructurada alrededor de una arquitectura integrada de servicios orientada a procesos ágiles, simplificados, comunes, donde la prioridad se centre en prestar unos servicios públicos de calidad, de forma personalizada, proactiva e, incluso, predictiva, que base su mejora continua en el análisis de datos y la rendición de cuentas permanente. Un modelo de gestión basado en la transparencia, el uso inteligente de los datos, la generación de valor y la innovación a través del empleo de los medios electrónicos que debería contar, como mínimo, con la interrelación de los siguientes elementos imprescindibles:

- Un modelo de datos estructurado, que garantice el dato único y la calidad de los datos, en su máxima expresión.
- Gestión de procesos y documentos, que permita el control de toda la información que se genera o ingresa en la corporación, incorporando los instrumentos necesarios que posibiliten su

gestión (simplificación, automatización, clasificación, conservación y acceso).

- Un Bus de servicios que permita integrar de forma automatizada las diferentes aplicaciones, en base a una arquitectura tecnológica que garantice tanto la interoperabilidad interna como la externa, con el objetivo de construir procesos de extremo a extremo, y garantizar a la ciudadanía derechos como el de no aportar documentación ni datos que ya estén en poder de otras Administraciones públicas; conocer el estado de tramitación de sus asuntos; etc., todo ello en un contexto de confianza, seguridad en las transacciones (cumplimiento requisitos ENS y ENI) y protección de datos personales.

Funcionar sin papeles implica, por tanto, dotarse de instrumentos que permitan controlar el ciclo de vida de la información, documentación y expedientes; pero sobre todo implica precisamente utilizar las tecnologías para transformar estos procesos desde una óptica innovadora con el objetivo de ofrecer servicios más personalizados y conseguir ser más proactivos de cara a la ciudadanía y las empresas.

El modelo debe permitir asegurar la disponibilidad de conjuntos de datos de buena calidad, fiables e interoperables e infraestructuras que lo posibiliten; favoreciendo la generación de valor a partir de los conjuntos de datos e identificando y gestionando un gran volumen de datos. La apertura y democratización de los datos y la información, con una gestión eficiente de los recursos, implica una estrategia de gestión integral de la información (datos) y de los documentos, así como de los procedimientos administrativos que permiten articular la gestión de forma transversal, facilitando una buena política de interoperabilidad, de cara a garantizar la interrelación administrativa tanto en los procesos de incorporación de información como en los que provienen de sistemas de información de terceros.

En definitiva, los elementos esenciales que debe contemplar un Sistema de Información estructurado son, como mínimo, los siguientes:

- Infraestructuras y herramientas orientadas a la simplificación, al desarrollo y apoyo en los procesos de toma de decisiones a todos los niveles: estratégico, operativo y productivo (*Business*

Intelligence, facility management, gestión inteligente de los recursos y la ciudad, reutilización de la información...).

- La gestión de procesos y documentos, entendida como la implementación de herramientas orientadas a la gestión automatizada de los flujos administrativos y los documentos asociados desde la demanda o el impulso por parte de la Administración, hasta su resolución y archivo, en términos de simplificación, eficiencia y eficacia.
- Una arquitectura de sistemas de información flexible, modular y escalable, requerimientos que son necesarios aunque no suficientes para poder concebir escenarios de mejora, integración y crecimiento, sin necesidad de replantear de forma periódica la capacidad productiva del sistema.
- Capacidad de integración, existente en la organización o bajo la modalidad de externalización, pero en cualquier caso en términos de viabilidad sobre los módulos estratégicos, lo que permite la incorporación de soluciones propias o de terceros, aspirando así a una superposición de soluciones que permitan completar el ciclo de la gestión.
- Y, por supuesto, el mencionado modelo de datos, como pieza clave también para construir un núcleo de información corporativa que permitirá desarrollar procesos de interoperabilidad y de agregación de datos

3.2.- *Un modelo de gobernanza de los datos*

La inteligencia artificial se aplica sobre los datos, y esta evolución hacia un modelo de gestión datacéntrico obliga a las entidades locales a establecer un modelo de gobernanza con la finalidad de controlar los datos que se generan e ingresan en la organización, independientemente del origen de la fuente. El objetivo último es consolidar una organización *Data Driven* donde la toma de decisiones estratégicas se fundamente en datos de calidad y, al mismo tiempo, la explotación de estos datos suponga un estímulo para la innovación de los productos y servicios que se prestan a la ciudadanía, posibilitando así la personalización, la proactividad e, incluso, la predicción.

A nivel europeo, la estrategia de datos se está materializando mediante dos instrumentos normativos fundamentales:

- El Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos y por el que se modifica el Reglamento (UE) 2018/1724 (Reglamento de Gobernanza de Datos publicado en DOUE núm. 152, de 3 de junio de 2022²⁵). Este Reglamento tiene como objetivo crear los procesos y estructuras para promover la disponibilidad de datos para su reutilización entre sectores y facilitar el intercambio de datos de forma transfronteriza, aumentando así la confianza en los servicios de intermediación de datos, fomentando la cesión altruista de datos por todo el territorio de la UE, con la finalidad de posibilitar la creación de espacios comunes de datos interoperables en sectores estratégicos.
- La Ley de Datos (*Data Act*), cuya propuesta fue adoptada el 23 de febrero de 2022 por parte de la Comisión Europea [COM (2022) 68 final], que aún está en tramitación y que pretende armonizar todo lo relativo al acceso justo y al uso equitativo de los datos.

Se trata pues de dos normas complementarias que tienen como objetivo crear un mercado único que facilite la libre circulación de datos dentro de la UE y entre sectores en beneficio de las empresas, investigadores, Administraciones públicas y sociedad en general, siempre de acuerdo con los principios y valores de la UE.

Si tomamos como posible definición de gobernanza de los datos la que sugiere el *Data Governance Institute*: “*a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods.*”, se puede concluir que es objeto de la gobernanza de los datos establecer en las organizaciones aspectos relacionados con:

²⁵ <https://www.boe.es/buscar/doc.php?id=DOUE-L-2022-80835>

- La estructura organizativa de la propia gobernanza
- Políticas, estándares, directrices y protocolos para toda la organización
- Roles y responsabilidades

En cuanto a los principios que sustentan la gobernanza de los datos, conviene que éstos estén alineados, por un lado, con los principios que rigen la administración digital como por ejemplo:

- Personalización en la prestación de servicios públicos. La entidad local se adapta a las necesidades específicas de cada persona para la prestación proactiva de servicios.
- Eficacia y eficiencia. La actuación administrativa municipal se lleva a cabo de forma ágil, eliminando tareas superfluas, con el menor coste posible y satisfaciendo las necesidades reales de la ciudadanía.
- Inclusión. Se crean las condiciones materiales e institucionales necesarias para favorecer la participación de todos los grupos a fin de promover y garantizar la diversidad y la inclusión efectiva en el ámbito municipal.
- Transparencia. La entidad local actúa de forma proactiva para dar a conocer la información relativa a sus ámbitos de actuación, con carácter permanente y actualizado, a través de los canales de difusión más adecuados.
- Datos abiertos y reutilización de la información del sector público. Se promueve la apertura de datos públicos para su reutilización por parte de la ciudadanía, las empresas y otras administraciones.
- Calidad en la prestación de los servicios públicos. La entidad local implementa un sistema de gestión de la calidad para la prestación de los servicios públicos que ofrece, a partir de la elaboración y publicación de cartas de servicios y de la evaluación permanente y mejora continua de los servicios que presta.
- Colaboración y cooperación interadministrativas. Se establecen y promueven mecanismos de cooperación y trabajo en red con otras administraciones.

- Participación ciudadana. Se potencia la participación ciudadana en la toma de decisiones públicas a fin de apoderar e implicar a la ciudadanía en los objetivos colectivos.

Por otra parte, principios rectores que aseguran una adecuada gestión y uso de los datos, como:

- Valor estratégico de los datos. Los datos municipales son un activo de valor estratégico que tiene un alto impacto para la ciudadanía y para la toma de decisiones públicas basadas en estos datos.
- Responsabilidad y compromiso ético institucional. Se instaaura una gestión y uso ético y responsable de los datos, que se integra en un sistema más amplio de integridad institucional.
- Conocimiento compartido. La entidad local facilita, a través de los canales de difusión más adecuados, que la ciudadanía pueda acceder de forma directa a los datos municipales, entendiendo que esta transparencia generalizada incrementa las posibilidades de trabajar de forma colaborativa.
- Orientado a resultados. Los datos municipales son un medio para asegurar la transparencia y la rendición de cuentas. También son un recurso para afianzar la gestión y mejora continua de los servicios públicos, el bienestar de la ciudadanía y la consecución de los objetivos municipales.
- Trazabilidad. Se establecen auditorías periódicas con el fin de asegurar que los procesos y actividades sujetas al modelo de gobernanza de los datos se efectúan de acuerdo a las políticas y normas establecidas.
- Estándares abiertos. Se difunde la información en formatos reutilizables, susceptibles de tratamiento informatizado en formatos legibles por máquina que sean abiertos e interoperables, de acuerdo al catálogo de estándares de la Norma Técnica de Interoperabilidad.
- Seguridad integral. Se establecen medidas y políticas de seguridad a fin de minimizar los riesgos a los que están sometidos los datos públicos. Asimismo, en aplicación del Esquema Nacional de Seguridad, la seguridad se concibe como un proceso integral

constituido por todos los elementos técnicos, humanos, materiales y organizativos vinculados al sistema.

- Privacidad desde el diseño. Se garantiza la protección de los datos personales durante todo su ciclo de vida de forma comprobable y aplicando los principios y medidas técnicas y organizativas necesarias, en cumplimiento de la normativa en materia de protección de datos.
- Gestión de la calidad de los datos. La entidad local vela por la calidad de los datos en todo su ciclo de vida para que sean adecuados para su fin y cumplan con los requisitos establecidos. El aseguramiento de la calidad de los datos municipales implica la implementación de un sistema de mejora continua.
- Control del ciclo de vida de los datos. Los datos municipales disponen de un ciclo de vida, que puede variar según la tipología de cada dato. Desde la entidad local se gestiona este ciclo, aplicando procesos y acciones concretas para garantizar el uso, disponibilidad, accesibilidad y preservación de los datos.
- Estandarización. La entidad local se compromete a trabajar en la estandarización de términos y en la unificación de la estructura de datos a fin de asegurar modelos de datos comunes que permitan la integración y la interoperabilidad, sujeta a los requerimientos técnicos y legales del momento.

Por lo tanto, mediante la gobernanza de los datos se trata de establecer el marco decisorio y estratégico que debe guiar las demás funciones de la gestión de los datos, tal y como se visualiza en la rueda DAMA evolucionada:

Figura 1. Rueda DAMA evolucionada (DAMA²⁶ International, 2017)

Asimismo, según el punto de partida de cada entidad local, construir un modelo propio de gobernanza de datos deberá tener en consideración, al menos, los siguientes aspectos:

- El marco normativo vigente y en construcción
- La estructura organizativa
- Los instrumentos necesarios para el gobierno del dato
- La gestión del ciclo de vida de los datos

26

De entre todos los modelos de referencia, el DAMA es probablemente el más extendido. “*DAMA International's Guide to the Data Management Body of Knowledge (DAMA-DMBOK Guide)*” es un documento que recopila procesos y áreas de conocimiento generalmente aceptados como mejores prácticas de la gobernanza de los datos.

- La arquitectura y modelado de los datos
- La gestión de los metadatos
- La calidad de los datos
- La seguridad, privacidad y ética de los datos

Por último, y no menos importante, la gobernanza de los datos requiere revisar, definir y, en su caso, incorporar nuevos perfiles profesionales y funciones que deben velar por el correcto funcionamiento del modelo, y que deberá adaptarse a la estructura organizativa de cada entidad local (en ocasiones, como está ocurriendo por ejemplo con la función de Delegado/a de Protección de Datos, en ciertos ayuntamientos pequeños algunas de estas funciones se podrían asumir por entidades supramunicipales). En todo caso, las funciones identificadas como necesarias son:

- *Chief Data Officer* (CDO): responsable de asegurar la correcta administración y gestión de los datos y de la información a lo largo de su ciclo de vida
- Delegado/a de Protección de Datos (DPD): figura de obligado cumplimiento según la normativa vigente
- Propietarios de los datos: deben actuar como responsables de una serie de sets (o grupos) de datos, desde su administración, control, calidad, acceso y gobernanza
- Gestor de explotación de datos: debe asegurar el conocimiento de los datos de la organización respecto a sus ámbitos de actuación, para que los requerimientos de gestión y explotación sean identificados de la forma más efectiva y productiva posible
- Ingeniero/a – Científico/a de datos: tendrán que dar soporte a la preparación de los datos y modelos analíticos, su catalogación en fuentes de datos existentes, y en el acceso a fuentes de datos internas y externas
- Equipos de administración de datos (personal técnico de administración de datos)

3.3.- *Una metodología*

La transformación digital, en un marco de gestión inteligente y de gobernanza de los datos, además del cumplimiento normativo y de los elementos necesarios que hemos visto anteriormente, requiere también dotarse de normativa interna, políticas, protocolos y metodologías que faciliten este cambio cultural. Asimismo, la incorporación progresiva de las tecnologías disruptivas, en especial de la Inteligencia Artificial, precisa también que las entidades locales adopten una metodología que, de forma innovadora y transversal, facilite el desarrollo de pruebas piloto que permitan ensayar el uso de dichas tecnologías y evaluar los efectos e impactos de su uso en la prestación de los servicios, tanto internamente como por parte de las personas usuarias del servicio.

Esta metodología puede tener un esquema clásico en base a cuatro fases:

- Análisis: identificación y priorización de los servicios y procesos en los que es conveniente aplicar la IA; analizar experiencias similares; normativa vigente; alcance, ámbitos afectados y dimensiones del proyecto...
- Diseño: constitución de un equipo de trabajo multidisciplinar (colaboración público-privada); identificación de fuentes de datos afectados; recursos tecnológicos, presupuestarios y humanos; objetivos, acciones y propuestas de mejora, tanto organizativos como procedimentales; análisis de los riesgos y evaluación del impacto, si procede; incorporación de la ciudadanía para testear la propuesta de diseño...
- Implementación: distribución de tareas entre perfiles técnicos (depuración de datos, diseño o adquisición de los algoritmos necesarios, herramientas de gestión y control...); perfiles jurídicos y organizativos (normativa interna, procedimientos, metodologías, protocolos, instrucciones, formación, comunicación, participación...); desarrollo de pruebas piloto...
- Seguimiento y evaluación: garantizando la transparencia de todo el proceso y su trazabilidad; estableciendo indicadores que permitan evaluar el impacto y efectividad de la incorporación de la IA; la valoración de la ciudadanía...

4.- POSIBLES USOS DE LA INTELIGENCIA ARTIFICIAL Y DESAFÍOS

Es innegable que la inteligencia artificial está ya presente en muchos de los ámbitos de nuestra vida cotidiana y que, sin duda, lo estará mucho más en el futuro. Para las Administraciones públicas la incorporación de esta tecnología supone una oportunidad histórica de cambio de modelo radical que materialice la eterna reforma tantas veces anunciada y diagnosticada, pero con escasos resultados desde el punto de vista de la innovación en la prestación de los servicios públicos.

La inteligencia artificial puede ser utilizada en diferentes estadios y grados, desde la automatización de tareas rutinarias, procesos de análisis de grandes cantidades de datos, hasta la posibilidad de toma de decisiones públicas de forma automatizada, prestación de servicios públicos personalizados y proactivos, realizar simulaciones, predicciones, etc. Sin olvidar el debate actual lógico y absolutamente necesario respecto a qué tipo de decisiones se pueden tomar basándose en algoritmos, y que sin duda está cuestionando los cimientos sobre los que se sustenta el Estado democrático y de derecho, resulta conveniente que las entidades locales aborden su proceso de transformación digital incorporando acciones que contemplen, al menos, la introducción de esta tecnología de forma progresiva, mediante pruebas piloto que les permita, por un lado, testear sus sistemas de gestión, y por otro, acercar al personal de su organización al conocimiento de su potencial.

Es evidente que queda mucho camino por recorrer en esta materia, y que es fundamental el esfuerzo tanto académico como gubernamental²⁷ por intentar dar respuesta a los innumerables riesgos ya planteados que puede suponer el uso de esta tecnología.

En todo caso, una de las primeras acciones que sería conveniente abordar por parte de las entidades locales sería identificar, para cada una de las fases de la tramitación del procedimiento administrativo,

²⁷ En este sentido, la propuesta de Reglamento sobre inteligencia artificial [COM(2021) 206 final], introduce un enfoque basado en el riesgo que diferencia entre los usos de la inteligencia artificial que crean un riesgo inaceptable, los que generan un riesgo alto y aquellos que entrañan un riesgo bajo o mínimo.

qué actuaciones pueden realizarse de forma automatizada, es decir, sin intervención humana²⁸. Ello permitiría, como mínimo, simplificar la tramitación administrativa en ciertas actuaciones de trámite (numeración de resoluciones y decretos; expedición de copias auténticas y de recibos en el registro electrónico; consultas internas y a otras Administraciones públicas; transcripción a libros de resoluciones y actas de órganos de gobierno, la emisión de certificaciones, la realización de notificaciones de resoluciones y acuerdos, el índice y foliado de expedientes, la expedición de volantes y certificados de empadronamiento, etc.)

Por otro lado, disponer de una administración plenamente digital y un modelo de gestión basado en datos debidamente estructurados y de calidad, hace posible comenzar a experimentar la reformulación de ciertos procesos (como las ayudas y subvenciones, matriculaciones, ciertas ocupaciones de la vía pública, etc.) para aplicar los principios de la personalización y la proactividad. En efecto, ya hay experiencias locales en las que ciertos trámites realizan comprobaciones automáticas tanto de los datos que dispone la propia administración como de las consultas realizadas en otras bases de datos, mostrando a la ciudadanía la información para su validación o modificación.

Hay incluso entidades locales que han implementado herramientas predictivas de situaciones de vulnerabilidad en el ámbito de los servicios sociales, mediante el análisis de diferentes fuentes de datos y tendencias sociales, identificando perfiles poblacionales en el territorio, relacionándolo con la información disponible, y, aplicando una serie de algoritmos de aprendizaje automático (*machine learning*), lo que permite realizar simulaciones en tiempo real y redimensionar el servicio.

Otro posible uso de la inteligencia artificial en las entidades locales podría ser la realización de análisis de riesgos (por ejemplo en la función interventora) que permitiría la generación de forma automatizada de informes de control y automatizar también la planificación de futuras auditorías internas.

²⁸ Artículo 41.1 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público.

Otras experiencias interesantes en el ámbito local se centran en la atención ciudadana, mediante la incorporación de *chatbots* o robots conversacionales que gracias a la inteligencia artificial pueden dar respuesta en tiempo real a cierto tipo de demandas ciudadanas.

Es destacable también, en el ámbito de la gestión y mantenimiento de edificios, los sistemas inteligentes basados en la metodología BIM²⁹ (*Building Innovation Management*), que proporciona una nueva forma de modelar y gestionar todo el proceso edificatorio mediante herramientas que incorporan la inteligencia artificial combinada con otras tecnologías disruptivas, integrando la definición arquitectónica y constructiva, el cálculo de instalaciones, las mediciones y los presupuestos, la planificación, modificación y gestión de obra y la gestión del mantenimiento durante la vida útil del edificio en un único archivo digital, evitando problemas de descoordinación al trabajar directamente sobre una realidad virtual de la edificación global y garantizando un gran ahorro de costes en todo el proceso (Alguacil y Moro, 2020).

Las entidades locales también deben ser conocedoras de los riesgos que la implementación de la inteligencia artificial puede suponer para los derechos y libertades de las personas. Sin duda, el uso de algoritmos para la toma de decisiones exige poder auditarlos y conocer en qué se ha basado esa decisión y si se garantizan los derechos y la seguridad jurídica de las personas interesadas. Por otro lado, la necesidad de disponer de datos reales que “alimenten” esa inteligencia artificial, puede tener como consecuencia la utilización de datos con un propósito distinto e incompatible con la finalidad por la que fueron recabados. Ello hace imprescindible asegurarse que se dispone de la legitimación para dicho tratamiento, y si ésta se basa en el consentimiento de los usuarios, debe realizarse con la máxima transparencia para que las personas usuarias sean plenamente conscientes del uso de sus datos y de los mecanismos para ejercer sus derechos³⁰.

²⁹ <https://www.mitma.gob.es/el-ministerio/sala-de-prensa/noticias/mar-26012021-1009>

³⁰ Como material de apoyo en materia de protección de datos, es recomendable el documento elaborado por la AEPD <https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia.pdf>

En todo caso, la incorporación de la inteligencia artificial requiere, sin duda, la aplicación del principio de prudencia, el conocimiento de las directrices éticas impuestas desde Europa en lo que tiene que ver con el desarrollo y uso de la inteligencia artificial para que ésta sea fiable, que se sustentan en tres componentes que deben estar presentes a lo largo de todo su ciclo de vida: la inteligencia artificial debe ser lícita, ética y robusta, tanto desde el punto de vista técnico como social, para evitar, entre otras consecuencias, los potenciales daños accidentales.

5.- CONCLUSIONES

Es evidente que nos estamos enfrentando a un cambio de paradigma sin precedentes donde el papel de la tecnología, en especial la inteligencia artificial, suscita intensos debates y posturas maniqueas respecto a lo que nos deparará el futuro y los riesgos que ciertos usos de la inteligencia artificial pueden suponer como herramienta para el control social y para sustentar regímenes totalitarios, además del posible impacto en la propia evolución de la humanidad en un mundo hiperconectado³¹.

Ante este inmenso desafío, el papel de lo público se erige como absolutamente primordial y determinante para garantizar derechos y libertades en un entorno donde cada vez es más difícil distinguir entre lo físico y lo virtual.

Por otro lado, es indiscutible que la tecnología supone una palanca de cambio fundamental para transitar hacia una Administración pública más eficiente, abierta, íntegra, inclusiva, accesible y sostenible, adaptando la prestación de los servicios públicos a las nuevas demandas y retos de la sociedad.

Por ello, las entidades locales, como el resto de Administraciones públicas, deben diseñar una hoja de ruta que, alineada con las agendas y estrategias globales, estatales y autonómicas, les permita gobernar todo este proceso de transformación integral desde, al menos,

³¹ <https://digitalfuturesociety.com/es/agenda/digital-future-society-summit-2021-humanism-in-digital-age/>

una triple perspectiva: el buen gobierno y la buena administración, el nuevo modelo de gestión pública y de relación con la ciudadanía, y la gestión inteligente y sostenible de la ciudad.

Al final de lo que se trata es de saber qué hacer con esta tecnología, qué implicaciones tiene a nivel jurídico, técnico y organizativo, cómo se debe hacer, y qué medidas se deben tomar para minimizar los posibles riesgos asociados.

Porque la inteligencia artificial por sí sola no va a solucionar los problemas endémicos de las Administraciones públicas, sino se acompaña de un cambio cultural progresivo, de una nueva cultura de gestión de los datos a los que, convenientemente gobernados, podamos aplicar inteligencia para generar valor público para la sociedad.

En definitiva, un viaje apasionante que las entidades locales deben abordar lo antes posible transformando sus cimientos en beneficio de la sociedad y de la ciudadanía.

BIBLIOGRAFÍA CITADA

- Agencia Española de Protección de Datos. (2020). “Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción”.
- ALGUACIL SANZ, M. (2014). Administración, administración electrónica, administración digital. Archivamos: Boletín ACAL, 94, 26-27.
- ALGUACIL SANZ, M. (2015). “De la gestión inteligente al compromiso de servicio”. Archivamos: Boletín ACAL, 97, 39-40.
- ALGUACIL SANZ, M. (2017). “La transformación digital en el sector público”. El Consultor de los Ayuntamientos, 5, 563-571.
- ALGUACIL SANZ, M. y MORO CORDERO, M.^a A. (2020). “Cuatro décadas de cambio culminadas en dos meses: la experiencia del Ayuntamiento de Sant Feliu de Llobregat”. En V. Almonacid Lamelas (ed.). Guía práctica para la implantación de la administración electrónica en el ámbito local. Madrid: Wolters Kluwer.
- ALGUACIL SANZ, M., GONZÁLEZ MAYA, M. y REAL OSUNA, R. (2021). “SmartCityMAP, la plataforma para la gestión inteligente del territorio de Sant Feliu de Llobregat”. En M. Á. de Bas Sotelo (ed.). Buenas prácticas en la innovación pública (50+1 Experiencias para la transformación). Madrid: Wolters Kluwer.
- CERRILLO I MARTÍNEZ, A. y MORO CORDERO, M.^a A. (2021). “La transformación digital de la Administración local”, en CERRILLO I

- MARTÍNEZ, A. (coord.) Barcelona: Fundación Democracia y Gobierno Local
- CERRILLO I MARTÍNEZ, A. y MORO CORDERO, M^a A. (2022). “El Reglamento de Gobernanza de Datos y su impacto en las Administraciones públicas”. Revista El Consultor de los Ayuntamientos. Madrid: Wolters Kluwer
- “La gobernanza inteligente: una transformación multinivel y multidimensional en tiempos de metaverso” (2022) en COLÓN DE CARVAJAL FIBLA, B., MAGRO PEDROVIEJO, R. y MORO CORDERO, M^a A. (coord.). Revista El Consultor de los Ayuntamientos. Madrid: Wolters Kluwer
- MANZANO PELÁEZ, E. (2022) “*La ciutadania al centre? Comencem per l'escolta activa*”. Blog EAPC
- PANIAGUA ALARIO, A. M^a. (2021) “Inteligencia Artificial en el sector público (parte II)”. Blog PRODAT
- PONCE SOLÉ, J. (2019a). “Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico”. Revista General de Derecho Administrativo, 50.
- PONCE SOLÉ, J. (2019b). “La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida”. Madrid: Universidad de Alcalá-Defensor del Pueblo.

QUINTA PARTE:
DERECHO EUROPEO Y
PROPUESTAS REGULATORIAS
DE FUTURO

Capítulo XVI

Algoritmos, inteligencia artificial y procedimiento administrativo: principios comunes en el Derecho de la Unión Europea

ORIOI MIR PUIGPELAT¹

*Catedrático de Derecho Administrativo
Universidad Pompeu Fabra de Barcelona*

Resumen: El presente trabajo examina y subraya el papel crucial que corresponde a la institución clásica del procedimiento administrativo en el debate actual sobre el uso responsable de los algoritmos en el sector público. Propone algunas distinciones fundamentales relativas a la utilización de los algoritmos en el marco del procedimiento y sugiere introducir determinadas adaptaciones procedimentales para evitar los riesgos que dicho uso puede comportar. Se presta una especial atención a los principios que cabe extraer del Derecho de la Unión y a la actuación de la Administración europea, al hilo de un estudio de campo que se ha llevado a cabo en los últimos meses y que se resume en el Anexo.

Abstract: This paper examines and underlines the crucial role of the classical institution of administrative procedure in the current debate on the responsible use of algorithms in the public sector. It proposes some basic distinctions concerning the use of algorithms within the administrative procedure and suggests certain procedural adaptations in order to avoid the risks that such use may entail. Particular attention is paid to the principles to be drawn from EU law and the activities of the European administration, in the light of a case study carried out in recent months and summarised in the Annex.

Palabras-clave: Inteligencia artificial, algoritmos, automatización, procedimiento administrativo, Derecho de la Unión Europea, Administración de la Unión Europea

Key-words: Artificial intelligence, algorithms, automation, administrative procedure, European Union law, European Union Administration

SUMARIO.- 1. INTRODUCCIÓN.- 2. EL PAPEL CENTRAL DEL PROCEDIMIENTO ADMINISTRATIVO EN LA ADMINISTRACIÓN ANALÓGICA Y DIGITAL.- 2.1. Actos administrativos.- 2.2. Normas reglamentarias.- 3. DISTINCIONES RELEVANTES RELATIVAS AL USO DE ALGORITMOS EN EL MARCO DEL PROCEDIMIENTO ADMINISTRATIVO.- 3.1. Distinción en función de la fase del procedimiento en que se emplea el algoritmo y de su grado de incidencia sobre la decisión

¹ ORCID: 0000-0001-5807-4748. Trabajo elaborado en el marco de los Proyectos AEI-PID2019-108274GB-I00 / AEI / 10.13039/501100011033 y PCI2020-112207 / AEI / 10.13039/501100011033.

final.- 3.2. Distinción en función del tipo de algoritmo empleado.- 3.3. Distinción en función del carácter favorable o desfavorable de la decisión automatizada.- 3.4. Distinción en función del carácter discrecional o reglado de la decisión automatizada.- 4. ADAPTACIONES PROCEDIMENTALES NECESARIAS EN CASO DE UTILIZACIÓN DE ALGORITMOS.- 4.1. La necesidad de informar adecuadamente a los interesados y al público sobre los sistemas automatizados empleados.- 4.2. El establecimiento de un principio de supervisión humana.- 4.3. La necesidad de realizar evaluaciones de impacto antes y después de automatizar la toma de decisiones administrativas.- 5. ANEXO: EL USO DE IA POR PARTE DE LA ADMINISTRACIÓN DE LA UNIÓN: UN ESTUDIO DE CAMPO.- 5.1. Finalidad del estudio y casos examinados.- 5.1.1. Caso 1 (DG-Agri/ESA).- 5.1.2. Caso 2 (EFSA).- 5.1.3. Caso 3 (EUIPO).- 5.1.4. Caso 4 (eu-LISA).- 5.2. Algunas conclusiones que cabe extraer del estudio de campo realizado. BIBLIOGRAFÍA CITADA

1.- INTRODUCCIÓN

La automatización de la actuación administrativa no es un fenómeno nuevo. En su tesis doctoral de 1966 Luhmann ya pronosticó el gran impacto que tendría para la organización administrativa², y la Ley federal alemana de procedimiento administrativo le dedicó algunos preceptos relevantes en su versión inicial de 1976³.

El gran interés que su regulación jurídica suscita en la actualidad responde a dos grandes factores. El primero, por supuesto, es el importante desarrollo tecnológico producido en los últimos años en el ámbito de la computación, que ha abaratado y multiplicado las capacidades de automatización tanto en el sector público, como privado, y que ha posibilitado el salto cualitativo que suponen los algoritmos de *machine learning* o aprendizaje automático.

El segundo factor, ligado al anterior, son los riesgos asociados a este cambio cuantitativo y cualitativo, y que se han materializado ya en numerosos episodios de mal funcionamiento administrativo. Tales episodios han tenido lugar tanto en Europa⁴ como, sobre todo, en

² LUHMANN, N. (1966): *Recht und Automation in der öffentlichen Verwaltung*, Duncker & Humblot, Berlin.

³ §§ 28.2.4, 37.4 y 39.2.3 *Verwaltungsverfahrensgesetz* (VwVfG).

⁴ Especialmente conocido es el caso del sistema holandés SyRI de detección del fraude en la solicitud de prestaciones sociales, declarado ilegal por el Tribunal de Distrito de La Haya mediante sentencia de 5 de mayo de 2020

los Estados Unidos⁵, cuyas agencias han apostado fuertemente por la automatización, a menudo para reducir costes de personal. Con frecuencia han comportado errores masivos y difícilmente detectables en la reducción de prestaciones sociales a colectivos desfavorecidos. La proliferación de dichos episodios ha llegado a poner en cuestión la propia legitimidad del Estado administrativo⁶.

En el presente trabajo se subrayará la importante contribución que corresponde a la institución clásica del procedimiento administrativo en el debate actual sobre el uso responsable de los algoritmos (*algorithmic accountability*) en el sector público. Se defenderá que el procedimiento administrativo constituye un instrumento crucial para lograr un desarrollo adecuado de la automatización y la inteligencia artificial en la toma de decisiones administrativas, que extraiga su gran potencial sin merma de los derechos de los ciudadanos afectados.

Se prestará una especial atención a los principios que cabe extraer del Derecho de la Unión y a la actuación de la Administración europea (integrada principalmente por la Comisión y las agencias europeas), al hilo de un estudio de campo que se ha llevado a cabo en los últimos meses y que se resume en el Anexo. Y se explorarán las posibilidades de mejora de la Propuesta de Reglamento sobre Inteligencia Artificial (RIA) que está tramitando actualmente el legislador de la Unión⁷ con algunas sugerencias que podrían incorporarse a un nuevo Título específico sobre el uso de algoritmos avanzados por parte de la Administración europea.

(ECLI:NL:RBDHA:2020:1878). Vid. <https://algorithmwatch.org/en/story/syri-netherlands-algorithm/> (última visita: 31/08/2022).

⁵ Vid. los numerosos casos documentados por CITRON, D. K. (2008), “Technological Due Process”, *Washington Law Review*, n° 85, pp. 1249 ss. (1256 ss., 1267 ss.); CALO, R.; CITRON, D. K. (2021), “The Automated Administrative State: A Crisis of Legitimacy”, *Emory Law Journal*, n° 70, pp. 797 ss. (799 ss., 818 ss.), relacionados principalmente con agencias estatales.

⁶ CALO, R.; CITRON, D. K. (2021), *op. cit.*, *passim*.

⁷ Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la unión, COM/2021/206 final, de 21 de abril de 2021.

2.- EL PAPEL CENTRAL DEL PROCEDIMIENTO ADMINISTRATIVO EN LA ADMINISTRACIÓN ANALÓGICA Y DIGITAL

El procedimiento administrativo ha constituido una garantía básica del Estado de Derecho en el marco de la Administración analógica y debe seguir siéndolo en el contexto de una Administración digital, tanto en la adopción de actos administrativos como de normas reglamentarias.

2.1.- *Actos administrativos*

Como es sabido, hace décadas que las Administraciones públicas de ambos lados del Atlántico deben observar una serie de trámites y garantías procedimentales antes de adoptar decisiones que puedan tener efectos jurídicos relevantes sobre ciudadanos y personas jurídicas particulares, tales como la imposición de sanciones y prohibiciones, la concesión y retirada de autorizaciones y subvenciones, la expropiación forzosa de sus bienes por razones de interés público, etc. (decisiones singulares o actos administrativos). Tanto en los Estados Unidos, como en la mayoría de países de la Unión, existen incluso leyes generales de procedimiento administrativo que codifican sus principios rectores.

El Derecho de la Unión ha reforzado la importancia del procedimiento administrativo de adopción de decisiones singulares a través de dos vías principales. Por un lado, mediante la aprobación de una gran cantidad de normas sectoriales de Derecho derivado que imponen detallados procedimientos tanto a la Administración de la Unión, como a las de los Estados miembros (procedimientos que a menudo tienen una fase nacional y otra comunitaria –procedimientos compuestos–); por otro, a través de los principios generales del Derecho de la Unión elaborados a lo largo de los años por el Tribunal de Justicia, y que deben ser también observados tanto por la Administración de la Unión como por las de los Estados miembros cuando ejecutan el Derecho europeo⁸. Algunos de dichos principios han sido

⁸ GALETTA, D.-U.; HOFMANN, H. C. H.; MIR, O.; ZILLER, J. (2015), “The General Principles of EU Administrative Procedural Law. An in-depth analysis”,

consagrados al máximo rango normativo al pasar a integrar el derecho fundamental a una buena administración del art. 41 de la Carta de los Derechos Fundamentales de la Unión (CDFUE). Por su peso y tradición destacan, en particular, las garantías contenidas en el apartado 2 de dicho precepto: el *derecho de audiencia* de los interesados, su *derecho de acceso al expediente* y el *deber de la Administración de motivar* sus decisiones. De especial importancia en la jurisprudencia del TJUE resulta también el *deber de diligencia* en la tramitación del procedimiento, y que obliga a investigar y examinar de forma minuciosa e imparcial todos los elementos relevantes del asunto, también los que sean favorables a los interesados. Este deber ha sido erigido por el Tribunal en *principio general del Derecho de la Unión* al considerarlo inherente al principio de buena administración⁹.

El especial rango normativo de los principios generales del Derecho de la Unión y del art. 41 CDFUE comporta que tales garantías procedimentales deban ser observadas por toda autoridad europea o nacional que pretenda automatizar la toma de decisiones con efectos jurídicos sobre personas concretas en ejecución del Derecho de la Unión. Es más: dado su rango constitucional, incluso el propio legislador de la Unión y de los Estados miembros deben respetarlas cuando habiliten tal automatización.

Estas garantías son fundamentales para evitar los riesgos de mal funcionamiento de los sistemas automatizados a que antes se ha hecho referencia. Así lo ha entendido también la doctrina norteamericana, que ha subrayado la importancia de observar escrupulosamente las garantías del *due process* derivadas de las enmiendas V y XIV de la Constitución de los Estados Unidos, así como las reglas procedimentales más concretas que contienen la Ley federal de procedimiento administrativo de 1946 y sus homólogas estatales¹⁰.

Rivista italiana di diritto pubblico comunitario, n° 5, pp. 1421 ss.

⁹ STJUE de 4 de abril de 2017, Defensor del Pueblo c. Claire Staelen, C-337/15 P, ECLI:EU:C:2017:256, apartados 34 ss.

¹⁰ CITRON, D. K. (2008), *op. cit.*, pp. 1278 ss.; MULLIGAN, D. K.; BAMBERGER, K. A. (2019), "Procurement as Policy: Administrative Process for Machine Learning", *Berkeley Technology Law Journal*, n° 34, pp. 773 ss. (pp. 801 ss.), formulando propuestas para asegurar su cumplimiento en el caso de decisiones automatizadas o semiautomatizadas.

Los casos de uso de algoritmos de *machine learning* por parte de la Administración de la Unión que se han identificado en el marco del estudio de campo realizado permiten constatar dicha afirmación.

En el caso de la utilización de herramientas de monitorización por satélite de los cultivos europeos¹¹, el deber de diligencia obliga a la Administración a comprobar de oficio –a través de una revisión humana de las imágenes o, en su caso, de una inspección sobre el terreno– las alertas sobre posible incumplimiento de la normativa de subvenciones que emita el sistema, y el derecho de audiencia de los agricultores posibilita que puedan señalar y rebatir los errores que se produzcan antes de que se les revoque una subvención previamente concedida.

En el caso del registro de marcas por parte de EUIPO¹², el citado deber de diligencia obliga al funcionario o funcionarios que deciden el registro a revisar la propuesta formulada por el sistema de IA, a la luz de los argumentos invocados por el solicitante y quien se opone al registro solicitado. A su vez, el deber de motivación obliga a dichos funcionarios a fundamentar las razones de su decisión y posibilita su impugnación por el solicitante o el oponente ante las Salas de Recurso de EUIPO y, en su caso, ante el Tribunal General de la Unión.

Finalmente, el examen de los sistemas de información operados por eu-LISA¹³ ilustra cómo el Reglamento del ETIAS toma en consideración las referidas garantías procedimentales, p. ej., al obligar a Frontex a verificar que no se haya producido ningún error cuando el sistema arroje una respuesta positiva (*hit*) que impida la concesión automática de la autorización de viajar al territorio de la Unión, al exigir a la autoridad nacional competente que motive la denegación de dicha autorización y a que queden registradas en el expediente de la solicitud todas las operaciones realizadas por el sistema automatizado.

¹¹ Vid. Anexo, caso 1.

¹² Vid. Anexo, caso 3.

¹³ Vid. Anexo, caso 4.

2.2.- Normas reglamentarias

El procedimiento administrativo y sus garantías no se exigen solo respecto de la adopción de decisiones singulares dotadas de efectos jurídicos. En muchos ordenamientos existe también una regulación más o menos prolija del procedimiento que debe seguir la Administración para aprobar normas generales-abstractas que desarrollen las previsiones legales.

El supuesto paradigmático es el de Ley federal norteamericana de procedimiento administrativo, que establece desde hace casi ochenta años un procedimiento de *notice and comment* que ha tenido gran influencia y que permite la participación de la ciudadanía y de los sectores afectados durante la elaboración de normas por parte de las agencias federales. A diferencia de lo que ocurre con las decisiones singulares, el Derecho de la Unión no contempla un precepto equivalente al art. 41 CDFUE que imponga de forma general obligaciones procedimentales a las Administraciones nacionales cuando aprueben normas que desarrollen las europeas. Tampoco hay previsiones de Derecho originario o derivado que impongan tales exigencias a la Comisión, más allá de las obligaciones de consulta del Parlamento Europeo, del Consejo y de los Estados miembros que se desprenden de los arts. 290 y 291 TFUE y del Reglamento sobre comitología¹⁴. Sin embargo, la Comisión lleva ya años observando reglas procedimentales destacables en la elaboración de sus propuestas legislativas y de (solo) algunos de sus actos delegados y de ejecución. Se trata, fundamentalmente, de amplias *consultas* a la ciudadanía y los sectores afectados, y de *evaluación ex ante* y *ex post* de los distintos impactos económicos, sociales y medioambientales de las normas propuestas y aprobadas¹⁵. Se ha comprometido formalmente a ello en el Acuerdo interinstitucional sobre la mejora de la legislación de 2016¹⁶.

¹⁴ Reglamento (UE) n° 182/2011 del Parlamento Europeo y del Consejo, de 16 de febrero de 2011, por el que se establecen las normas y los principios generales relativos a las modalidades de control por parte de los Estados miembros del ejercicio de las competencias de ejecución por la Comisión.

¹⁵ https://ec.europa.eu/info/law/law-making-process/planning-and-proposing-law/better-regulation-why-and-how_es (última visita: 31/08/2022).

¹⁶ Acuerdo interinstitucional entre el Parlamento Europeo, el Consejo de la Unión Europea y la Comisión Europea sobre la mejora de la legislación (DO L 123,

Tales trámites procedimentales pueden ayudar a detectar errores que se produzcan como consecuencia de la utilización de sistemas algorítmicos en algún momento del proceso de gestación de nuevas normas. Puede traerse aquí a colación el caso de estudio de la automatización parcial del proceso de revisión sistemática de la literatura científica existente acometido por la EFSA cuando emite un informe científico evaluando un determinado riesgo¹⁷. Cuando se efectúe dicha revisión sistemática en el marco de un procedimiento normativo (para decidir, p. ej., si se prohíbe o no el uso alimentario de una determinada sustancia), el hecho de que el informe científico se integre en la documentación sometida a consulta pública facilita que se detecten eventuales omisiones que pueda contener (como trabajos científicos previos y relevantes que no hayan sido considerados en su evaluación de riesgos, debido a un error en el cribado de la literatura efectuado por el algoritmo).

3.- DISTINCIONES RELEVANTES RELATIVAS AL USO DE ALGORITMOS EN EL MARCO DEL PROCEDIMIENTO ADMINISTRATIVO

Cuando se trata de identificar los requerimientos jurídicos relativos al uso de algoritmos en el marco de procedimientos regidos por el Derecho de la Unión resulta de utilidad efectuar ulteriores distinciones, más allá de la realizada entre procedimientos de adopción de actos singulares y de normas administrativas.

3.1.- Distinción en función de la fase del procedimiento en que se emplea el algoritmo y de su grado de incidencia sobre la decisión final

El impacto jurídico del uso de algoritmos varía en función de la fase del procedimiento en que tenga lugar y de su grado de incidencia sobre el contenido de la decisión final. No es lo mismo, obviamente, automatizar el acuse de recibo que se envía al solicitante que la

12/05/2016, pp. 1-14).

¹⁷ Vid. Anexo, caso 2.

adopción de la decisión final sobre dicha solicitud. Tampoco es lo mismo automatizar la imposición de sanciones o la revocación de beneficios sociales que emplear algoritmos –como se hace a menudo– para identificar posibles infracciones y fraudes sobre los que iniciar las correspondientes investigaciones y procedimientos. En general, no deben exigirse a las actuaciones automatizadas requerimientos que no deben cumplir esas mismas actuaciones cuando son realizadas por humanos¹⁸.

Ello guarda relación con los distintos grados de automatización que pueden producirse en un determinado procedimiento y de incidencia en la decisión final. Cabe distinguir, así, entre sistemas algorítmicos que se limitan a proporcionar un *input* más al humano que debe decidir, los que proponen una decisión por defecto que el humano puede corregir, y los que adoptan la decisión sin que intervenga humano alguno¹⁹. En el último caso nos encontraremos ante un procedimiento completamente automatizado, mientras que los dos primeros constituirían modalidades de procedimientos parcialmente automatizados o semiautomatizados.

Como se ha señalado antes y se insistirá después, la intervención humana en el procedimiento es importante y puede servir para corregir los errores en que pueden incurrir los sistemas algorítmicos. Pero no debe ser sobrevalorada. Los juristas, marcados por el modelo judicial, en que una o varias personas resuelven definitivamente las controversias jurídicas, tendemos a incurrir en un cierto *sesgo de humanidad* y a confiar en que el hecho de que sea un humano el que finalmente decida solvente los problemas que pueda presentar un sistema algorítmico. El propio legislador incurre a menudo en dicho sesgo y contempla garantías solo cuando el procedimiento se

¹⁸ HUERGO LORA, A. (2021), “Administraciones Públicas e inteligencia artificial: ¿más o menos discrecionalidad?”, *El Cronista del Estado Social y Democrático de Derecho*, n° 96-97, pp. 78 ss. (pp. 7 y 12 de la versión publicada en línea por el INAP), poniendo el ejemplo de actuaciones que no deben ser objeto de motivación y que no son impugnables como, en España, la decisión sobre los sujetos que van a ser objeto de una inspección administrativa.

¹⁹ COGLIANESE, C. (2020): *A Framework for Governmental Use of Machine Learning*, Report for the Administrative Conference of the United States, ACUS, Washington, pp. 72-73.

encuentra totalmente automatizado²⁰. La realidad evidencia una y otra vez, sin embargo, que el tan conocido riesgo de sesgo de automatización (*automation bias*) es muy cierto, y que es muy habitual que los funcionarios que resuelven los procedimientos acepten, sin cuestionarlas ni examinarlas, las propuestas formuladas por sistemas automatizados. En muchos de los casos de mal funcionamiento documentados en Estados Unidos, la decisión final correspondía a humanos, que no advirtieron los errores del sistema²¹. Por ello, como se verá también después, es especialmente importante tomar medidas (procedimentales) adecuadas para garantizar el buen funcionamiento de los sistemas algorítmicos, tanto cuando se empleen en procedimientos totalmente automatizados, como solo semiautomatizados.

Los sistemas algorítmicos deben quedar sometidos a control jurídico (y, en particular, en lo que aquí interesa, a las garantías propias del procedimiento administrativo, como la transparencia y la motivación) siempre que contribuyan a determinar el contenido de la decisión administrativa²², y no solo cuando la dicten de forma automatizada. En la motivación de la decisión debe explicarse debidamente el uso del algoritmo y la influencia que ha ejercido en aquella, y tanto los interesados como, en su caso, el juez, han de poder comprobar que su funcionamiento ha sido correcto y ajustado a la ley.

3.2.- *Distinción en función del tipo de algoritmo empleado*

Desde el punto de vista de las exigencias procedimentales resulta también relevante tener en cuenta el tipo de algoritmo utilizado por el sistema automatizado. Junto a los algoritmos condicionales tradicionales propios de los sistemas expertos, en los últimos años se han desarrollado enormemente, como es sabido, los de aprendizaje automá-

²⁰ Es el caso de la regulación española (art. 41 LRJSP), ampliamente criticada por ello, y que provoca una *huida* de la Administración hacia sistemas semiautomatizados (COTINO HUESO, L. [2021], “El nuevo reglamento de Administración electrónica, que no innova en tiempos de transformación digital”, *Revista Catalana de Dret Públic*, n° 63, pp. 117 ss. [p. 131]).

²¹ CITRON, D. K. (2008), *op. cit.*, p. 1271.

²² HUERGO LORA, A. (2021), *op. cit.*, p. 12.

tico, y, dentro de estos, los de aprendizaje profundo (*deep learning*), basados en redes neuronales inspiradas por el cerebro humano. Los primeros operan sobre la base de reglas predeterminadas por los programadores (“*if-then*”), mientras que los de aprendizaje automático establecen sus propias reglas a partir de las correlaciones que inferen de grandes cantidades de datos (*big data*) con los que han sido entrenados²³. En el caso del aprendizaje profundo, el proceso tiene lugar a través de múltiples capas sucesivas y resulta especialmente complejo y opaco, sin que los propios programadores puedan explicar a menudo por qué el sistema ha sugerido un determinado resultado. Por ello se califica a tales sistemas como cajas negras, y por ello pueden resultar muy problemáticos desde el punto de vista del cumplimiento de las exigencias de motivación de las decisiones administrativas²⁴.

Dicha motivación es más sencilla cuando se emplean algoritmos condicionales tradicionales, puesto que, cuando están bien diseñados, codifican las normas legales y reglamentarias que rigen en cada momento la actuación de la Administración, y que son las que se aplicarán al caso concreto²⁵.

A menudo será preferible emplear un sistema automatizado más simple, basado en este tipo de algoritmos condicionales tradicionales, que un sistema opaco de aprendizaje automático. Es lo que sucede con el sistema ETIAS de autorización automatizada de viajes para los ciudadanos extracomunitarios²⁶, que combina adecuadamente la eficacia administrativa con las garantías procedimentales de los solicitantes y la necesaria predeterminación normativa (por parte del

²³ MULLIGAN, D. K.; BAMBERGER, K. A. (2019), *op. cit.*, pp. 814 ss.

²⁴ GAMERO CASADO, E. (2021), “Compliance (o Cumplimiento Normativo) de desarrollos de Inteligencia Artificial para la toma de decisiones administrativas”, *Diario La Ley*, n° 50, 19/04/2021 (p. 8). Relativiza el problema de la motivación de los algoritmos de *machine learning* COGLIANESE, C. (2020), *op. cit.*, p. 53.

²⁵ Por supuesto, puede suceder que el algoritmo esté mal diseñado e incluya reglas contrarias a las leyes vigentes. Además, la comprobación de su correcto diseño puede ser técnicamente compleja y venir dificultada por cláusulas de confidencialidad de las empresas que los desarrollen. Así lo ilustran los casos estudiados por CITRON, D. K. (2008), *op. cit.*, que no eran todavía, en su mayoría, de *machine learning*.

²⁶ *Vid.* Anexo, caso 4.

propio legislador) de los motivos que pueden fundar una resolución negativa²⁷.

Recientemente, en la importante sentencia *Ligue des droits humains* de 21 de junio de 2022, el Tribunal de Justicia se ha mostrado contrario al uso de sistemas de aprendizaje automático en la evaluación automatizada de los riesgos para la seguridad pública que pueden suponer los pasajeros aéreos, al considerar que son incompatibles con la exigencia de que dicha evaluación se base en criterios predeterminados (en la medida en que estos sistemas pueden modificar tales criterios o su ponderación sin intervención o revisión humana), con la necesaria revisión humana de los resultados positivos que pueda dar el sistema y con el derecho fundamental a la tutela judicial efectiva (debido a la opacidad que caracteriza a dichos sistemas, y que impide conocer la razón por la que han dado un resultado positivo)²⁸.

3.3.- *Distinción en función del carácter favorable o desfavorable de la decisión automatizada*

Conviene asimismo distinguir las decisiones o propuestas de decisión automatizadas en función de si son favorables o desfavorables a los interesados en el procedimiento. No es lo mismo, evidentemente, automatizar la concesión de una ayuda o autorización que su denegación o revocación. Tampoco es lo mismo automatizar la imposición de una sanción o prohibición que su revocación. Las garantías procedimentales devienen especialmente necesarias en el caso de las decisiones desfavorables para los particulares, como evidencia el art. 41.2.a) CDFUE al circunscribir el derecho de audiencia a las deci-

²⁷ Sin perjuicio de los riesgos de errores y discriminación apuntados por VAVOULA, N. (2021), “Artificial Intelligence (AI) at Schengen Borders: Automated Processing, Algorithmic Profiling and Facial Recognition in the Era of Techno-Solutionism”, *European Journal of Migration and Law*, <https://ssrn.com/abstract=3950389>; y DERAIVE, C.; GENICOT, N.; HETMANSKA, N. (2022), “The Risks of Trustworthy Artificial Intelligence: The Case of the European Travel Information and Authorisation System”, *European Journal of Risk Regulation*, n° 13(3), pp. 389 ss. Tales riesgos aconsejan efectuar un seguimiento detenido de su puesta en funcionamiento.

²⁸ STJUE de 21 de junio de 2022, *Ligue des droits humains*, C-817/19, ECLI:EU:C:2022:491, apartados 194 y 195.

siones administrativas “que [...] afecte[n] desfavorablemente” a las personas a las que van dirigidas.

El caso de ETIAS²⁹ suministra nuevamente un ejemplo interesante al automatizar únicamente la concesión de la autorización, pero no su denegación, que solo puede producirse de forma manual, a través de un humano. El agente que debe decidir sólo queda vinculado por el resultado del tratamiento automatizado en una serie de supuestos claros (en los que debe denegar la autorización), y en el más dudoso (cuando se produzca una coincidencia con los indicadores de riesgo específico del art. 33 del Reglamento del ETIAS) se señala expresamente que no podrá asumir de forma automática dicho resultado, y que deberá efectuar una evaluación individual de los riesgos para la seguridad, de inmigración ilegal y de epidemia que pueda suponer el solicitante³⁰. Esta evaluación individual deriva del deber procedimental de diligencia a que antes se ha hecho referencia.

Naturalmente, existe una gran cantidad de procedimientos administrativos multipolares, en los que hay distintas partes con intereses contrapuestos, y en los que la decisión administrativa puede ser favorable para uno o varios interesados y desfavorable para otros. Tampoco hay que olvidar que el procedimiento administrativo no solo trata de proteger los derechos de las personas afectadas por las decisiones administrativas, sino también la debida satisfacción de los intereses públicos. Ello significa que deben articularse los mecanismos para asegurar la corrección de la decisión automatizada, aunque sea favorable al interesado.

3.4.- Distinción en función del carácter discrecional o reglado de la decisión automatizada

Finalmente, importa distinguir si la decisión automatizada presenta o no elementos discrecionales. La adopción automatizada de decisiones discrecionales resulta mucho más controvertida que la de decisiones que están completamente predeterminadas por la

²⁹ Vid. Anexo, caso 4.

³⁰ Art. 26 del Reglamento (UE) 2018/1240 del Parlamento Europeo y del Consejo, de 12 de septiembre de 2018, por el que se establece un Sistema Europeo de Información y Autorización de Viajes (SEIAV).

legislación aplicable. Numerosos autores consideran que solo cabe o debería caber automatizar totalmente la toma de decisiones administrativas no discrecionales³¹, como establece expresamente el Derecho alemán tras la adición del § 35a a la Ley federal alemana de procedimiento administrativo³².

Esta es una cuestión sustantiva relacionada con la naturaleza jurídica de los algoritmos empleados por la Administración y la competencia de esta para diseñarlos³³. Tiene también, sin embargo, una dimensión procedimental, ya que las decisiones discrecionales de la Administración vienen siempre asociadas a exigencias reforzadas de motivación, a fin de descartar la existencia de arbitrariedad. En el caso de que se pueda automatizar una decisión discrecional, deberá contar con una motivación al menos tan detallada como la que se exigiría a un humano.

4.- ADAPTACIONES PROCEDIMENTALES NECESARIAS EN CASO DE UTILIZACIÓN DE ALGORITMOS

Además de tenerse que cumplir las exigencias tradicionales examinadas, que siguen plenamente vigentes y que la tecnología no puede convertir en papel mojado³⁴, parece necesario introducir algunas adaptaciones con tal de garantizar el pleno respeto de los fines perseguidos con la institución del procedimiento administrativo.

³¹ CITRON, D. K. (2008), *op. cit.*, p. 1304; HUERGO LORA, A. (2021), *op. cit.*, p. 15; PONCE SOLÉ, J. (2019), “Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico”, *Revista General de Derecho Administrativo*, n° 50, p. 14, fundándolo en la falta de empatía de las máquinas.

³² Vid. sobre este precepto, en inglés, BUOSO, E. (2020), “Fully Automated Administrative Acts in the German legal System”, *European Review of Digital Administration & Law*, n° 1, pp. 113 ss.

³³ Vid. apartado 4.3.

³⁴ De hecho, la tecnología facilita enormemente el cumplimiento de muchas de estas exigencias, abaratando y facilitando, p. ej., la audiencia a los interesados o su acceso al expediente (donde los procesos automatizados pueden quedar perfectamente registrados). Los algoritmos también pueden ayudar a motivar con mucha mayor profundidad las decisiones administrativas, al ofrecer datos e instrumentos analíticos antes inexistentes o muy difíciles de obtener.

4.1.- La necesidad de informar adecuadamente a los interesados y al público sobre los sistemas automatizados empleados

La primera adaptación consiste en informar específicamente a los interesados y al público en general sobre las decisiones adoptadas de forma automatizada.

Para evitar dudas, podría exigirse expresamente que dicha información se facilite a los interesados en la motivación de la decisión. Como mínimo, debería indicarse que esta ha sido producida de forma automatizada y ofrecerles la posibilidad, si lo desean, de solicitar el acceso a los detalles del tratamiento automatizado.

Esta comunicación individualizada a las personas afectadas por decisiones singulares podría combinarse con una información genérica al público sobre los principales tipos de decisiones automatizadas adoptadas por cada Administración, y que estaría disponible en su web respectiva.

En esta línea se sitúa la vigente legislación francesa, al disponer que la Administración debe informar a los interesados cuando adopte una decisión basada en tratamientos algorítmicos, y que aquellos podrán conocer, si lo solicitan, las normas reguladoras de dicho tratamiento y las principales características de su aplicación al caso concreto³⁵. El reglamento que ha concretado esta disposición legal señala que ello implica obtener información inteligible sobre el grado y modo de contribución del tratamiento algorítmico a la toma de la decisión; los datos tratados y sus fuentes; los parámetros del tratamiento y, en su caso, su ponderación, aplicados a la situación del interesado; y las operaciones realizadas por el tratamiento³⁶.

Esta misma ley francesa obliga también a las Administraciones a publicar en línea las normas reguladoras de los principales tratamientos

³⁵ Art. 4 de la Ley n° 2016-1321 de 7 de octubre de 2016 para una República Digital (1) (*Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique [1]*), que introdujo dicho precepto como art. L311-3-1 del Código de relaciones entre el público y la Administración (*Code des relations entre le public et l'administration*), la ley francesa de procedimiento administrativo.

³⁶ Art. R311-3-1-2 del referido Código, añadido mediante el Decreto n° 2017-330 de 14 de marzo de 2017.

algorítmicos empleados por ellas en la toma de decisiones individuales³⁷. Una obligación similar, algo más detallada, se contempla en el ordenamiento español³⁸.

Esta información específica sobre las decisiones automatizadas es una medida de transparencia algorítmica importante por diversas razones. Posibilita la adecuada motivación de las decisiones administrativas y el cumplimiento de las dos grandes funciones que corresponden a esta (la garantía de la defensa de los interesados y del control judicial). Alerta a los interesados sobre la posibilidad de que se produzca alguno de los errores absurdos que solo pueden cometer las máquinas, y les evita comprobar aquellos aspectos en que estas son infalibles. Permite el escrutinio público de la automatización acometida por la Administración. Facilita, cuando se constate el buen funcionamiento de la automatización, su consolidación y aceptación por parte de la ciudadanía. No es descabellado imaginar que, dentro de unos años, la indicación de que una determinada decisión ha sido adoptada de forma automatizada sea valorada positivamente por los interesados, al ser percibida como una garantía de acierto y objetividad.

Naturalmente, esta información detallada no debe comportar una relajación de las medidas de control del buen funcionamiento del sistema automatizado por parte de la Administración y un traslado de dicha carga a los interesados y la ciudadanía.

Debe también cohonestarse con los límites existentes al derecho de acceso al expediente y al derecho general de acceso a la información pública, en especial con los límites destinados a asegurar la eficacia de las investigaciones efectuadas por la Administración³⁹. Los intereses comerciales de los desarrolladores de los sistemas algorítmicos deben ceder ante el derecho de defensa de los interesados, y así debe indi-

³⁷ Art. 6 de la citada Ley n° 2016-1321, que introdujo dicho precepto como art. L312-1-3 del Código de relaciones entre el público y la Administración.

³⁸ Art. 11.1.i) RAFESP.

³⁹ Cfr. art. 41.2.b) CDFUE y, en relación con la Administración de la Unión, art. 4 del Reglamento (CE) n° 1049/2001 del Parlamento Europeo y del Consejo, de 30 de mayo de 2001, relativo al acceso del público a los documentos del Parlamento Europeo, del Consejo y de la Comisión.

carse en los pliegos contractuales, a fin de evitar que luego puedan invocarse cláusulas de confidencialidad⁴⁰.

4.2.- El establecimiento de un principio de supervisión humana

Una segunda adaptación importante sería la de establecer expresamente el principio de supervisión humana de las decisiones administrativas automatizadas. Este principio se deriva con toda claridad del art. 47 CDFUE en relación con el control judicial de la actuación administrativa, que necesariamente ha de ser ejercido por humanos según el marco normativo vigente en Europa. No hay ninguna duda de que las decisiones administrativas han de poder ser impugnadas judicialmente ante humanos, sin que sea necesario especificar que ello incluye también a las adoptadas de forma automatizada. En el caso de las decisiones de la Administración de la Unión, tales jueces y juezas son los integrantes del Tribunal General y del Tribunal de Justicia (arts. 253 y ss. TFUE).

No es tan claro, sin embargo, que dicho principio se extienda a la fase administrativa previa al recurso judicial. Puede sostenerse que se encuentra implícito en el citado deber de diligencia, cuando obliga a la Administración a investigar y examinar de forma minuciosa e imparcial todos los elementos relevantes del asunto. Al menos, cuando el interesado aporte, a lo largo del procedimiento, elementos fácticos que no hayan sido considerados por el sistema automatizado, tal como dispone el § 24.1.3 de la Ley federal alemana de procedimiento administrativo desde 2016. Pero una afirmación expresa de dicho principio por parte del legislador europeo o del Tribunal de Justicia dispararía toda duda.

Ello sería coherente con la finalidad perseguida por el art. 22 del Reglamento General de Protección de Datos⁴¹ y el precepto equivalente

⁴⁰ Así lo recomienda la Administrative Conference of the United States (ACUS, la agencia federal encargada de formular propuestas de mejora de la actuación de la Administración norteamericana) a las distintas agencias federales (ACUS [2021]: *Administrative Conference Statement #20, Agency Use of Artificial Intelligence*, 86 Fed. Reg. 6616, apartado 1).

⁴¹ Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta

aplicable a la Administración de la Unión⁴², cuando disponen que “[t]odo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado [...] que produzca efectos jurídicos en él o le afecte significativamente de modo similar” (apartado 1), y reconocen expresamente el derecho a obtener intervención humana (además de poder expresar su opinión e impugnar la decisión) en dos de los tres supuestos en que se permite la decisión automatizada (apartado 3). El tercer supuesto requiere en todo caso autorización por parte del Derecho de la Unión o de un Estado miembro y que se prevean “medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado” (apartado 2.b)]. La supervisión humana es una de tales posibles medidas, como se desprende del apartado 3.

El principio de supervisión humana puede inferirse también de la jurisprudencia del Tribunal de Justicia dictada al hilo del tratamiento automatizado de datos personales por razones de seguridad pública. En su sentencia *La Quadrature du Net* de 6 de octubre de 2020 ha señalado que los errores que pueden cometer los sistemas automatizados comportan que “cualquier resultado positivo obtenido gracias a un tratamiento automatizado debe reexaminarse individualmente por medios no automatizados antes de adoptar una medida individual que produzca efectos perjudiciales para las personas afectadas”⁴³. Esta di-

al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos).

⁴² Art. 24 del Reglamento (UE) 2018/1725 del Parlamento Europeo y del Consejo, de 23 de octubre de 2018, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por las instituciones, órganos y organismos de la Unión, y a la libre circulación de esos datos, y por el que se derogan el Reglamento (CE) n° 45/2001 y la Decisión n° 1247/2002/CE.

⁴³ STJUE de 6 de octubre de 2020, *La Quadrature du Net*, C-511/18, C-512/18 y C-520/18, ECLI:EU:C:2020:791, apartado 182. Antes ya en su Dictamen 1/15 (Acuerdo PNR UE-Canadá), de 26 de julio de 2017, ECLI:EU:C:2017:592, apartado 173. Y más recientemente, de nuevo y con mayor detalle, en la citada STJUE de 21 de junio de 2022, *Ligue des droits humains*, C-817/19, ECLI:EU:C:2022:491, apartados 179 y 202 ss. En esta última se señala que el número de falsos positivos en la detección automática de pasajeros aéreos sospechosos afecta nada más y nada menos que a cinco de cada seis personas identificadas (apartado 106).

rectriz del Tribunal de Justicia, que supone una especie de *Grundrechtsschutz durch Verfahren* del derecho fundamental a la protección de datos personales, parece perfectamente extensible a toda decisión administrativa desfavorable, aunque afecte a personas jurídicas y no a personas físicas. Ha sido recogida por el Reglamento del ETIAS, como se ha visto antes.

La supervisión humana no solo permite corregir errores de los sistemas automatizados. También acerca a los ciudadanos a las Administraciones y facilita que acepten (y cumplan) las decisiones adoptadas por estas, que es otra función importante del procedimiento administrativo. Una medida introducida en Italia hace ya muchos años y que se ha propuesto extender a la Administración de la Unión es, precisamente, la de designar a un responsable del procedimiento⁴⁴, un concreto empleado público que se responsabiliza de la tramitación y la interlocución con los interesados, a fin de evitar la despersonalización administrativa, incrementada enormemente con la digitalización y la automatización. Esta intervención humana es especialmente necesaria en el caso de la Administración de la Unión, tan distante para muchos europeos. Es también más sencilla, porque no tramita procedimientos tan masivos como las Administraciones de los Estados miembros.

No parece, en definitiva, que pueda desarrollarse una “IA centrada en el ser humano”, como pretende la Comisión⁴⁵, si se saca completamente a los humanos de la ecuación administrativa.

Esta supervisión humana puede tener diferentes configuraciones e intensidades dependiendo del tipo de procedimiento y de actuación administrativa de que se trate. Parece que, como mínimo, debería implicar

⁴⁴ MIR PUIGPELAT, O.; HOFMANN, H. C. H.; SCHNEIDER, J.-P.; ZILLER, J.; AUBY, J.-B.; CRAIG, P.; CURTIN, D.; DELLA CANANEA, G.; GALETTA, D.-U.; MENDES, J.; STELKENS, U.; WIERZBOWSKI, M. (Dir.) (2015): *Código ReNEUAL de procedimiento administrativo de la Unión Europea*, INAP, Madrid, art. III-7 y las explicaciones correspondientes (marginales 33 y 34).

⁴⁵ COMISIÓN EUROPEA (2019): *Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones—Generar confianza en la inteligencia artificial centrada en el ser humano*, COM(2019)168, 8 de abril de 2019, incluyendo expresamente la “supervisión humana” entre los requisitos esenciales de una IA fiable (p. 4).

el derecho a un recurso ante uno o varios humanos de la Administración en cuestión antes de acudir a la vía judicial⁴⁶.

Al producirse la intervención humana debe evitarse, por un lado, el sesgo de automatización a que se ha hecho antes referencia⁴⁷. Por otro, deben tenerse en cuenta las limitaciones humanas que el concreto algoritmo trata de superar, y que justifican su utilización⁴⁸: el humano no puede echar a perder, con su intervención, el valor añadido aportado por el sistema algorítmico.

4.3.- La necesidad de realizar evaluaciones de impacto antes y después de automatizar la toma de decisiones administrativas

Los humanos no solo han de participar, según lo expuesto, en el procedimiento automatizado. También intervienen de forma determinante en un momento previo, crucial: el diseño del sistema automatizado y del algoritmo en que se basa.

Ello suscita numerosos problemas de interés, como la naturaleza jurídica de los algoritmos, el tipo de reglas que pueden contener, el alcance de la competencia administrativa para configurarlos o la intervención que pueden tener los contratistas privados a los que muy frecuentemente se encarga su diseño.

El hecho de que los algoritmos empleados por la Administración predeterminen el contenido de las decisiones automatizadas que se adoptarán en los casos concretos futuros ha llevado a numerosos autores a sostener que tienen naturaleza normativa y que pueden ser equiparados a los reglamentos administrativos⁴⁹. Otros, en cambio, lo

⁴⁶ Como se desprende, en España, del art. 41.2 LRJSP (GAMERO CASADO, E. [2021], *op. cit.*, p. 3).

⁴⁷ *Vid.* apartado 3.1.

⁴⁸ Extensamente, al respecto, para justificar un amplio uso de los algoritmos de *machine learning* por parte de la Administración, COGLIANESE, C. (2020), *op. cit.*, pp. 8 ss.

⁴⁹ Así, p. ej., en la doctrina española, BOIX PALOP, A. (2020), “Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la administración para la adopción de decisiones”, *Revista de Derecho Público: Teoría y Método*, nº 1, 2020, pp. 223 ss. (*passim*); VAQUER CABALLERÍA, M. (2020): “¿Para qué sirve el procedi-

niegan, al considerar que son meros instrumentos técnicos de los que se vale la Administración⁵⁰.

A mi modo de ver, hay que distinguir el algoritmo informático⁵¹ de las reglas que en cada caso codifica –en los algoritmos condicionales tradicionales– o que infiere durante su aprendizaje –en los algoritmos de *machine learning*–, y el plano fáctico del plano jurídico. Tales reglas sí pueden predeterminar el contenido de las decisiones administrativas futuras, siendo las “normas” (legítimas o no) que el sistema automatizado aplicará a los casos concretos que se le planteen. Cumplen así, *de facto*, una función similar a las normas que rigen *de iure* la actuación administrativa. Pero no deben confundirse con ellas, ya que el ordenamiento jurídico atribuye dicha condición, y validez, únicamente a las reglas aprobadas por una serie de sujetos determinados (el legislador y la Administración en cada caso competente), tras la tramitación de un procedimiento preestablecido (el legislativo o reglamentario que en cada caso corresponda) y su publicación en un diario oficial. Dicho de otro modo: las normas que rigen o deben regir la actuación administrativa no son las que codifiquen o determinen sus algoritmos, sino las que cumplan estos requisitos organizativos y procedimentales decantados progresivamente a lo largo de la ya dilatada historia del Derecho público.

miento administrativo?”, en PAREJO ALFONSO, L.; VAQUER CABALLERÍA, M. (Dirs.): *Estudios sobre el procedimiento administrativo. III. Instituciones*, Tirant lo Blanch, Valencia, pp. 53 ss. (p. 67).

⁵⁰ Nuevamente, en la doctrina española, PONCE SOLÉ, J. (2019), *op. cit.*, p. 16; VALERO TORRIJOS, J. (2019), “Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración”, *Revista Catalana de Dret Públic*, nº 58, pp. 82 ss. (pp. 87-88); HUERGO LORA, A. (2020): “Una aproximación a los algoritmos desde el Derecho administrativo”, en HUERGO LORA, A. (Dir.): *La regulación de los algoritmos*, Aranzadi Thomson Reuters, Cizur Menor, pp. 23 ss. (pp. 64 ss.); ARROYO JIMÉNEZ, L. (2020), “Algoritmos y reglamentos”, *Almacén de Derecho*, <https://almacenederecho.org/algoritmos-y-reglamentos>; CERRILLO I MARTÍNEZ, A. (2021), “Robots, asistentes virtuales y automatización de las administraciones públicas”, *Revista Galega de Administración Pública*, nº 61, pp. 271 ss. (p. 292).

⁵¹ Es evidente que el algoritmo de un procesador de textos o una hoja de cálculo empleados por la Administración no constituye ninguna norma jurídica, como tampoco lo es un algoritmo de reconocimiento biométrico, de predicción de riesgos o de selección de personal utilizado por una empresa privada.

Por ello, las reglas algorítmicas no pueden contradecir las normas formalmente aprobadas como tales por el legislador y la propia Administración (normalmente, una Administración distinta de la que diseña y emplea el algoritmo, como sucederá a menudo a nivel de la Unión⁵²). En caso de no observarlas, y de introducir p. ej. requisitos ilegales para la obtención de ayudas, o contemplar sanciones para conductas no tipificadas como infracción en un acto legislativo, ello será motivo de anulación de la decisión automatizada subsiguiente.

Ocurre, sin embargo, que las normas (formales) aplicables a la Administración no predeterminan completamente su actuación, y le confían amplios ámbitos de discrecionalidad⁵³. Es entonces cuando se plantea si cabe que cada Administración ejerza dicha discrecionalidad de forma “anticipada” a través de la definición de los algoritmos que empleará en la toma de decisiones automatizadas (especificando los criterios discrecionales que el sistema se limitará a aplicar de forma automática), o si solo puede hacerlo mediante decisiones humanas individuales adoptadas en los casos concretos que deba resolver, a la vista de sus particularidades⁵⁴. Se trata de la cuestión sustantiva relativa a la admisibilidad de las decisiones algorítmicas discrecionales a que antes se ha hecho referencia⁵⁵, y sobre la que no es posible profundizar en estas páginas.

Lo que aquí interesa subrayar es que, pese a estas divergencias teóricas sobre la naturaleza de los algoritmos empleados por la Administración, existe un amplio consenso en la literatura académica acerca de la conveniencia de extender algunas de las garantías propias

⁵² Como es sabido, solo la Comisión, y no las agencias europeas, puede adoptar los principales actos no legislativos de alcance general (actos delegados y de ejecución).

⁵³ Aunque el Derecho europeo, en la línea de los ordenamientos continentales, conceda una importancia mayor que el norteamericano a la reserva de ley y la predeterminación legislativa de la actuación administrativa (cfr. art. 52.1 CDFUE), reconoce también amplios espacios de discrecionalidad a la Comisión y las agencias europeas.

⁵⁴ Cabe también, por supuesto, que la Administración apruebe un reglamento formal, cuando sea competente para ello. Esta opción no plantea problemas: el algoritmo deberá recoger las normas contenidas en este reglamento.

⁵⁵ *Vid.* apartado 3.4.

del procedimiento administrativo a la fase de su diseño⁵⁶. Aunque no se les considere verdaderos reglamentos, hay poderosos argumentos para someterlos, en particular, a trámites procedimentales habituales en la elaboración de estos y de actos singulares de especial trascendencia, como la autorización de actividades industriales de elevado impacto ambiental. Se trata, sobre todo, de someter la elaboración de algoritmos a *evaluaciones de impacto* y a *consultas públicas*, dos instrumentos típicos de la preparación de normas de la Comisión, como ya nos consta. La propuesta más detallada es la que ha presentado recientemente el European Law Institute⁵⁷, a partir de experiencias previas como la del Gobierno canadiense⁵⁸.

Esta procedimentalización de la elaboración de los algoritmos administrativos potencialmente más peligrosos parece plenamente justificada. Como se ha visto, los algoritmos no solo determinan las decisiones automatizadas de la Administración, sino que condicionan de forma muy significativa muchas de las que adopta su personal, en el amplio abanico de las actuaciones semiautomatizadas⁵⁹. Los algoritmos son también cada vez más importantes en actuaciones administrativas que no se traducen en decisiones singulares amparadas por procedimientos administrativos, como la prestación de servicios públicos o la formulación de advertencias⁶⁰.

⁵⁶ Vid. los trabajos citados en las notas 49 y 50, así como CITRON, D. K. (2008), *op. cit.*, pp. 1308 ss.; REISMAN, D.; SCHULTZ, J.; CRAWFORD, K.; WHITTAKER, M. (2018): *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability*, AI Now Report; MULLIGAN, D. K.; BAMBERGER, K. A. (2019), pp. 835 ss.; GAMERO CASADO, E. (2021), pp. 9-10.

⁵⁷ EUROPEAN LAW INSTITUTE (2022): *Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration* (Normas Modelo del ELI).

⁵⁸ Directive on Automated Decision-Making (Government of Canada). Aunque de forma más laxa, también la ACUS ha recomendado a las distintas agencias federales que evalúen adecuadamente los distintos riesgos de los sistemas de inteligencia artificial antes de su puesta en funcionamiento (ACUS [2021], *op. cit.*, *passim*).

⁵⁹ Según el art. 2.1 de las Normas Modelo del ELI, estas se aplican, acertadamente, no solo a los sistemas totalmente automatizados, sino también a los que apoyan la toma de decisiones humanas (“*support[...] human decision-making*”).

⁶⁰ El art. 2.3 de las Normas Modelo del ELI maneja también acertadamente un concepto amplio de decisión, definida como cualquier decisión de una autoridad

Someter a los algoritmos a esta detenida evaluación previa reduce el riesgo de errores y de mal funcionamiento a que antes se ha aludido. De entrada, exige que la Administración respectiva cuente con personal suficientemente cualificado para llevar a cabo la evaluación y valorar debidamente el impacto que tendrá el uso del algoritmo⁶¹. Obliga a considerar los riesgos que pueda comportar y a ponderarlos con las ventajas que presente. La sumisión a consulta pública y/o a auditorías de expertos externos⁶² refuerza el control del diseño del algoritmo y permite someter a discusión pública los usos que puedan resultar más controvertidos.

En términos económicos, esta evaluación detenida se encuentra justificada por el uso masivo que tendrá luego el algoritmo y los costes sociales elevados que un diseño defectuoso puede tener. Es necesaria porque no todos los destinatarios de las decisiones que se adopten después sobre la base del algoritmo –en particular, los colectivos más vulnerables– estarán en condiciones de defenderse adecuadamente en el marco del procedimiento administrativo correspondiente.

En todo caso, el hecho de que se evalúe detenidamente el algoritmo no exime de tramitar el procedimiento requerido para la toma posterior de decisiones individuales, examinado en los apartados precedentes de este trabajo. Ello es así porque debe poder verificarse que la utilización del algoritmo en el caso concreto ha sido correcta y adecuada. Si la tramitación de procedimiento para la aprobación de un reglamento no exime de la tramitación de procedimiento para su aplicación a casos individuales, menos aún debe producirse dicha exención cuando lo que se aplique sea un algoritmo.

pública de emprender o no emprender una actuación (“*any determination by a public authority to take or not to take action*”). *Vid.* la justificación que se ofrece en las pp. 41-42, con algunos ejemplos.

⁶¹ Sobre las dificultades de la Administración para contar con personal suficientemente especializado en inteligencia artificial, COGLIANESE, C. (2020), *op. cit.*, p. 40. Sobre la necesidad de que los sistemas algorítmicos sean desarrollados por equipos multidisciplinares integrados también por juristas, GAMERO CASADO, E. (2021), *op. cit.*, pp. 4, 9.

⁶² Que también contemplan los arts. 10 y 11 de las Normas Modelo del ELI para los sistemas que merezcan la calificación de alto riesgo. Sobre las auditorías algorítmicas *vid.* COSTANZA-CHOCK, S.; RAJI, I. D.; BUOLAMWINI, J. (2022), “Who Audits the Auditors? Recommendations from a field scan of the algorithmic auditing ecosystem”, *FACCT*’22, proponiendo que sean obligatorias.

Al igual que sucede con la evaluación de impacto normativo, la evaluación de los algoritmos no solo debe producirse *ex ante*, sino también *ex post*, esto es, una vez el algoritmo ha sido puesto en funcionamiento. Así lo defiende la doctrina y así se establece en las distintas iniciativas apuntadas⁶³. En esta evaluación periódica debe comprobarse que el sistema funcione correctamente, tal como se había previsto. En caso de que se adviertan disfunciones, ello permitirá no solo corregir el algoritmo⁶⁴, sino revisar de oficio las decisiones erróneas que hayan podido dictarse a su amparo. Naturalmente, también habrá que modificar el algoritmo siempre que se produzcan circunstancias que obliguen a ello, como p. ej. un cambio en las normas que regulen la actuación administrativa de que se trate.

A la vista de lo señalado, cabría estudiar la posibilidad de que esta evaluación previa y posterior del impacto de los nuevos algoritmos usados por la Administración fuera, al menos en ciertos casos, exigida legalmente⁶⁵. Ello contribuiría a incrementar la confianza de la población europea en su uso, que puede reportar grandes beneficios sociales, y a evitar que los algoritmos acaben siendo percibidos como la versión transgénica de los procesos decisorios tradicionales.

5.- ANEXO: EL USO DE IA POR PARTE DE LA ADMINISTRACIÓN DE LA UNIÓN: UN ESTUDIO DE CAMPO

5.1.- *Finalidad del estudio y casos examinados*

A los integrantes del proyecto de investigación en que se enmarca este trabajo nos ha parecido importante dedicar los primeros meses de su desarrollo a estudiar con cierto detalle el uso de algoritmos

⁶³ Normas Modelo del ELI (art. 14); Directive on Automated Decision-Making (Government of Canada), sección 6.3.2; ACUS (2021), *op. cit.*, apartado 9.

⁶⁴ Como señala COGLIANESE, C. (2020), *op. cit.*, pp. 24 y 49, es más fácil corregir algoritmos defectuosos o sesgados que reentrenar a los empleados públicos que incurran en esos mismos sesgos o errores.

⁶⁵ Como se propone para la Administración de la Unión en la p. 12 de las Normas Modelo del ELI.

avanzados por parte de la Administración de la Unión (principalmente, la Comisión Europea y las distintas agencias de la Unión). En particular, el uso de tales algoritmos para automatizar total o parcialmente la toma de decisiones administrativas con efectos jurídicos sobre los ciudadanos y las empresas, ya se trate de normas de alcance general (los actos no legislativos de alcance general que aprueba la Comisión, a menudo a propuesta de las agencias) o de decisiones singulares sobre casos concretos. Aunque la ejecución administrativa del Derecho de la Unión corresponda sobre todo a los Estados miembros, los Tratados y el Derecho derivado otorgan también competencias destacadas a la Comisión y las agencias europeas para adoptar este tipo de decisiones. En el Código ReNEUAL⁶⁶ propusimos una serie de reglas procedimentales mínimas aplicables a la adopción de tales decisiones por parte de la Administración de la Unión, extraídas de las normas sectoriales ya existentes y de las leyes de procedimiento nacionales que deben observar las Administraciones de los Estados miembros.

Esta tarea de estudio del uso de herramientas de IA por parte de la Administración de la Unión ha tenido, metodológicamente, dos fases diferenciadas. En primer lugar, se han revisado detenidamente la literatura existente y las páginas web de la Comisión y las distintas agencias, en búsqueda de información adecuada sobre el uso de tales herramientas. Se ha prestado una atención especial al suplemento del DOUE dedicado a la contratación pública (TED), teniendo en cuenta que la experiencia, a nivel nacional, demuestra que muchas de las herramientas de IA que usan las Administraciones públicas no son desarrolladas por ellas mismas, sino que las adquieren en el mercado de proveedores externos. Esta primera fase ha aportado muy poca luz. La información disponible en línea o en publicaciones previas es muy escasa y fragmentaria. Existen informes detallados e interesantes sobre experiencias nacionales coordinados por las instituciones de la UE⁶⁷, pero ninguno sobre el uso que estas mismas hagan de sistemas

⁶⁶ Vid. *supra*, nota 44.

⁶⁷ Destacan especialmente los elaborados por el Joint Research Centre (JRC) de la Comisión, y que integran la serie AI Watch. En particular, los informes de MISURACA, G.; VAN NOORDT, C. (2020): *Overview of the use and impact of AI in public services in the EU*, EUR 30255 EN, Publications Office of the

de IA. La situación contrasta ostensiblemente con la existente en los Estados Unidos, donde abunda la información relativa a las agencias federales. De hecho, el informe más detallado sobre la utilización de herramientas de IA a nivel federal ha sido encargado y facilitado por una de sus agencias, la ACUS⁶⁸. No hay nada parecido a nivel de la Administración de la UE.

Ante la poca información disponible, se ha pasado a una segunda fase de entrevistas semiestructuradas con distintos responsables de la Comisión y de algunas agencias europeas. Tras contactar con DG-Connect, la Dirección General de la Comisión Europea responsable de la elaboración de la importante Propuesta de Reglamento de IA (RIA) llamado a establecer el marco regulatorio general del uso de sistemas de IA tanto en el sector público como privado, y después de múltiples peticiones a sucesivos interlocutores potenciales de la Comisión y de distintas agencias, entre los meses de julio y diciembre de 2021 consiguieron celebrarse entrevistas con representantes de la DG de Agricultura (DG-Agri) y de las agencias EFSA (Autoridad Europea de Seguridad Alimentaria), EUIPO (Oficina de Propiedad Intelectual de la UE) y eu-LISA (Agencia Europea para la Gestión Operativa de Sistemas Informáticos de Gran Magnitud en el Espacio de Libertad, Seguridad y Justicia)⁶⁹. Tales entrevistas fueron muy ilustrativas y conviene resumirlas brevemente.

European Union, Luxembourg, doi:10.2760/039619, JRC120399; y TANGI L.; VAN NOORDT C.; COMBETTO M.; GATTWINKEL D.; PIGNATELLI F. (2022): *AI Watch. European Landscape on the Use of Artificial Intelligence by the Public Sector*, EUR 31088 EN, Publications Office of the European Union, Luxembourg, doi:10.2760/39336, JRC129301.

⁶⁸ ENGSTROM, D. F.; HO, D. E.; SHARKEY, C. M.; CUÉLLAR, M.-F. (2020): *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*, Report submitted to the Administrative Conference of the United State, ACUS, Washington.

⁶⁹ Se trata de Doris Marquardt (DG-Agri, 30.7.2021), Ermanno Cavalli (EFSA, 14.10.2021), Rahul Bhartiya (EUIPO, 30.11.2021) y Aleksandrs Cepilovs (eu-LISA, 22.7.2021 y 14.12.2021), a los que agradecemos muy sinceramente la excelente colaboración brindada.

5.1.1.- Caso 1 (DG-Agri/ESA): la utilización de IA para monitorizar por satélite los cultivos europeos y el cumplimiento de la normativa sobre subvenciones agrícolas de la PAC

El primer caso de uso se refiere a una experiencia piloto en el ámbito de la Política Agrícola Común (PAC) destinada a monitorizar por satélite los cultivos europeos y el cumplimiento de la normativa sobre subvenciones agrícolas. En la actualidad, los Estados miembros tienen la obligación de inspeccionar sobre el terreno el cinco por ciento de los cultivos subvencionados a fin de comprobar el cumplimiento de los requisitos exigidos y evitar el fraude. El nuevo sistema, que se vale de algoritmos de aprendizaje automático para mejorar la precisión del reconocimiento de las imágenes tomadas por satélite⁷⁰, busca, entre otras finalidades, monitorizar la totalidad de los campos europeos, también de los que son más difícilmente accesibles, y reducir y optimizar el número de inspecciones sobre el terreno, en beneficio de las Administraciones nacionales y de los propios agricultores, a los que puede facilitar también la obtención de las subvenciones.

El sistema, impulsado por la Agencia Espacial Europea (ESA) y orientado por un comité de dirección integrado por las tres Direcciones Generales de la Comisión involucradas en la PAC (DG-Agri, DG-Grow y DG-JRC), está siendo desarrollado tecnológicamente por un consorcio público-privado liderado por una universidad belga, en el marco de un proyecto de investigación del programa europeo Horizon 2020. Se está implementando de forma piloto en seis Estados miembros: España, Italia, Lituania, Países Bajos, República Checa y Rumanía.

El sistema no adopta decisiones automatizadas, sino que se limita a emitir alertas en caso de posible incumplimiento. Tales alertas son verificadas por humanos a través de la revisión o ampliación de las imágenes o, en su caso, de una inspección sobre el terreno, antes de que se produzca la decisión jurídica de denegar la subvención solicitada o de obligar a reintegrar la previamente concedida. La monitorización por satélite puede formar parte, pues, de los complejos procedimientos de otorgamiento, control y revocación de subvenciones de

⁷⁰ Principalmente, los satélites Sentinel-1 y Sentinel-2 de la ESA.

la PAC. Constituye un medio adicional de prueba del cumplimiento o incumplimiento de la normativa y, como tal, formaría parte de la fase de obtención de información del procedimiento de adopción de decisiones contemplado en el Libro III del Código ReNEUAL (Capítulo 3, arts. III-10 y ss.). Dicha monitorización por satélite ha sido admitida y regulada por un Reglamento de ejecución de la Comisión de 2018⁷¹, que no se ocupa de la tecnología empleada ni, en particular, del uso de herramientas de IA para el análisis de las imágenes tomadas⁷².

5.1.2.- Caso 2 (EFSA): la utilización de IA para el análisis de publicaciones científicas relevantes en la evaluación de riesgos alimentarios

El segundo caso de uso se refiere a la automatización, mediante algoritmos de *machine learning*, de parte del proceso de análisis de publicaciones científicas relevantes que lleva a cabo la EFSA cuando efectúa evaluaciones del riesgo de determinadas sustancias o productos. Esta revisión exhaustiva de la literatura científica, conocida como revisión sistemática, constituye una parte fundamental de la actuación de evaluación de riesgos que caracteriza a la EFSA y a agencias europeas similares como la Agencia Europea de Medicamentos (EMA), el Centro Europeo para la Prevención y el Control de las Enfermedades (ECDC) y la Agencia Europea de Sustancias y Mezclas Químicas (ECHA). Consume una gran parte de sus recursos, al obligar a los expertos que las llevan a cabo a examinar un volumen ingente de publicaciones, que crece exponencialmente. Se trata de un proceso lento, tedioso y que a menudo se encuentra ya obsoleto cuando finaliza.

⁷¹ Reglamento de Ejecución (UE) 2018/746 de la Comisión, de 18 de mayo de 2018, por el que se modifica el Reglamento de Ejecución (UE) n° 809/2014 en lo que respecta a la modificación de las solicitudes únicas y solicitudes de pago y a los controles. Véase especialmente el nuevo art. 40 bis sobre los controles por medio de monitorización.

⁷² Para más información sobre este primer caso de uso puede verse la web del proyecto que lo desarrolla (<http://esa-sen4cap.org>), así como el Informe Especial 04/2020 del Tribunal de Cuentas Europeo, que lo valora positivamente y recomienda su potenciación.

La EFSA lleva ya varios años trabajando en la automatización parcial de dicho proceso⁷³. Ya utiliza con normalidad la automatización de la fase inicial de selección de las publicaciones relevantes, y que opera sobre la base de un análisis de su título y de su resumen. Esta selección excluye los trabajos considerados irrelevantes y reduce normalmente de varios miles a pocos centenares aquellos que serán objeto de estudio. Ello se realiza mediante el programa DistillerSR comercializado por la empresa Evidence Partners, y permite sustituir a uno de los dos expertos requeridos habitualmente en la revisión. En su informe final, los expertos indican que han empleado la herramienta.

La EFSA querría llegar a automatizar fases posteriores del proceso de revisión, como la de extracción de datos relevantes de los trabajos previamente seleccionados, e incluso la de valoración crítica de tales trabajos, que permite determinar su calidad. En lo concerniente a la extracción de datos, están colaborando con la agencia federal estadounidense de protección ambiental (EPA) en el suministro de datos sobre seguridad alimentaria para entrenar adecuadamente a un programa de *machine learning* (Fiddle) desarrollado por la empresa Sciome con una subvención concedida por la EPA⁷⁴.

El dictamen científico final sobre la evaluación del riesgo es siempre elaborado por un experto humano, aunque un error en el cribaje automatizado de las publicaciones relevantes puede, por supuesto, dejar fuera trabajos y evidencias científicas importantes que no han podido ser considerados en la elaboración de aquel.

⁷³ Vid. JASPERS, S.; DE TROYER, E.; AERTS, M. (2018): *Machine learning techniques for the automation of literature reviews and systematic reviews in EFSA*, EFSA supporting publication 2018:EN-1427, doi:10.2903/sp.efsa.2018.EN-1427.

⁷⁴ Vid. la convocatoria de la EFSA GP/EFSA/AMU/2020/03 – Support for Automating some specific steps of Systematic Review process using Artificial Intelligence (que ya no se encuentra disponible en la web de la EFSA), convocando una subvención para el desarrollo de estos conjuntos de datos de entrenamiento del referido programa informático.

5.1.3.- Caso 3 (EUIPO): la utilización de IA en el procedimiento de registro de marcas y diseños

La agencia europea de propiedad intelectual (EUIPO) registra anualmente alrededor de 135.000 marcas y 100.000 diseños, tramitando solicitudes presentadas en 23 lenguas distintas, por lo que no es de extrañar la importante apuesta que ha efectuado por la introducción de herramientas de IA destinadas a facilitar la labor de sus empleados y de los solicitantes⁷⁵. Entre julio de 2020 y junio de 2025 está desarrollando un proyecto de implementación de soluciones de IA en distintos ámbitos de su actuación que cuenta con un presupuesto de 2.860.000 euros y 24,5 empleados a tiempo completo⁷⁶.

Entre las distintas herramientas que se están introduciendo pueden destacarse dos relacionadas con el procedimiento de registro. La primera consiste en la posibilidad de efectuar búsquedas de imágenes similares a través de la base de datos eSearchPlus, disponible en la web de EUIPO para cualquier persona que se plantee registrar una determinada marca o diseño y desee comprobar si la que tiene en mente se encuentra ya registrada.

La segunda herramienta posibilita una comparación de bienes y servicios basada en IA que permite a los funcionarios de la EUIPO evaluar los casos de oposición (aquellos en los que un tercero se opone a la marca que se pretende registrar) con mayor facilidad y calidad. Los solicitantes de una determinada marca deben indicar los bienes y servicios que esta pretende cubrir, y hay marcas que pueden abarcar hasta dos mil bienes y servicios distintos. En caso de oposición, los funcionarios deben acometer la comparación de los bienes y servicios cubiertos por las marcas respectivas, lo que resulta lento y tedioso, además de complejo en los muchos supuestos en que no hay una distinción clara entre dos bienes o servicios. La herramienta de IA implementada facilita esta comparación al sugerir al funcionario

⁷⁵ Sobre los casos de uso desarrollados por su homóloga estadounidense, la US Patent and Trademark Office, *vid.* ENGSTROM, D. F.; *et al.* (2020), *op. cit.*, pp. 46 ss.

⁷⁶ https://euiipo.europa.eu/tunnel-web/secure/webdav/guest/document_library/contentPdfs/Strategic_Plan_2025/project_cards/SD3_Artificial_Intelligence_implementation_PC_en.pdf (última visita: 31/08/2022).

una respuesta al par de bienes y servicios en conflicto sobre la base de los miles de resoluciones dictadas con anterioridad por la EUIPO. El sistema incluso proporciona la motivación aportada en las decisiones precedentes, a fin de facilitar la redacción de la decisión, que compete en todo caso al funcionario o funcionarios que integran cada División de Oposición⁷⁷. Tales decisiones son impugnables por los interesados ante las Salas de Recurso de EUIPO, integrados también por una o tres personas físicas⁷⁸.

Destaca el hecho de que, a diferencia de lo que suele ser habitual, EUIPO está desarrollando estas herramientas de IA con sus propios medios, sin adquirirlas de terceros.

5.1.4.- Caso 4 (eu-LISA): la utilización de IA para el reconocimiento biométrico de personas en las fronteras de la Unión

El cuarto y último caso de uso corresponde a eu-LISA, la agencia europea responsable de la gestión de sistemas de información básicos para las autoridades fronterizas y policiales de los Estados miembros, como son los sistemas de información de Schengen (SIS II), de visados (VIS) y de asilo (Eurodac). Está también desarrollando, para su próxima puesta en funcionamiento, nuevos sistemas de información ya regulados por el Derecho de la Unión, como el Sistema de Entrada/Salida (SES; EES en inglés), el Sistema Europeo de Información y Autorización de Viajes (SEIAV, más conocido por sus siglas en inglés ETIAS) y el Sistema Europeo de Información de Antecedentes Penales–Nacionales de Terceros Países (ECRIS-TCN).

La IA se utiliza en los tres primeros sistemas y en los próximos EES y ECRIS-TCN para la identificación y verificación biométrica de las personas en las fronteras de la Unión y dentro de los Estados miembros⁷⁹. Todos ellos emplean sistemas de correspondencia biométrica,

⁷⁷ Vid. la noticia publicada en la web de EUIPO “New AI-based comparison of goods and services”, de 29 de marzo de 2022 (última visita: 31/08/2022).

⁷⁸ Arts. 66 ss. y 159 ss. del Reglamento (UE) 2017/1001 del Parlamento Europeo y del Consejo, de 14 de junio de 2017, sobre la marca de la Unión Europea.

⁷⁹ Sobre los casos de uso de reconocimiento facial por parte de la agencia federal estadounidense de control de fronteras, Customs and Border Protection (CBS),

que utilizan algoritmos avanzados de aprendizaje automático para cotejar las imágenes faciales y huellas dactilares tomadas en las fronteras con las almacenadas en dichos sistemas de información. Cada sistema cuenta con su propio servicio de correspondencia biométrica⁸⁰, pero las empresas que están desarrollando el sistema biométrico del EES también están ocupándose de implementar una herramienta que posibilite la búsqueda y comparación simultánea de datos biométricos en todos estos sistemas de información a la vez⁸¹. Se trata del servicio de correspondencia biométrica compartido (SCB compartido, sBMS en inglés), previsto y regulado en los arts. 12 y ss. de los Reglamentos que permiten la interoperabilidad entre todos estos sistemas de información⁸².

Los referidos sistemas de correspondencia biométrica no son desarrollados por eu-LISA, sino por contratistas privados, sobre la base de las especificaciones técnicas fijadas por esta, que verifica también su buen funcionamiento. El contrato de desarrollo del sistema del EES y del sBMS se ha adjudicado por un importe de 302 millones de euros a un consorcio de empresas europeas⁸³. Como es sabido, un aspecto esencial de todo sistema de *machine learning* es su entrenamiento, que debe realizarse con una gran cantidad de datos de calidad para que el rendimiento del sistema sea adecuado. La formación de estos datos de entrenamiento es muy costosa y está cubierta por el secreto empresarial de los contratistas, que no permiten que eu-LISA acceda

vid. ENGSTROM, D. F.; *et al.* (2020), *op. cit.*, pp. 30 ss.

⁸⁰ Eu-LISA (2018): *Shared Biometric Matching Service (sBMS) Feasibility Study – final report*, doi:10.2857/84504, p. 5.

⁸¹ Eu-LISA, Call for Tender – Framework contract for implementation and maintenance in working order of the biometrics part of the Entry Exit System and future Shared Biometrics Matching System, LISA/2019/RP/05 EES BMS and sBMS, Executive Summary, pp. 7-8.

⁸² Reglamento (UE) 2019/817 del Parlamento Europeo y del Consejo, de 20 de mayo de 2019, relativo al establecimiento de un marco para la interoperabilidad de los sistemas de información de la UE en el ámbito de las fronteras y los visados, y Reglamento (UE) 2019/818 del Parlamento Europeo y del Consejo, de 20 de mayo de 2019, relativo al establecimiento de un marco para la interoperabilidad entre los sistemas de información de la UE en el ámbito de la cooperación policial y judicial, el asilo y la migración.

⁸³ <https://ted.europa.eu/udl?uri=TED:NOTICE:200083-2020:TEXT:EN:HTML> (última visita: 31/08/2022).

a ellos. Eu-LISA, por tanto, ignora cuáles son los datos empleados por sus contratistas para entrenar los sistemas, y si adolecen de los sesgos (principalmente raciales y de género) que se han observado con frecuencia en el entrenamiento de sistemas de reconocimiento biométrico⁸⁴. Para paliarlo, eu-LISA piensa llevar a cabo una evaluación y pruebas independientes del rendimiento del sBMS, en las que, entre otros parámetros, comprobará los posibles sesgos de género y raza.

En todo caso, los sistemas que eu-LISA pone a disposición de los Estados miembros serían de los más avanzados del mundo y tendrían un rendimiento muy elevado, superior al del policía de fronteras más experimentado. Su precisión se habría multiplicado por diez desde que tales sistemas empezaron a ser usados por eu-LISA en 2014, y vendría facilitada por los entornos controlados en que operan (aeropuertos que cuentan con buenas cámaras, en que las imágenes se toman sin movimiento, con una iluminación adecuada, etc., a diferencia de lo que sucede con las cámaras de videovigilancia).

La regulación europea existente de estos sistemas de correspondencia biométrica empleados en las fronteras de la Unión no se ocupa de las particularidades suscitadas por el hecho de que se basen en algoritmos de aprendizaje automático, o de que sean desarrollados por contratistas externos. Sí establece, entre otros aspectos, los requisitos de calidad que deben cumplir las huellas dactilares e imágenes faciales empleadas, la tasa de falsos positivos y negativos admitidos y el control periódico (al menos mensual) que eu-LISA debe efectuar del rendimiento del sistema⁸⁵.

⁸⁴ Vid. p. ej. el conocido trabajo de BUOLAMWINI, J.; GEBRU, T. (2018), “Gender shades: Intersectional accuracy disparities in commercial gender classification”, Conference on Fairness, Accountability and Transparency, *PMLR*, n° 81, pp. 77 ss, que llevó a la primera de las autoras a declarar ante el Congreso de los Estados Unidos sobre el impacto de la tecnología de reconocimiento facial en los derechos individuales.

⁸⁵ Vid., para el EES, el Anexo de la Decisión de Ejecución (UE) 2019/329 de la Comisión, de 25 de febrero de 2019, por la que se establecen las especificaciones para la calidad, resolución y uso de impresiones dactilares e imágenes faciales, para la verificación e identificación biométrica, en el Sistema de Entradas y Salidas (SES).

Importa destacar que el otro gran sistema de información que está implementando actualmente eu-LISA, el ETIAS⁸⁶, no cuenta con algoritmos de aprendizaje automático. El Reglamento que lo regula predetermina con detalle los aspectos que deberá comprobar el sistema cuando un nacional de un tercer país solicite la autorización para viajar al territorio de la Unión⁸⁷. El sistema informático concederá automáticamente la autorización para viajar cuando dichas comprobaciones regladas arrojen un resultado negativo. Cuando el resultado sea positivo y se produzca un “*hit*” (porque, p. ej., el solicitante emplee un pasaporte que consta en la base de datos de Interpol de pasaportes perdidos o robados, pertenezca a la lista de alerta rápida de ETIAS por ser sospechoso de terrorismo, o encaje en alguno de los indicadores específicos de riesgo que se elaboren con arreglo al art. 33 del Reglamento), el sistema informará a Frontex para que efectúe las verificaciones pertinentes y, de confirmarse el resultado positivo, transmita la solicitud al Estado miembro competente para que decida la solicitud manualmente (esto es, a través de un humano) y de forma motivada. Se trata, pues, de un sistema algorítmico tradicional, perfectamente rastreable, que se limita a comprobar que concurren las condiciones previamente establecidas por el legislador-programador (sistema “*if-then*”), sin que establezca reglas nuevas a partir de correlaciones que pueda extraer de grandes cantidades de datos, como sucede con los algoritmos de aprendizaje automático⁸⁸.

Resulta cuestionable que los sistemas de IA que se integren en estos sistemas de información gestionados por eu-LISA antes de que transcurran 36 meses desde la entrada en vigor de la Propuesta de

⁸⁶ El equivalente europeo del Electronic System for Travel Authorization (ESTA) norteamericano. ETIAS exigirá obtener una autorización para viajar al territorio de la Unión, por un período máximo de 90 días, a los ciudadanos no comunitarios procedentes de países exentos de visado. Está previsto que entre en funcionamiento en mayo de 2023.

⁸⁷ *Vid.* los arts. 20 ss. del Reglamento (UE) 2018/1240 del Parlamento Europeo y del Consejo, de 12 de septiembre de 2018, por el que se establece un Sistema Europeo de Información y Autorización de Viajes (SEIAV).

⁸⁸ Un análisis pormenorizado y crítico de los algoritmos de reconocimiento facial y de evaluación de riesgos empleados por estos sistemas de información operados por eu-LISA puede encontrarse en VAVOULA, N. (2021), *op. cit.* y en DERAVE, C.; GENICOT, N.; HETMANSKA, N. (2022), *op. cit.*

RIA queden excluidos de esta⁸⁹, pese a tener la consideración de alto riesgo con arreglo a su Anexo III⁹⁰.

5.2.- Algunas conclusiones que cabe extraer del estudio de campo realizado

Cabe extraer algunas conclusiones del estudio de campo realizado.

La primera es la poca información disponible sobre los casos de uso de IA existentes en el seno de la Administración de la Unión. Llama la atención que dicha información no solo no esté disponible en internet, sino que ni siquiera se encuentre al alcance de ningún servicio centralizado de la propia Unión. Existe una red informal entre determinadas agencias europeas (“*AI Virtual Community*”) que intercambia experiencias sobre casos de uso de IA, pero no participan en ella ni la Comisión ni todas las agencias. Resulta cuestionable que la DG que ha impulsado la importante Propuesta de RIA desconozca los casos de uso existentes a nivel de la Unión y los problemas que pueden suscitar. Disponer de dicha información es básico para poder evaluar el impacto que la nueva Propuesta tendrá para la propia Administración de la Unión, así como para considerar posibles reglas específicas aplicables a la utilización de sistemas de IA por parte de las autoridades públicas.

No es de extrañar, por ello, que la referida Propuesta ignore prácticamente las especificidades del uso de IA por parte del sector público, y se centre sobre todo en el sector privado. El establecimiento por parte del art. 60 de la Propuesta de una base de datos centralizada en el seno de la Comisión con los casos de uso existentes tanto en el sector público como en el privado es un paso positivo para superar la falta de información actual, pero en el caso de las autoridades públicas podría extenderse a todos los sistemas de IA y no limitarse solo a los que merezcan la (esquiva) calificación de alto riesgo. Las eventuales objeciones competenciales que podrían oponerse a una regulación por parte del legislador europeo de la utilización de IA por parte de las Administraciones nacionales no serían aplicables a la Administración

⁸⁹ Art. 83 de la Propuesta, en relación con su art. 85.2 y su Anexo IX.

⁹⁰ Apartados 1 y 7 del Anexo III de la Propuesta.

de la propia Unión: el legislador europeo puede regular como desee a su propia Administración (art. 298 TFUE).

El estudio de campo también ha revelado que existe un notable interés y una utilización creciente de herramientas de IA por parte de la propia Administración de la Unión. Aunque es una utilización todavía esporádica, y que no responde a una política centralizada y consciente de la Comisión, sino que es fruto de las iniciativas individuales de las distintas Direcciones Generales y agencias, en ocasiones en colaboración con sus homólogas de otras regiones (como atestigua el caso de la EFSA y su colaboración con la EPA norteamericana). La utilizan tanto las autoridades que disponen de competencias decisorias propias (EUIPO, EFSA –en lo concerniente a la emisión de dictámenes científicos–), como las que proporcionan sistemas de información a los Estados miembros para que adopten las decisiones correspondientes (DG-Agri/ESA, eu-LISA).

El estudio confirma asimismo la importancia que tiene la externalización en esta materia, y las limitadas capacidades de la Administración de la Unión para desarrollar sus propios sistemas de IA. Con la excepción destacada de EUIPO, las demás autoridades han de acudir a la contratación pública (eu-LISA, por importes muy considerables) o a colaboradores externos no comerciales (DG-Agri/ESA, EFSA) para poderlos desarrollar⁹¹. Ello plantea en ocasiones, como se ha visto, el problema de no poder acceder a los datos de entrenamiento de los sistemas de *machine learning*, protegidos por secreto empresarial.

Los casos de uso examinados evidencian a su vez el gran potencial que la IA puede tener para la mejora de determinadas funciones administrativas, incrementando su calidad y eficacia y no solo reduciendo su coste. En algunos casos resulta ya inimaginable, incluso reproducible, no acudir a la IA. Es lo que sucede con la traducción automática de textos, en la que la Comisión está invirtiendo grandes cantidades

⁹¹ Sobre la situación en las agencias federales de EE.UU., *vid.* ENGSTROM, D. F.; *et al.* (2020), *op. cit.*, pp. 88 ss.: más de la mitad de los 157 casos de uso de IA identificados (53%) fueron desarrollados internamente por tecnólogos de la agencia, y casi otros tantos procedieron de fuentes externas, con un 33% procedente de fuentes comerciales privadas a través de contratación pública y un 14% resultante de colaboraciones no comerciales, incluyendo concursos organizados por las agencias y asociaciones entre la Administración y el mundo académico.

de recursos, según confirmaron varios entrevistados. En los supuestos estudiados, la IA permite reforzar notablemente el control de las subvenciones agrícolas y de las fronteras de la Unión, así como acelerar el proceso de evaluación de riesgos alimentarios y facilitar la coherencia de las decisiones relativas al registro de marcas.

El estudio revela que el uso de IA también plantea riesgos, riesgos que van más allá de la vulneración del derecho a la protección de datos personales y de los que los entrevistados eran perfectamente conscientes. Los casos analizados tienen un alcance limitado y no han aflorado supuestos de mal funcionamiento. Tampoco existe un reemplazo completo de los humanos, que son quienes acaban tomando las decisiones finales. Pero se ha constatado la inexistencia de un marco regulatorio específico o incluso de directrices internas en el seno de cada autoridad destinados a evitar la producción de tales riesgos, y que establezcan p. ej. la obligación de efectuar una evaluación de impacto antes de introducir un nuevo sistema de IA⁹², las condiciones que deberán exigirse a los contratistas a los que se encargue su desarrollo, las pruebas que deberán realizarse antes de su puesta en funcionamiento o las medidas destinadas a evitar una confianza excesiva del personal en los sistemas automatizados que se introduzcan (*automation bias*).

El estudio confirma, en fin, la importancia de las normas de procedimiento administrativo para evitar los riesgos apuntados, como se ha visto en la parte principal del presente trabajo. Las garantías procedimentales, lejos de ser vistas como rémoras de una Administración analógica ya superada, constituyen requisitos fundamentales de la nueva Administración digital, debiendo ser mantenidas y adaptadas en aquello que resulte necesario.

BIBLIOGRAFÍA CITADA

- ACUS (2021): *Administrative Conference Statement #20, Agency Use of Artificial Intelligence*, 86 Fed. Reg. 6616.
- ARROYO JIMÉNEZ, L. (2020), “Algoritmos y reglamentos”, *Almacén de Derecho*, <https://almacenederecho.org/algoritmos-y-reglamentos>.

⁹² Vid. *supra* el apartado 4.3.

- BOIX PALOP, A. (2020), “Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la administración para la adopción de decisiones”, *Revista de Derecho Público: Teoría y Método*, n° 1, 2020, pp. 223 ss.
- BUOLAMWINI, J.; GEBRU, T. (2018), “Gender shades: Intersectional accuracy disparities in commercial gender classification”, Conference on Fairness, Accountability and Transparency, *PMLR*, n° 81, pp. 77 ss.
- BUOSO, E. (2020), “Fully Automated Administrative Acts in the German legal System”, *European Review of Digital Administration & Law*, n° 1, pp. 113 ss.
- CALO, R.; CITRON, D. K. (2021), “The Automated Administrative State: A Crisis of Legitimacy”, *Emory Law Journal*, n° 70, pp. 797 ss.
- CERRILLO I MARTÍNEZ, A. (2021), “Robots, asistentes virtuales y automatización de las administraciones públicas”, *Revista Galega de Administración Pública*, n° 61, pp. 271 ss.
- CITRON, D. K. (2008), “Technological Due Process”, *Washington Law Review*, n° 85, pp. 1249 ss.
- COGLIANESE, C. (2020): *A Framework for Governmental Use of Machine Learning*, Report for the Administrative Conference of the United States, ACUS, Washington.
- COMISIÓN EUROPEA (2019): *Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones—Generar confianza en la inteligencia artificial centrada en el ser humano*, COM(2019)168, 8 de abril de 2019.
- COSTANZA-CHOCK, S.; RAJI, I. D.; BUOLAMWINI, J. (2022), “Who Audits the Auditors? Recommendations from a field scan of the algorithmic auditing ecosystem”, *FAccT’22*.
- COTINO HUESO, L. (2021), “El nuevo reglamento de Administración electrónica, que no innova en tiempos de transformación digital”, *Revista Catalana de Dret Públic*, n° 63, pp. 117 ss.
- DERAVE, C.; GENICOT, N.; HETMANSKA, N. (2022), “The Risks of Trustworthy Artificial Intelligence: The Case of the European Travel Information and Authorisation System”, *European Journal of Risk Regulation*, n° 13(3), pp. 389 ss.
- ENGSTROM, D. F.; HO, D. E.; SHARKEY, C. M.; CUÉLLAR, M.-F. (2020): *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*, Report submitted to the Administrative Conference of the United State, ACUS, Washington.
- Eu-LISA (2018): *Shared Biometric Matching Service (sBMS) Feasibility Study – final report*, doi:10.2857/84504.
- EUROPEAN LAW INSTITUTE (2022): *Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration*.

- GALETTA, D.-U.; HOFMANN, H. C. H.; MIR, O.; ZILLER, J. (2015), "The General Principles of EU Administrative Procedural Law. An in-depth analysis", *Rivista italiana di diritto pubblico comunitario*, n° 5, pp. 1421 ss.
- GAMERO CASADO, E. (2021), "Compliance (o Cumplimiento Normativo) de desarrollos de Inteligencia Artificial para la toma de decisiones administrativas", *Diario La Ley*, n° 50, 19/04/2021.
- HUERGO LORA, A. (2020): "Una aproximación a los algoritmos desde el Derecho administrativo", en HUERGO LORA, A. (Dir.): *La regulación de los algoritmos*, Aranzadi Thomson Reuters, Cizur Menor, pp. 23 ss.
- HUERGO LORA, A. (2021), "Administraciones Públicas e inteligencia artificial: ¿más o menos discrecionalidad?", *El Cronista del Estado Social y Democrático de Derecho*, n° 96-97, pp. 78 ss.
- JASPERS, S.; DE TROYER, E.; AERTS, M. (2018): *Machine learning techniques for the automation of literature reviews and systematic reviews in EFSA*, EFSA supporting publication 2018:EN-1427, doi:10.2903/sp.efsa.2018.EN-1427.
- LUHMANN, N. (1966): *Recht und Automation in der öffentlichen Verwaltung*, Duncker & Humblot, Berlin.
- MIR PUIGPELAT, O.; HOFMANN, H. C. H.; SCHNEIDER, J.-P.; ZILLER, J.; AUBY, J.-B.; CRAIG, P.; CURTIN, D.; DELLA CANANEA, G.; GALETTA, D.-U.; MENDES, J.; STELKENS, U.; WIERZBOWSKI, M. (Dirs.) (2015): *Código ReNEUAL de procedimiento administrativo de la Unión Europea*, INAP, Madrid.
- MISURACA, G.; VAN NOORDT, C. (2020): *Overview of the use and impact of AI in public services in the EU*, EUR 30255 EN, Publications Office of the European Union, Luxembourg, doi:10.2760/039619.
- MULLIGAN, D. K.; BAMBERGER, K. A. (2019), "Procurement as Policy: Administrative Process for Machine Learning", *Berkeley Technology Law Journal*, n° 34, pp. 773 ss.
- PONCE SOLÉ, J. (2019), "Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico", *Revista General de Derecho Administrativo*, n° 50.
- REISMAN, D.; SCHULTZ, J.; CRAWFORD, K.; WHITTAKER, M. (2018): *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability*, AI Now Report.
- TANGI L.; VAN NOORDT C.; COMBETTO M.; GATTWINKEL D.; PIGNATELLI F. (2022): *AI Watch. European Landscape on the Use of Artificial Intelligence by the Public Sector*, EUR 31088 EN, Publications Office of the European Union, Luxembourg, doi:10.2760/39336, JRC129301.
- VALERO TORRIJOS, J. (2019), "Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración", *Revista Catalana de Dret Públic*, n° 58, pp. 82 ss.

- VAQUER CABALLERÍA, M. (2020): “¿Para qué sirve el procedimiento administrativo?”, en PAREJO ALFONSO, L.; VAQUER CABALLERÍA, M. (Dir.): *Estudios sobre el procedimiento administrativo. III. Instituciones*, Tirant lo Blanch, Valencia, pp. 53 ss.
- VAVOULA, N. (2021), “Artificial Intelligence (AI) at Schengen Borders: Automated Processing, Algorithmic Profiling and Facial Recognition in the Era of Techno-Solutionism”, *European Journal of Migration and Law*, <https://ssrn.com/abstract=3950389>.

Capítulo XVII

Hacia un marco regulatorio holístico de la Inteligencia Artificial fiable en la Unión Europea: una perspectiva científico-técnica

EMILIA GÓMEZ, ISABELLE HUPONT TORRES,
IGNACIO SÁNCHEZ, DAVID FERNÁNDEZ LLORCA
European Commission, Joint Research Centre

Resumen: El enfoque de la Unión Europea sobre la Inteligencia Artificial (IA) se centra en la excelencia y la confianza, con el objetivo de impulsar la investigación y la capacidad industrial al tiempo que se garantiza la seguridad y los derechos fundamentales. Desde el año 2018, la Comisión Europea, en colaboración con sus colegisladores, ha ido dando una serie de pasos para desplegar un marco regulatorio holístico en torno a los sistemas de Inteligencia Artificial, incluyendo tres propuestas legislativas: la Ley de IA (*AI Act*), la Directiva sobre responsabilidad en materia de IA (*AI Liability Directive*) y la revisión de la Directiva sobre responsabilidad por los daños causados por productos defectuosos (*Product Liability Directive*). En este trabajo se presentan, de forma general, y desde una perspectiva científico-técnica, los aspectos principales de estas tres iniciativas, y se discuten los pasos futuros necesarios para que estas propuestas acaben consolidándose para definir el futuro de la Inteligencia Artificial fiable en la UE.

Abstract: The European Union's approach to Artificial Intelligence (AI) focuses on excellence and trust, with the aim of boosting research and industrial capacity while ensuring safety and fundamental rights. Since 2018, the European Commission, in collaboration with co-legislators, has been taking a series of steps to deploy a holistic regulatory framework around Artificial Intelligence systems, including three legislative proposals: the Artificial Intelligence Act (AI Act), the Directive on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive) and the revision of the Directive on liability for defective products (Product Liability Directive). This paper outlines the main aspects of these three initiatives from a scientific and technical perspective and discusses the future steps necessary for these proposals to be consolidated in order to define the future of a trustworthy Artificial Intelligence in the EU.

Palabras-clave: Inteligencia Artificial Fiable, Ley de Inteligencia Artificial, Responsabilidad de la Inteligencia Artificial, Directiva de Responsabilidad por Productos Defectuosos, Unión Europea.

Keywords: Trustworthy Artificial Intelligence, Artificial Intelligence Act, Artificial Intelligence Liability, Product Liability Directive, European Union.

SUMARIO.- 1. EL ENFOQUE EUROPEO PARA LA INTELIGENCIA ARTIFICIAL. - 2. LA LEY DE INTELIGENCIA ARTIFICIAL. - 3. LA DIRECTIVA SOBRE RESPONSABILIDAD EN MATERIA DE INTELIGENCIA ARTIFICIAL. - 4. LA REVISIÓN DE

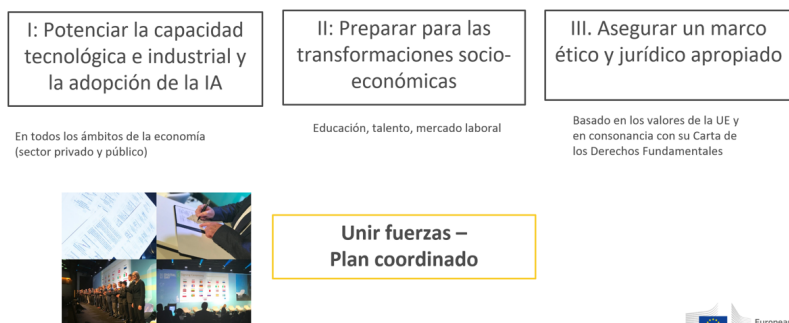
1.- EL ENFOQUE EUROPEO PARA LA INTELIGENCIA ARTIFICIAL

Los sistemas de inteligencia artificial (IA) son sistemas capaces de analizar su entorno y pasar a la acción, con cierto grado de autonomía, con el fin de alcanzar objetivos específicos (Annoni, 2018). El desarrollo de este tipo de sistemas ha evolucionado notablemente en los últimos años con el auge del área del aprendizaje automático, potenciado por la creciente disponibilidad de tres elementos esenciales: datos, capacidades de computación y el diseño de nuevos algoritmos de aprendizaje profundo basados en redes neuronales.

La Comisión Europea publicó en 2018 una Comunicación (Comisión Europea, 2018) denominada *Inteligencia Artificial para Europa* que definía tres objetivos fundamentales para una estrategia europea de IA:

1. Potenciar la capacidad tecnológica e industrial de la Unión Europea (UE) e impulsar la adopción de la IA en todos los ámbitos de la economía, tanto en el sector privado como en el público. Este objetivo incluye una serie de inversiones en investigación e innovación y en la mejora del acceso a los datos.
2. Preparar a Europa para las transformaciones socioeconómicas que origina la IA, fomentando la modernización de los sistemas de educación y formación, favoreciendo el talento, previendo los cambios en el mercado laboral y prestando apoyo a las transiciones que se operen en él y a la adaptación de los sistemas de protección social.
3. Garantizar el establecimiento de un marco ético y jurídico apropiado, basado en los valores de la Unión y en consonancia con la Carta de los Derechos Fundamentales de la UE.

Figura 1. Ilustración de los tres pilares de la estrategia europea de IA definidos por la Comunicación COM(2018)237



Posteriormente, la IA se estableció como una de las áreas prioritarias en los primeros cien días de mandato de la Presidenta Von Der Leyen. En febrero de 2020, una segunda Comunicación (Comisión Europea, 2020) denominada *Libro Blanco sobre la Inteligencia Artificial – un enfoque europeo orientado a la excelencia y la confianza* definía dos pilares fundamentales de la IA en Europa: la excelencia, con el diseño de un Plan Coordinado de IA con los Estados Miembros (Comisión Europea, 2018b), y la confianza, con el desarrollo por parte de un grupo de expertos de alto nivel de unas directrices éticas para una IA fiable (Comisión Europea, 2019).

Un año más tarde, y en el contexto de la situación global de pandemia y su gran impacto socioeconómico, la Comisión reiteró su apuesta por la IA definiendo en 2021 un *Paquete de IA* (Comisión Europea, 2021a) que incluía en primer lugar una Comunicación (Comisión Europea, 2021c) para *fomentar un planteamiento europeo en materia de IA*. En dicho paquete se reforzaba y adaptaba al contexto actual el ecosistema europeo de excelencia en IA con una revisión del plan coordinado de la IA con los Estados Miembros.

El paquete, dentro del pilar relativo al ecosistema de confianza, presentaba la primera propuesta legislativa a nivel mundial en materia IA, la Ley de Inteligencia Artificial o *AI Act*, reforzando la apuesta por un ecosistema de confianza. Dicho ecosistema de confianza se ha enriquecido posteriormente con una armonización específica de las normas nacionales de responsabilidad en materia de IA,

la Directiva sobre responsabilidad en materia de IA o *AI Liability Directive* (Comisión Europea, 2022a) y una propuesta de revisión de la Directiva sobre responsabilidad por los daños causados por productos defectuosos o *Product Liability Directive - PLD* (Comisión Europea, 2022b).

Este trabajo se centra en este pilar de la estrategia europea de IA, el ecosistema de confianza, para introducir sus diferentes iniciativas legislativas y reflexionar sobre los retos científico-técnicos específicos de este tipo de sistemas.

2.- CARACTERÍSTICAS DE LOS SISTEMAS DE INTELIGENCIA ARTIFICIAL

La IA aporta beneficios a los consumidores, los negocios y la sociedad, pero también crea algunos riesgos. Por ejemplo, en la consulta pública del Libro Blanco de la IA (Comisión Europea 2020), el 90% de los encuestados estaban preocupados por que la IA pudiera violar los derechos fundamentales (ej. los sistemas basados en biometría), y el 82% de los encuestados mencionaron los riesgos a la seguridad (ej. en la salud).

Algunos de estos riesgos se derivan de determinadas características prevalentes en los sistemas de IA como son su opacidad, complejidad, impredecibilidad, ausencia de causalidad, autonomía, en algunos casos la capacidad de autoaprendizaje en línea, o el hecho de que muchas de las técnicas actuales de aprendizaje automático se basan en datos. Dichas características están descritas en detalle en (Fernández Llorca, 2023) y (Estevez-Almenzar, 2022) y se resumen a continuación.

La **opacidad** de los sistemas de IA, conocida como el paradigma de la “caja negra”, se define como la imposibilidad o dificultad de comprensión del sistema. Incluso cuando los desarrolladores o los propios creadores del sistema de IA examinan por ejemplo un modelo de aprendizaje automático (ej. código fuente, proceso de entrenamiento, arquitectura, parámetros, datos) es difícil o imposible entender cómo los diferentes componentes se combinan para generar una decisión, salida o predicción. Dicha dificultad o imposibilidad se debe a la incapacidad de generar razonamiento a escala humana a partir de sis-

temas que, en la actualidad, tienen un alto grado de **complejidad** (los modelos de aprendizaje automático basados en redes neuronales, por ejemplo, se componen de decenas de millones de parámetros). Estas características intrínsecas de los actuales sistemas de IA han motivado el desarrollo de requisitos de transparencia para asegurar la confianza y la responsabilidad en los sistemas de IA.

La **impredecibilidad** de un sistema de IA que aprende de datos se debe principalmente a dos razones. En primer lugar, en el caso de que el conjunto de datos no sea lo suficientemente representativo del problema que debe abordar el sistema de IA. Independientemente de la capacidad de generalización que tenga el sistema y de su proceso de entrenamiento a partir de datos, la solución encontrada en las regiones mal representadas con los datos disponibles generará resultados imprevisibles, ya que el sistema no se ha entrenado bien en dichas regiones. Por otro lado, puede darse un problema de falta de generalización o sobreajuste (*overfitting*). Esto es, el sistema se ha ajustado de forma demasiado literal a los pocos datos que utiliza (se dice que ha memorizado) en lugar de haber obtenido una función general aplicable a datos no usados durante el entrenamiento.

La **capacidad de autoaprendizaje en línea** de algunos sistemas de IA implica que, por un lado, el sistema es capaz de “auto aprender” por sí mismo, por ejemplo, reconociendo patrones en los datos de entrenamiento sin supervisión humana. Por otro lado, el aprendizaje “en línea” implica una adaptación del sistema de IA mientras se encuentra en operación. Esta característica permite su adaptación y mejora constante del sistema según está en uso, y es fundamental en algunos ámbitos de aplicación como los sistemas de recomendación o de generación de perfil de usuario. Sin embargo, en casos donde la seguridad de las personas está en juego, el autoaprendizaje en línea puede conllevar cambios drásticos en el comportamiento del sistema.

Por último, la **autonomía** de un sistema de IA es la capacidad del sistema para operar con la ausencia o un bajo grado de intervención externa como por ejemplo la intervención humana (Estevez-Almenzar, 2022). Si bien la autonomía es una decisión de diseño, solo los sistemas más avanzados (normalmente diseñados con sistemas de IA muy sofisticados y complejos) son capaces de operar sin necesidad

de intervención o supervisión humana. Es por ello que aunque la autonomía no sea una característica intrínseca de la IA, se suele tratar como tal.

Las características mencionadas se añaden y relacionan con algunos retos que tienen que abordar las políticas públicas de IA como la aplicación de políticas, la certitud legal, la falta de confianza de la sociedad o el riesgo de la fragmentación regulatoria en los diferentes países. Finalmente, dichos riesgos y retos se encuentran en sistemas de IA que operan en diferentes sectores, desde la educación hasta la seguridad, lo que justifica un enfoque transversal y no sectorial para la regulación de dichos sistemas.

3.- LA LEY DE INTELIGENCIA ARTIFICIAL (AI ACT)

La ley de Inteligencia Artificial publicada por la Comisión Europea el 21 de abril de 2021, es una propuesta legislativa horizontal que se basa en los siguientes principios. Por un lado, en la legislación del mercado interno (basada principalmente en el Art. 114 del Tratado de Funcionamiento de la UE) que establece las normas para poner en el mercado o en servicio los sistemas de IA. Por otro lado, la ley de Inteligencia Artificial se encuentra alineada con la legislación relativa a la seguridad de producto, por ejemplo, IA integrada en diferentes productos.

Como característica distintiva, se propone una aproximación por capas o niveles de riesgo, donde no se regula la tecnología en sí sino casos de usos concretos, con el objetivo de abordar los riesgos a la salud, a la seguridad y a los derechos fundamentales asociados a su uso. Finalmente, la aplicación de la regulación no depende del origen de los productos, estableciendo unas reglas independientes del origen del proveedor o usuario del sistema de IA.

3.1.- Definición de IA

El artículo 3 de la propuesta de la Comisión Europea para la Ley de Inteligencia Artificial define «*Sistema de inteligencia artificial (sistema de IA)*»: *el software que se desarrolla empleando una o varias de las técnicas y estrategias que figuran en el anexo I y que puede, para*

un conjunto determinado de objetivos definidos por seres humanos, generar información de salida como contenidos, predicciones, recomendaciones o decisiones que influyan en los entornos con los que interactúa.

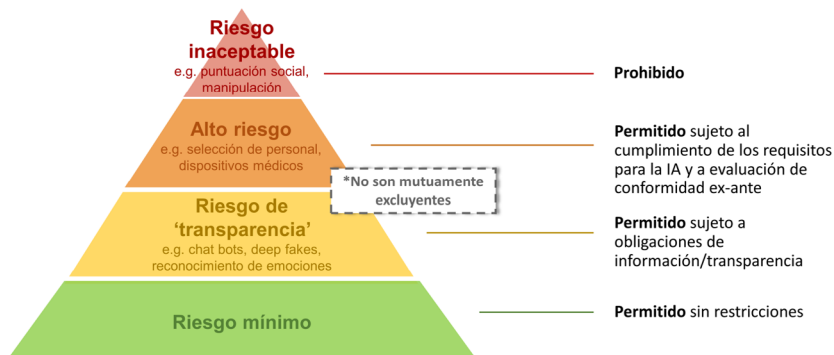
Esta definición engloba una serie de requisitos. Primeramente, se trata de una definición general que busca cubrir tanto los desarrollos actuales como futuros, incorporando técnicas del ámbito de la IA simbólica, los sistemas aprendizaje automático (incluyendo los sistemas de aprendizaje profundo basados en arquitecturas de redes neuronales) y los sistemas mixtos. El Anexo I de la propuesta legislativa de la Comisión incluye una lista detallada de técnicas y aproximaciones, que podrían adaptarse en el futuro si fuera necesario.

3.2.- Niveles de riesgo y sus requisitos

El citado artículo 3 establece además el concepto de «Finalidad prevista» del sistema de IA como *el uso para el que un proveedor concibe un sistema de IA, incluido el contexto y las condiciones de uso concretas, según la información facilitada por el proveedor en las instrucciones de uso, los materiales y las declaraciones de promoción y venta, y la documentación técnica.*

Este concepto es muy relevante en la legislación ya que, en función de la finalidad del sistema, se derivarán los correspondientes niveles de riesgo y los requisitos que un sistema de IA debe cumplir antes de poder ser comercializado en Europa (Hupont, 2022). El reglamento define cuatro niveles de riesgo según la finalidad prevista del sistema, como se muestra en la Figura 2:

Figura 2. Ilustración de los niveles de riesgo definidos en la propuesta de la Comisión Europea para la Ley de Inteligencia Artificial, con sus respectivos requisitos.



1. **Prácticas prohibidas** según el reglamento (Título II Artículo 5). Aquí se incluyen sistemas de manipulación subliminal que puedan resultar en un daño físico o psicológico, sistemas que se aprovechen de las vulnerabilidades de las personas, sistemas en los que las autoridades realicen una puntuación social, y los sistemas de identificación biométrica remotos, en espacios públicos, utilizados para la seguridad, con algunas excepciones.
2. **Sistemas de IA de alto riesgo** (Título II, Capítulo 1 y Anexos II y III), que se permiten siempre y cuando dichos sistemas cumplan una serie de requisitos y pasen por un procedimiento de evaluación de conformidad ex-ante. Dichos sistemas incluyen los componentes de seguridad de productos regulados y ciertos sistemas de IA en ámbitos como el acceso a servicios o beneficios públicos, la inmigración, el asilo y el control de fronteras, la educación, el empleo, la identificación y caracterización biométrica, el cumplimiento de la ley o la gestión y operación de infraestructuras críticas. Los requisitos para sistemas de IA de alto riesgo están definidos en el Título II, Capítulo 2, y dependen de la finalidad prevista. Entre los mismos se incluyen: establecer e implementar un sistema de gestión de riesgos; utilizar datos de alta calidad para el entrenamiento, validación y evaluación de los sistemas de IA; creación de documentación técnica relevante y registros de datos (*logs*) para la auditabilidad; asegurar el

grado de transparencia adecuado y proporcionar al usuario la información necesaria sobre las capacidades y limitaciones del sistema de IA; asegurar una vigilancia humana; y proporcionar una precisión, solidez y ciberseguridad adecuadas. El proveedor de un sistema de IA de alto riesgo (Título II, Capítulo 4, Artículo 49) debe, por tanto, determinar el nivel de riesgo de su sistema, asegurar el diseño y el desarrollo del mismo de acuerdo con los requisitos de la regulación, someterse a un procedimiento de validación de conformidad, y firmar dicha declaración antes de poner el sistema en el mercado o en servicio. A su vez, el usuario de dicho sistema tiene una serie de obligaciones (Título III, Capítulo 3) como la de operar el sistema según las instrucciones de uso, asegurar la vigilancia humana, efectuar un seguimiento del sistema para detectar posibles riesgos, informar al proveedor o distribuidor sobre posibles incidentes o deficiencias en el funcionamiento y respetar las obligaciones legales, como aquellas derivadas de la Ley General de Protección de Datos.

3. **Sistemas de IA con riesgo relacionado con la transparencia**, que están igualmente permitidos pero sujetos a obligaciones de información y transparencia. Dichas obligaciones, definidas en el Artículo 52, incluyen el notificar a las personas que están interaccionando con un sistema de IA o expuestas a un reconocimiento de emociones o categorización biométrica, así como aplicar etiquetas a las *ultrafalsificaciones* o *deep fakes*.
4. **Sistemas de IA con riesgo mínimo**, los cuales no están sujetos a restricciones. Aun así, se establecen códigos de conducta (Artículo 69) no obligatorios de ciertos requisitos mencionados anteriormente.

La Ley de Inteligencia Artificial se encuentra actualmente en proceso de negociación con los legisladores, el Parlamento Europeo y el Consejo. Por tanto, la propuesta puede ser ligeramente modificada y algunos aspectos están actualmente en discusión. Una vez adoptada, se establecerá un proceso de transición de dos años hasta ser aplicada. En paralelo, se están elaborando estándares armonizados para guiar y ayudar a proveedores y usuarios a cumplir con los requisitos establecidos en la regulación.

4.- LA DIRECTIVA SOBRE RESPONSABILIDAD EN MATERIA DE INTELIGENCIA ARTIFICIAL (*AI LIABILITY DIRECTIVE*)

El 28 de septiembre de 2022 la Comisión Europea adoptó, por primera vez, una propuesta de armonización específica de las normas nacionales de responsabilidad en materia de Inteligencia Artificial, la Directiva sobre responsabilidad en materia de IA o *AI Liability Directive* (Comisión Europea, 2022a) con el objetivo principal de que los estándares de protección para las víctimas perjudicadas por productos o servicios desplegados mediante Inteligencia Artificial (IA) sean los mismos que en los casos donde el perjuicio viene de productos defectuosos sin IA. Esta propuesta está alineada con los objetivos establecidos en el Libro Blanco sobre la Inteligencia Artificial (Comisión Europea, 2020) y con la propuesta de Ley de Inteligencia Artificial o *AI Act* (Comisión Europea, 2021). No obstante, esta norma va necesariamente más lejos que la Ley de la Inteligencia Artificial, en tanto en cuanto debe contemplar mecanismos para todos los sistemas de IA independientemente de su nivel de riesgo.

Por un lado, se asume que los sistemas basados en Inteligencia Artificial tienen ciertas características intrínsecas resumidas en el apartado 2 y descritas en detalle en (Fernández Llorca, 2023) tales como opacidad, ausencia de causalidad, impredecibilidad y, en algunos casos, la capacidad de autoaprendizaje en línea, que añaden una gran complejidad y dificultad a la hora de demostrar posibles comportamientos negligentes (esto es, faltas), así como la relación causal con el daño generado. Por otro lado, se entiende que independientemente de las regulaciones específicas de Inteligencia Artificial o de las regulaciones sectoriales de seguridad (por ejemplo, para el caso de los sistemas de conducción autónoma (Fernández Llorca, 2023b)), los sistemas basados en Inteligencia Artificial sean de alto riesgo o no, una vez en el mercado pueden producir daños a personas que deben poder activar mecanismos de responsabilidad para buscar compensación. La norma pretende complementar aquellos casos que estén fuera del ámbito de aplicación de la Directiva sobre responsabilidad por productos defectuosos (basada en defecto y relación causal), esto es, casos en los que los daños a la víctima se deben a un comportamiento negligente.

Las principales líneas de actuación son dos. Por un lado, se establecen normas uniformes para facilitar el acceso a información y documentación por parte de las víctimas. A este respecto, el apartado 1 del Artículo 3 de la Directiva establece que un tribunal puede ordenar la revelación de evidencias pertinentes sobre los sistemas de IA de alto riesgo de los que se sospeche que han causado daños. Las solicitudes de revelación de pruebas deben estar respaldadas por hechos y evidencias suficientes para establecer la verosimilitud de la reclamación de daños y perjuicios contemplada. En el apartado 4 del mismo Artículo se contempla el legítimo interés de todas las partes involucradas para que el derecho al acceso a evidencias por parte de la víctima no conlleve riesgos a la revelación de secretos industriales. Desde un punto de vista técnico cabe resaltar que lo que aquí se denomina como “evidencia pertinente” puede requerir distintos niveles de acceso a documentación, e incluso a registros de datos (*logs*) dependiendo del caso. En los escenarios más complejos, la documentación técnica necesaria para poder demostrar un comportamiento negligente por parte del responsable del sistema basado en IA podría tener que incluir datos tales como la arquitectura del sistema, bases de datos utilizadas para entrenar, validar y testear los modelos de aprendizaje, formato y registros de datos inmediatamente anteriores y posteriores al accidente, etc. Estos datos deberían ser legibles por parte de una opinión experta.

La segunda línea de actuación de la Directiva se centra en tratar de aliviar la carga de la prueba para la víctima. Para ello se desarrolla el Artículo 4 que presupone una relación causal entre la falta o comportamiento negligente y el daño, en diversos escenarios, por ejemplo, en el caso de que el demandado no haya proporcionado la evidencia necesaria establecida en el Artículo 3. También se asume en los casos en los que la falta pueda considerarse como la causa más probable del daño a partir de las circunstancias generales del caso. Para los sistemas basados en IA de alto riesgo, esta presunción se aplica siempre y cuando se pueda demostrar que el responsable del sistema no ha cumplido con las obligaciones establecidas en la Ley de Inteligencia Artificial. Para el resto de los sistemas (de menor riesgo), la presunción solo se aplica si el tribunal considera excesivamente difícil para la víctima establecer dicha relación causal. Este enfoque es lo suficientemente flexible para poder aplicarse caso a caso. Sin

embargo, es muy probable que según aumente la escala de casos de daños producidos por sistemas basados en IA, se haga necesario un desarrollo procedimental más específico para poder demostrar tanto el incumplimiento de obligaciones, como el grado de dificultad para demostrar la relación causal entre la falta y el daño.

5.- LA REVISIÓN DE LA DIRECTIVA DE RESPONSABILIDAD POR PRODUCTOS DEFECTUOSOS

Junto a la propuesta de armonización específica de las normas nacionales de responsabilidad en materia de Inteligencia Artificial, el mismo 28 de septiembre de 2022, la Comisión Europea presentó la propuesta de revisión de la Directiva sobre responsabilidad por los daños causados por productos defectuosos o *Product Liability Directive* (Comisión Europea, 2022b). Es importante destacar que la normativa vigente data del año 1985, por lo que su capacidad de adaptación al tipo de productos actuales es muy limitada. Las deficiencias de la norma se han ido identificando en numerosas ocasiones, incluyendo la evaluación dentro del marco del programa de adecuación y rendimiento (REFIT) llevada a cabo por la propia Comisión (Comisión Europea, 2018), así como el informe publicado por el Grupo de Expertos en Responsabilidad y Nuevas Tecnologías (Comisión Europea, 2019).

La propuesta de revisión amplía el rango de lo que se considera un producto dentro de la directiva, incluyendo software y servicios digitales, ya sean embebidos (integrados en el producto) o remotos (por conexión a Internet), que afecten al funcionamiento del producto. El concepto de “defecto” se amplía también para considerar la posibilidad de que el producto pueda integrar aprendizaje en línea (sistemas de aprendizaje máquina mientras el producto está en operación, o tras una actualización), así como requisitos específicos de seguridad y ciberseguridad. Por otro lado, también amplía el tipo de daño que puede ser compensado. Por ejemplo, se clarifica que los daños personales incluyen los daños a la salud psicológica reconocidos médicamente. Además, se añade la pérdida o la corrupción de datos que no se utilicen exclusivamente con fines profesionales.

Para poder garantizar que las personas tienen un nivel de protección similar ante el daño producido por productos de distinta complejidad, incluyendo sistemas basados en Inteligencia Artificial, la propuesta de revisión contempla en su Artículo 8 requisitos específicos para la revelación de información de los fabricantes de los productos. Si el fabricante no proporciona la información necesaria para que la víctima pueda demostrar que no se ha cumplido con el estándar de seguridad, los apartados 1 y 2 del Artículo 8 contemplan la presunción de defecto y relación causal con el daño. Al igual que sucede con la *AI Liability Directive*, determinar el nivel de información necesario en cada caso dependerá de muchos factores tales como la complejidad del sistema y del escenario donde se ha producido el daño. De igual forma, podría ser necesario disponer de documentación relacionada con la propia arquitectura del sistema de IA, los datos usados para entrenar, validar y testear, así como registros de datos de los sensores y capas intermedias instantes anteriores y posteriores al evento en el que se produjo el daño. Además, toda la información debe ser interpretable por una opinión experta.

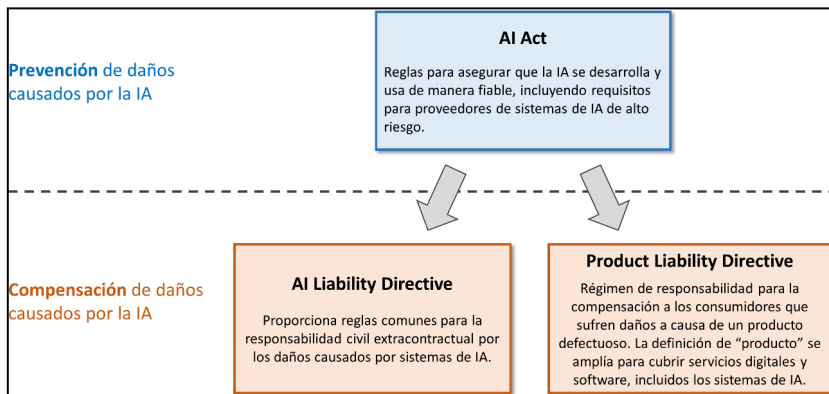
Por último, cabe destacar que la excepción clásica de responsabilidad para el fabricante cuando se puede demostrar que el defecto no existía en el momento en el que el producto se puso en el mercado, o debido a que el estado del conocimiento científico y técnico no era el adecuado, se deroga en el apartado 2 del Artículo 10 para los casos en los que el defecto se deba a un servicio, a un software, incluyendo actualizaciones, o la falta de actualizaciones de software necesarias para mantener la seguridad. Se pretende con este apartado incentivar a los fabricantes para que actualicen el software cuando sea posible, actuando de forma activa ante problemas de seguridad o ciberseguridad.

6.- CONCLUSIONES

En este estudio hemos resumido las iniciativas europeas de regulación en Inteligencia Artificial. La Figura 3 ilustra de forma gráfica dichas iniciativas, que conforman el ecosistema europeo de confianza para la IA. Dicha estructura propuesta holística aborda los retos que

plantea la IA y algunas de sus características técnicas como la opacidad y complejidad descritas en este trabajo.

Figura 3. Resumen de las iniciativas legislativas que conforman el ecosistema europeo de confianza en materia de Inteligencia Artificial.



El presente estudio se basa en las propuestas de la Comisión Europea publicadas en el boletín oficial de la Unión Europea en abril de 2021 (la Ley de Inteligencia Artificial) y septiembre (la directiva sobre responsabilidad en materia de IA, y la revisión de la directiva sobre responsabilidad por los daños causados por productos defectuosos) de 2022. Dichas propuestas legislativas se encuentran en fase de negociación en el Parlamento Europeo y Consejo Europeo, que deberá aprobarlas para que entren en vigor. Se trata de la primera estrategia legislativa holística a nivel mundial relativa a la inteligencia artificial, que pensamos tendrá un impacto muy positivo en la innovación responsable con la tecnología, con las personas en el centro, pensando en la protección de nuestros derechos fundamentales y seguridad física y mental.

BIBLIOGRAFÍA CITADA

ANNONI, A., BENCZUR, P., BERTOLDI, P., DELIPETREV, B., DE PRATO, G., FEIJOO, C., FERNANDEZ MACIAS, E., GOMEZ GUTIERREZ, E., IGLESIAS PORTELA, M., JUNKLEWITZ, H., LOPEZ COBO, M., MARTENS, B., FIGUEIREDO DO NASCIMENTO, S., NATIVI,

- S., POLVORA, A., SANCHEZ MARTIN, J.I., TOLAN, S., TUOMI, I. AND VESNIC ALUJEVIC, L. (2018), *Artificial Intelligence: A European Perspective*, Craglia, M. editor(s), EUR 29425 EN, Publications Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-97219-5, doi:10.2760/936974, JRC113826.
- Comisión Europea (2018^a): Comunicación: *Inteligencia artificial para Europa*. <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=COM%3A2018%3A237%3AFIN>
- Comisión Europea (2018b): Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. *Plan coordinado sobre la inteligencia artificial*. COM(2018) 795 final. https://eur-lex.europa.eu/resource.html?uri=cellar:22ee84bb-fa04-11e8-a96d-01aa75ed71a1.0022.02/DOC_1&format=PDF
- Comisión Europea (2018c), Commission Staff Working Document (SWD): *Evaluation of Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=SWD:2018:157:FIN>
- Comisión Europea (2019a). *Directrices éticas para una IA fiable*, Oficina de Publicaciones, <https://data.europa.eu/doi/10.2759/14078>
- Comisión Europea. 2019b. Directorate-General for Justice and Consumers, *Liability for artificial intelligence and other emerging digital technologies*, Publications Office, <https://data.europa.eu/doi/10.2838/573689>
- Comisión Europea (2020): *Libro Blanco sobre la Inteligencia Artificial – Un enfoque europeo orientado a la excelencia y la confianza*. https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf
- Comisión Europea (2021a): *Paquete de IA: Un enfoque europeo de la inteligencia artificial* <https://digital-strategy.ec.europa.eu/es/policies/european-approach-artificial-intelligence>
- Comisión Europea (2021b): *Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión*. COM(2021) 206 final. https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0008.02/DOC_2&format=PDF
- Comisión Europea (2021c): Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité económico y social europeo y al Comité de las Regiones. Fomentar un planteamiento europeo en materia de inteligencia artificial. https://eur-lex.europa.eu/resource.html?uri=cellar:01ff45fa-a375-11eb-9585-01aa75ed71a1.0023.02/DOC_1&format=PDF

- Comisión Europea (2022a): *Propuesta de directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial* (Directiva sobre responsabilidad en materia de IA). COM(2022) 496 final. <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52022PC0496&from=EN>
- Comisión Europea (2022b): *Propuesta de directiva del Parlamento Europeo y el Consejo sobre responsabilidad por los daños causados por productos defectuosos*. COM(2022) 495 final. <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52022PC0495&from=EN>
- FERNÁNDEZ LLORCA, D., CHARISI, V., HAMON, R., SÁNCHEZ, I., GÓMEZ, E. (2023): Liability regimes in the age of AI: a use-case driven analysis of the burden of proof. *Journal of Artificial Intelligence Research*.
- FERNÁNDEZ LLORCA, D., GÓMEZ, E. (2023b): Trustworthy Artificial Intelligence Requirements in the Autonomous Driving Domain, *Computer*, vol. 56, no. 2, pp. 29-39, doi: 10.1109/MC.2022.3212091.
- ESTEVEZ ALMENZAR, M., FERNANDEZ LLORCA, D., GOMEZ GU-TIERREZ, E. AND MARTINEZ PLUMED, F. (2022): *Glossary of human-centric artificial intelligence*, EUR 31113 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-53432-7 (online), doi:10.2760/860665 (online), JRC129614.
- FERNÁNDEZ LLORCA, D., CHARISI, V., HAMON, R., SÁNCHEZ, I., GÓMEZ, E. (2022): *Liability regimes in the age of AI: a use-case driven analysis of the burden of proof*. Arxiv preprint.
- FERNÁNDEZ LLORCA, D., GÓMEZ, E. (2021): *Trustworthy Autonomous Vehicles*, EUR 30942 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-46055-8 (online), doi:10.2760/120385 (online), JRC127051.
- HUPONT, I., TOLAN, S., GUNES, H., GÓMEZ, E. (2022): The Landscape of Facial Processing Applications in the Context of the European AI Act and the Development of Trustworthy Systems.

Capítulo XVIII

Hacia la regulación europea de la inteligencia artificial

ALEJANDRO HUERGO LORA

Catedrático de Derecho Administrativo. Universidad de Oviedo

Resumen: El texto destaca los aspectos comunes de las aplicaciones de la inteligencia artificial con relevancia jurídica. A partir de aquí se aíslan sus principales ventajas y riesgos. Estos resultados permiten identificar algunos parámetros de su marco jurídico, así como límites de su aplicación. Finalmente se abordan las características principales del proyecto de Reglamento Europeo sobre inteligencia artificial.

Palabras clave: Inteligencia artificial – Derecho de la UE – Intimidad – Discriminación – Protección de datos

Abstract: The text highlights the main features of artificial intelligence applications with legal relevance. Thereafter its main advantages and risks are isolated. These results make it possible to assert some parameters of its (future) legal framework, as well as limits of its application. Finally, the main characteristics of the draft European Regulation on artificial intelligence are addressed.

Keywords: Artificial intelligence – EU Law – Privacy – Discrimination – Data protection

SUMARIO: 1.- INTRODUCCIÓN. 2.- DE QUÉ ESTAMOS HABLANDO. 3.- LA IMPORTANCIA DEL CONTEXTO. 4.- SEGOS, ERRORES Y LÍMITES DE LAS PREDICCIONES ALGORÍTMICAS. 5.- EL ENCAJE JURÍDICO DE LAS PREDICCIONES ALGORÍTMICAS: ALGUNAS TENDENCIAS DEL REGLAMENTO EUROPEO DE INTELIGENCIA ARTIFICIAL.

1.- INTRODUCCIÓN

Creo que en los últimos treinta años ningún tema ha dado lugar a tantas publicaciones en la doctrina jurídica española como la inteligencia artificial a partir de 2020.¹ Las comunidades autónomas, el

¹ Este artículo se enmarca en el proyecto de investigación PID2021-126881OB-I00, “Herramientas algorítmicas para ciudadanos y Administraciones Públicas”,

“Derecho comunitario” o los contratos públicos han estado de moda en determinados momentos, pero ahora son constantes los libros colectivos o individuales sobre la inteligencia artificial en el Derecho público, abordada normalmente de forma general. Una de las causas son las ayudas a proyectos de investigación, que, aunque son un instrumento, acaban influyendo sobre el contenido de la propia investigación, no siempre en el modo pretendido por quienes dirigen esta política. Como consecuencia de los proyectos, y de la búsqueda de temas que resulten atractivos para los evaluadores o que encajen en listas de temas priorizados, la investigación se hace más procíclica (todo el mundo tiende a investigar simultáneamente sobre los mismos temas de moda) y se expresa a través de libros colectivos cuyo contenido tiende a repetirse (incluso con los mismos autores, combinados de forma diversa). Yo mismo he contribuido a ese aluvión de publicaciones y trato de llevar a buen puerto un proyecto de investigación sobre la materia, aunque en mi descargo debo añadir que cuando comencé a trabajar eran muy pocos los trabajos sobre este campo y también que he tratado de aportar un valor añadido que consiste en intentar entender y explicar qué son y para qué sirven las aplicaciones de inteligencia artificial (en lugar de trabajar a partir de bibliografía jurídica y pasar a explicar directamente los “riesgos” jurídicos de un fenómeno ignoto e inexplicado) y en adoptar un punto de vista un poco menos pesimista que la media.²

Aunque la ya mencionada abundancia de publicaciones ha dado motivo a chanzas en redes sociales, y no deja de resultar llamativa, tampoco carece de justificación, porque nos encontramos, sin duda, ante una gran innovación técnica y económica que se ha desarrollado en la última década, imprimiendo un giro nuevo al mundo digital y de internet. En una etapa histórica que, pese a las apariencias, no va

financiado por el Ministerio de Ciencia e Innovación, y en las actividades del grupo de investigación “Derecho Administrativo”, de la Universidad de Oviedo.

² Me remito a mis trabajos “Una aproximación a los algoritmos desde el Derecho administrativo”, en HUERGO LORA, Alejandro (Dir.)/DÍAZ GONZÁLEZ, Gustavo Manuel (Coor.), *La regulación de los algoritmos*, Aranzadi, Cizur Menor, 2020, págs. 23-87; “Administraciones Públicas e inteligencia artificial: ¿más o menos discrecionalidad?”, *El Cronista*, 96-97 (2021), págs. 78-95 o a la entrada “El proyecto de Reglamento sobre la Inteligencia Artificial” (publicada en el blog *Almacén de Derecho* el 17 de abril de 2021).

sobrada de avances científicos y tecnológicos y vive, en cierto modo, de las rentas de periodos anteriores, la inteligencia artificial es uno de los principales avances técnicos con inmediata repercusión en la economía. Así lo refleja la cotización de las grandes empresas que la explotan o, por mencionar otro indicador, la demanda de los títulos académicos relacionados con ella (empezando por los que tienen que ver con las matemáticas). Por ello, es normal que el Derecho se ocupe de la inteligencia artificial, porque es necesario y porque aquí puede aportar la doctrina mucho más valor añadido que repasando una y otra vez los temas tradicionales. Es cierto que escribir sobre la inteligencia artificial es mucho más expuesto que hacerlo sobre el acto administrativo y que, por qué no decirlo, el riesgo de hacer el ridículo es muy superior, pero el reto de la doctrina (no digamos de la que tiene su origen en las universidades) es aportar respuestas a los problemas reales, no jugar a las cuatro esquinas ni dar vueltas a las reglas de la geometría euclidiana en su versión jurídica.

2.- DE QUÉ ESTAMOS HABLANDO

Aunque esta cuestión la he desarrollado con detenimiento en mis anteriores trabajos, a los que he hecho referencia, me parece necesario repasar brevemente cuál es el objeto de estudio antes de centrar su análisis jurídico, porque ese repaso pone de manifiesto por sí solo cuáles son los problemas más relevantes.

De entrada, la expresión “inteligencia artificial”, si no es completamente inadecuada (como defiende más de uno), es una imagen muy imperfecta que introduce algunas distorsiones. La inteligencia artificial es una aspiración humana (o una pesadilla, según se vea) que siempre ha estado presente, como el sueño de volar. Pero, del mismo modo que este último se traduce a día de hoy en la aviación, que es una actividad importantísima pero que tiene poco que ver con ponerse unas alas y emprender el vuelo, la idea genérica de la inteligencia artificial se traduce en unas aplicaciones informáticas que nos ofrecen unos resultados que sustituyen al operador humano en ciertas tareas, o que emulan hasta cierto punto su labor, pero que lo hacen a través de unos procedimientos matemáticos e informáticos que poco tienen que ver con el pensamiento del ser humano.

Hablar de inteligencia artificial es un antropomorfismo, ciertamente justificado porque intenta realizar determinadas humanas ahorrando trabajo a las personas, pero inexacto porque el procedimiento que utilizan los ordenadores es distinto al de los humanos.

Los equívocos provocados por la expresión “inteligencia artificial” son fundamentalmente dos. El primero de ellos se refiere al ámbito, puesto que, frente a la “inteligencia artificial general”, que sustituiría al hombre en todo tipo de tareas, actualmente los ejemplos más logrados se centran en tareas relativamente sencillas que un humano llevaría a cabo en pocos segundos, como el reconocimiento facial o la confección o revisión de documentos rutinarios, que es uno de los principales campos de la llamada *Legaltech* o inteligencia artificial jurídica. No son desdeñables, con todo, los espectaculares avances realizados, por ejemplo, en la creación de vídeo con inteligencia artificial. Y el segundo equívoco, más relevante, es el que se refiere al cómo, a la naturaleza del proceso, porque las aplicaciones o modelos de inteligencia artificial ofrecen un resultado similar al del operador humano (un juicio, una predicción), pero al que se llega por un procedimiento distinto. Por ejemplo, un científico puede decir que alguien padece determinada enfermedad a partir de los resultados de un análisis en el que se detecta el virus o la bacteria que la provocan. En cambio, una aplicación de inteligencia artificial hace el mismo juicio después de analizar los datos de esa persona y llegar a la conclusión de que coinciden con los datos y características que suelen tener quienes padecen esa enfermedad, es decir, que existen correlaciones significativas entre esa persona y los positivos en esa enfermedad, más que entre esa persona y quienes no la padecen. (De hecho, un campo muy importante de la inteligencia artificial es el de la “emulación” o sustitución de pruebas o conocimientos. Una aplicación de inteligencia artificial puede averiguar si una calle está cortada sin tener constancia directa de ello, simplemente observando que los datos de tráfico que maneja indican que nadie circula por ella. Y puede detectar una enfermedad sin disponer de los resultados de un análisis que certifique la presencia del patógeno. Del mismo modo, y en sentido negativo, puede producir resultados que discriminen a los ciudadanos de una determinada etnia aunque no utilice expresamente ese criterio, si utiliza otros -de tipo físico, por ejemplo, o sociológico- que normalmente van unidos a esas personas).

Por otro lado, también hay un cierto equívoco, muchas veces subrayado, cuando se dice que la inteligencia artificial supone emitir juicios “sin intervención humana”. Sí hay intervención humana, lógicamente, para crear el programa que, a partir del análisis de datos del pasado, permitirá hacer juicios o predicciones futuras. Pero la aplicación del programa, una vez creado, no exige intervención humana (más allá de la introducción de los datos necesarios). Dicho de otro modo: cuando se aplica un modelo de inteligencia artificial a un análisis clínico (una resonancia, por ejemplo, o una ecografía) para saber si ese paciente tiene probabilidades de desarrollar un tumor, ese juicio de probabilidad se lleva a cabo sin intervención humana, a diferencia de lo que sucede cuando le enseñamos ese análisis a un médico. Pero eso no significa (es obvio) que, debido a esa falta de intervención humana, ese juicio sea “angelical” y libre de errores o sesgos, puesto que constituye el resultado de la aplicación de un artefacto matemático e informático en cuyo diseño, funcionamiento y premisas se han tomado muchas decisiones que pueden ser erróneas e incluso arbitrarias y dar lugar a errores en sus resultados.

Casi diariamente se publican nuevos avances o casos de aplicación de la inteligencia artificial en todos los campos, desde las finanzas al deporte pasando por la sanidad o la educación. También en el de la investigación científica, donde ha permitido hacer mapas de las proteínas³, identificar las mutaciones causantes de los distintos tipos de cáncer⁴ o avanzar en la búsqueda de remedios contra las superbacterias o bacterias resistentes a los antibióticos.⁵ Lo importante es identificar el elemento común a todas estas aplicaciones.

Las distintas aplicaciones suelen reducirse a un mismo esquema. Intentamos averiguar la respuesta a una pregunta: por ejemplo, qué personas son compradores potenciales de un producto (para dirigir a ellas la publicidad), qué averías se van a producir en una flota de

³ https://www.nationalgeographic.com.es/ciencia/inteligencia-artificial-crea-mapa-mas-completo-proteinas-humanas_17154

⁴ <https://elpais.com/ciencia/2021-07-28/un-sistema-de-inteligencia-artificial-identifica-las-mutaciones-causantes-del-cancer-para-cada-tipo-de-tumor.html>

⁵ MELO, M.C.R., MAASCH, J.R.M.A. & DE LA FUENTE-NÚÑEZ, C. “Accelerating antibiotic discovery through artificial intelligence”, *Commun Biol* 4, 1050 (2021).

aviones (para tratar de llevar los repuestos adecuados o incluso hacer un mantenimiento preventivo), qué clientes van a devolver un préstamo (para darles crédito a ellos y no a los demás), qué pacientes tienen más riesgo de desarrollar una determinada enfermedad (para aplicarles sólo a ellos una técnica de análisis que, por su coste y riesgos, no puede aplicarse a toda la población), qué empresas es más probable que estén cometiendo una infracción (para dirigir a ellas a los pocos inspectores disponibles) o cuál es el mejor momento -dentro de un margen de varias semanas- para realizar una cosecha y mejorar la calidad y el rendimiento.⁶ En todos los casos se trata de decisiones cuyo acierto permite obtener beneficios económicos (en unos casos) o sociales (cuando, por ejemplo, se detectan las infracciones realmente cometidas, o se logra identificar a los pacientes que necesitan tratamiento). Esas decisiones se vienen tomando normalmente con criterios que van desde el “ojo clínico” o la experiencia del responsable, hasta el azar (cuando, por ejemplo, se dirige la publicidad de manera general y aleatoria, sin seleccionar a sus destinatarios).

Las predicciones algorítmicas, que buscan dar mayor acierto a esas decisiones, se obtienen analizando el pasado: el historial de créditos de un banco, las infracciones detectadas y sancionadas en los últimos años, los datos de las personas que han desarrollado una determinada enfermedad o de quienes han comprado un determinado producto. Ello permite obtener un “retrato robot”, una serie de datos o características que se dan con más frecuencia en esas personas que en el resto. De lo que se trata es de utilizar ese retrato robot (que ha funcionado en el pasado) para intentar predecir el futuro. Del mismo modo, a través del análisis masivo de textos publicados en varios idiomas se han creado “traductores” informáticos (que todos utilizamos) que, aunque no conocen ni aplican las reglas de la gramática, saben, por

⁶ La definición de “sistema de inteligencia artificial” de la que parte el proyecto de Reglamento por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión, y que figura en su artículo 3, apartado 1, no se aparta de lo que se expone en el texto: “el software que se desarrolla empleando una o varias de las técnicas y estrategias que figuran en el anexo I y que puede, para un conjunto determinado de objetivos definidos por seres humanos, generar información de salida como contenidos, predicciones, recomendaciones o decisiones que influyan en los entornos con los que interactúa”.

ejemplo, cuál es la palabra inglesa que equivale a una española en un determinado contexto.

Este tipo de predicciones requieren tres elementos, que se han juntado en la última década: abundancia de datos analizables, capacidad computacional e instrumentos matemáticos (algoritmos) que permiten realizar ese análisis. Idealmente, disponemos así de predicciones “basadas en datos” frente al imperio de lo subjetivo, de los prejuicios o de los tópicos, a que estamos muy acostumbrados. Como predicciones que son, su éxito o fracaso depende de su tasa de acierto.

La inteligencia artificial, entendida como equivalente de predicciones algorítmicas, supone introducir un “baremo” para la toma de determinadas decisiones, es decir, una serie de datos o notas distintivas cuya aparición -o no- en un determinado sujeto o fenómeno nos ayudará a tomar una decisión respecto a él. Por ejemplo, concluiríamos que una empresa tiene más probabilidades de estar cometiendo una infracción en la medida en que reúna las características que en el pasado se observó que eran más frecuentes en las empresas infractoras. En unos casos, la utilización de ese baremo sustituye al uso, en la toma de esa decisión, de criterios subjetivos o intuitivos, al azar o a la pura y simple falta de criterios (por ejemplo, actualmente la legislación no obliga a razonar por qué se inspecciona una determinada empresa y no otra; la introducción de inteligencia artificial puede aportar un soporte racional a esa decisión). Se trata de un baremo que surge de los datos (del pasado), no de una decisión humana. En otros, la inteligencia artificial permite sustituir el uso de criterios que son objetivos pero que son muy caros o que requieren mucho tiempo (es lo que sucede en el campo científico, cuando la inteligencia artificial permite llegar a diagnósticos sin emplear análisis que serían muy costosos, o bien permite prescindir de miles o millones de experimentos cuya ejecución llevaría demasiado tiempo).

3.- LA IMPORTANCIA DEL CONTEXTO

Las aplicaciones de inteligencia artificial, entendidas como formas de producir, sin intervención humana (con los matices antes indicados), juicios o predicciones, tienen una trascendencia jurídica muy

diferente en función del contexto en que se empleen. El contexto pasa a ser, así, un elemento fundamental de su análisis jurídico.

Cuando la inteligencia artificial se utiliza para tomar decisiones que, aunque afecten a terceros, no interfieren en sus derechos o intereses jurídicamente protegidos, la inteligencia artificial entra dentro de la libertad del sujeto que decide utilizarla como guía para sus decisiones y plantea muy pocos problemas jurídicos. Así, cuando un inversor utiliza una aplicación de este tipo para decidir dónde invierte, o cuando una empresa se sirve de inteligencia artificial para decidir en qué ciudades abre tiendas, o cuando un abogado utiliza un programa para valorar la probabilidad de estimación de demanda y decidir, en función de ese análisis, presentarla o no presentarla, o hacerlo en uno u otro de los posibles foros, se trata de un modo, como otro posible, de adoptar una decisión libre y no se plantean mayores problemas a pesar de que, lógicamente, la decisión va a afectar a terceros (por ejemplo, a las empresas en cuyas acciones se invierte o no se invierte, a los dueños de locales en las ciudades en que se abren o no se abren tiendas, etc.).

Otra cosa sucede cuando el contexto jurídico en que se emplea la inteligencia artificial está regulado jurídicamente, lo que con frecuencia se debe a que el legislador considera que una de las partes de la relación jurídica (o de una relación jurídica potencial) se encuentra en situación de debilidad y requiere protección.

Por ejemplo, la utilización de inteligencia artificial para decidir a qué trabajadores se contrata o para decidir sobre su promoción laboral o sobre su despido sí es algo sobre lo que el Derecho tiene mucho que decir, porque se trata de decisiones que están sometidas a determinados principios y/o reglas jurídicas, empezando por la prohibición de discriminación. De ahí, por ejemplo, el artículo 64.4.d) del Estatuto de los Trabajadores (introducido por el Real Decreto-Ley 9/2021, de 11 de mayo), que obliga a las empresas a informar a los comités de empresa “de los parámetros, reglas e instrucciones en los que se basan los algoritmos o sistemas de inteligencia artificial que afectan a la toma de decisiones que pueden incidir en las condiciones de trabajo, el acceso y mantenimiento del empleo, incluida la elaboración de perfiles”.

Lo mismo pasa cuando la inteligencia artificial o “los algoritmos” se utilizan en relación con los consumidores, para decidir, por ejemplo, con quiénes se contrata o qué condiciones contractuales individualizadas se les ofrecen. A pesar de ello, es un campo en el que poco se ha avanzado jurídicamente, a pesar de que es notorio que las empresas utilizan algoritmos para determinar las condiciones contractuales o las ofertas que hacen a los potenciales clientes. Por ejemplo, recientemente se conocía cómo varios periódicos utilizan un algoritmo para decidir cuántas páginas dejan leer gratuitamente a través de su web, antes de exigir al lector que contrate una suscripción de pago si quiere seguir leyendo. A unos usuarios se les permite leer más páginas que a otros dependiendo de su mayor o menor predisposición a pagar (calculada en función de sus datos, es decir, en función de una predicción algorítmica), teniendo en cuenta que el periódico intenta maximizar tanto las suscripciones como la lectura de sus páginas (relevante para captar ingresos publicitarios), de modo que le interesa “no forzar” a aquellos lectores que no van a pagar, porque, si lo hace, no conseguirá una suscripción y perderá audiencia.⁷ El campo en el que seguramente más se ha avanzado en este campo es el del contrato de seguro.⁸

Un paso más se da cuando se pretende utilizar la inteligencia artificial en contextos especialmente sometidos al Derecho (o especialmente regulados), como es la actuación de los poderes públicos y, en particular, de las Administraciones Públicas. Éstas se encuentran sometidas al principio de legalidad (artículo 9.3 CE) y a la Ley y al Derecho (artículo 103.1 CE). Esto significa, como es bien sabido, que el Derecho no es para ellas un límite externo, como sucede con los sujetos privados, que pueden hacer todo aquello que no esté prohibido. Para las Administraciones Públicas el Derecho es la fuente habilitadora que les permite actuar, de modo que sólo pueden hacer aquello que les esté permitido. Es cierto que disponen de poderes que

⁷ Se trata del modelo conocido como Dynamic Meter, utilizado por el New York Times y descrito aquí: <https://open.nytimes.com/how-the-new-york-times-uses-machine-learning-to-make-its-paywall-smarter-e5771d5f46f8>.

⁸ Sobre esta materia, destaco especialmente los trabajos de M. L. Muñoz Paredes, empezando por “«Big Data» y contrato de seguro: los datos generados por los asegurados y su utilización por los aseguradores”, en HUERGO LORA, Alejandro (Dir.)/DÍAZ GONZÁLEZ, Gustavo Manuel (Coor.), *La regulación de los algoritmos*, Aranzadi, Cizur Menor, 2020, págs. 129-162.

son, con frecuencia, poderes muy amplios, pero las Administraciones deben justificar todas sus decisiones en una habilitación normativa, que en ocasiones determina con precisión la única decisión que pueden tomar (actos reglados) y, aunque no sea así (actos discrecionales), es necesario motivar el porqué de la decisión adoptada (dentro de las que eran posibles en ese marco normativo) y también el respeto de parámetros como el principio de proporcionalidad.

Es aquí donde resulta más criticable la utilización de modelos de inteligencia artificial, porque, si la norma establece que la decisión administrativa habrá de adoptarse cuando se dé un determinado supuesto de hecho, no cabe (salvo que la norma lo prevea o autorice) sustituir esa constatación por una predicción algorítmica. Esta es la gran dificultad para utilizar las predicciones algorítmicas como elemento en la toma de decisiones administrativas sin una previsión normativa en este sentido.⁹

⁹ E. GAMERO CASADO se ha referido a la motivación de las decisiones administrativas basadas en inteligencia artificial en “Necesidad de motivación e invalidez de los actos administrativos sustentados en inteligencia artificial o en algoritmos” (*Almacén de Derecho*, 4 de febrero de 2021). Es un enfoque distinto al que defiende en el texto y compatible con él. En la medida en que se utilicen algoritmos o predicciones algorítmicas para conformar el contenido de la decisión administrativa, tendrán que cumplirse unos requisitos de “calidad” (en sentido amplio), transparencia, control, etc., que E. GAMERO explica (y que están siendo codificados el Reglamento de Inteligencia Artificial). Así lo he puesto de manifiesto en mi trabajo de 2020: “en la medida en que el algoritmo aporte algún elemento nuevo en la aplicación de la norma, algún elemento que no sea replicable sin el algoritmo y comparable con los resultados de su aplicación, es exigible la total transparencia respecto a su funcionamiento, como cualquier elemento de un expediente administrativo” (página 75, nota 78). “En la medida en que los algoritmos predictivos contribuyan a determinar el contenido de las decisiones administrativas, estarán sujetos al mismo tipo de control que la decisión administrativa en su conjunto. Del mismo tipo que los informes técnicos pueden ser cuestionados en los recursos dirigidos contra las resoluciones dictadas con fundamento en los mismos (y ello por muy complejo que sea su contenido, o por muy especializada que sea la materia a la que se refieren), los modelos algorítmicos que sirven para determinar el contenido de la decisión administrativa pueden ser cuestionados por toda persona a la que dicha decisión afecte en sus derechos o intereses legítimos, o por cualquier otro legitimado para impugnarla (o por los órganos con funciones de control interno o externo). Un vicio o distorsión en el modelo provoca la disconformidad a derecho del acto o reglamento, exactamente igual que cualquier otro vicio” (página 80). “en la medida en que un algoritmo

Por ello, el campo de aplicación más adecuado para la inteligencia artificial es el de apoyo a decisiones administrativas de tipo instrumental o inicial, y se plasma, por ejemplo, en la selección de los casos en los que puede estar más indicada la iniciación de un procedimiento administrativo o la realización de una inspección, o bien de los supuestos en que puede estar más indicado llevar a cabo una determinada actuación material. Esa labor de identificación de “objetivos” puede producir efectos desfavorables o favorables para los interesados, según los casos. Son efectos desfavorables cuando los “señalados” van a ser objeto de una inspección, de la que podrá derivarse, en su caso, la iniciación de un procedimiento sancionador (aunque dicha iniciación se deberá, en su caso, a los hechos constatados en la inspección y no a la mera predicción algorítmica). Y son favorables cuando la predicción sirve para identificar a personas que, aunque no lo hayan manifestado, pueden estar en riesgo de suicidio, o de fracaso escolar, o de padecer violencia de género, o de desarrollar una determinada enfermedad, para hacerlas objeto de una actuación administrativa que intente adelantarse a esa situación.

Se trata en estos casos de utilizar la inteligencia artificial como soporte de “decisiones” administrativas que no constituyen actos administrativos en sentido estricto (no son resoluciones, sino actuaciones materiales o, en su caso, actos de trámite) y que se encuentran

sirva para tomar decisiones que afecten a un ciudadano, la Administración debe hacerlo público del mismo modo que cualquier otro documento que forme parte del expediente administrativo, en el que el algoritmo se integraría en virtud del artículo 70.1 LPAC, puesto que sirve «de antecedente y fundamento a la resolución administrativa» (página 85).

Lo que planteo en el texto es que, además de eso, la Administración no puede utilizar predicciones algorítmicas como fundamento de los actos administrativos cuando la programación normativa de éstos impone que se basen en la previa constatación de *hechos*. Otra cosa es que, con ciertas condiciones, se pueda utilizar una predicción algorítmica como *prueba indiciaria* de algún hecho o bien como soporte probatorio de un *pronóstico*, en aquellos actos o disposiciones que se basan precisamente en predicciones o valoraciones (por ejemplo, autorización o prohibición de fusiones empresariales en función de sus efectos esperados sobre la competencia en el mercado, o declaración de zonas inundables o sujetas a otro tipo de riesgos naturales). Sobre la primera de estas cuestiones, me remito a mi trabajo “Inteligencia artificial y sanciones administrativas”, en REBOLLO PUIG, M./HUERGO LORA, A./GUILLÉN CARAMÉS, J./CANO CAMPOS, T. (Dir.) *Anuario de Derecho Administrativo Sancionador* (2022), págs. 275-309, esp. págs. 297-298.

sometidas a un control jurídico muy limitado, hasta el punto de que en muchos casos no están formalizadas o no son recurribles autónomamente salvo que produzcan, por sí solas, efectos jurídicos directos sobre el particular. No me parece razonable exigir más requisitos a ese tipo de decisiones o actuaciones cuando se utiliza para su producción algún modelo de inteligencia artificial que cuando las adoptan humanos en función de criterios no explicados (porque no se exige), como sucede actualmente.

Algunas sentencias del Tribunal Supremo, no referidas a la inteligencia artificial pero sí relacionadas con ella, ofrecen puntos de interés. Así, en relación con la iniciación del procedimiento, la Sala Tercera se ha negado a anular una sanción por el hecho de que se haya iniciado el procedimiento sin ajustarse a los planes de inspección previamente aprobados, en los que no figuraba el sector de actividad al que se dedicaba el sancionado.¹⁰ En cambio, se ha mostrado reacia la misma Sala a admitir que decisiones administrativas que van más allá de la iniciación del procedimiento, como ocurre con las de autorización de entrada en domicilio para la búsqueda de pruebas de una infracción, se basen en consideraciones de tipo estadístico, referencia que se aproxima a las predicciones algorítmicas.¹¹

¹⁰ Sentencia de 19 de febrero de 2020 (recurso de casación 240/2018): “La introducción de los planes de actuación, de Control, parciales, pueden tener la finalidad de facilitar a la Administración Tributaria su funcionamiento, especialmente mediante la aplicación de programas informáticos que permitan un tratamiento más ágil de la fiscalización de un determinado sector de contribuyentes, pero su previsión, quizá no muy acertada, en normas legales y reglamentarias, no puede traducirse sin más en un derecho subjetivo del contribuyente a no ser investigado si no se encuentra incluido en dichos planes y programas, pues ello supondría el incumplimiento por parte de la Administración del deber de fiscalización de que todos los ciudadanos cumplan con el deber de contribuir, previsto en el artículo 31.1 de nuestra norma constitucional”.

¹¹ Sentencia de 1 de octubre de 2020 (recurso de casación 2966/2019): “[n]o pueden servir de base, para autorizar la entrada, los datos o informaciones generales o indefinidos procedentes de estadísticas, cálculos o, en general, de la comparación de la situación supuesta del titular del domicilio con la de otros indeterminados contribuyentes o grupos de estos, o con la media de sectores de actividad en todo el territorio nacional, sin especificación o segmentación detallada alguna que avale la seriedad de tales fuentes”.

4.- SESGOS, ERRORES Y LÍMITES DE LAS PREDICCIONES ALGORÍTMICAS

El concepto de sesgo como defecto inherente a las predicciones algorítmicas está siempre presente en su tratamiento doctrinal y en los instrumentos de *soft law* y proyectos normativos a ellas dedicados. Sin rechazarlo, es necesario ir más allá y colocarlo en el contexto -más amplio- de las limitaciones de dichas predicciones.

De entrada, hay que recordar que las decisiones humanas se encuentran sometidas a sesgos tanto o más profundos y abundantes que las algorítmicas y seguramente menos detectables. En todos aquellos casos (muy frecuentes) en los que existe un margen de decisión, es fácil constatar que el operador puede, de hecho, optar por varias decisiones, *todas ellas justificables en Derecho*, y que resulta un arcano completo el motivo real de la decisión tomada, que no necesariamente coincide con el exteriorizado en la motivación, que es la cobertura jurídica necesaria para superar los controles, pero que no tiene por qué ser el motivo real de la decisión. En este sentido, la insistencia en los sesgos algorítmicos puede resultar excesiva si se piensa en las alternativas (es decir, en las decisiones puramente humanas). Sin olvidar que, aunque las predicciones algorítmicas nacieron con la pretensión de superar la arbitrariedad típicamente humana, a través de auditorías algorítmicas se ha demostrado que muchas veces tienen sus propias arbitrariedades.

Más que de posibles errores, habría que preguntarse qué garantías de acierto existen y si las mismas justifican que la Administración tome auténticas decisiones sobre la base de estas predicciones. De ahí que, como vengo diciendo en este artículo, a falta de habilitación normativa para ello, las decisiones públicas se deben basar en la constatación del supuesto de hecho legalmente previsto, y no sustituirla por una predicción.

Las predicciones algorítmicas se basan en el análisis de datos que reflejan el pasado (más o menos próximo). El “algoritmo” *funciona* con los datos del pasado, en el sentido de que, con los datos pasados, “predice” adecuadamente lo que ocurrió. Por ejemplo: un modelo que establece qué características tienen las empresas que cometen más infracciones (y deben, por tanto, ser inspeccionadas en primer lugar) se basa en los datos del pasado (es decir, los datos de las infracciones

detectadas y las sanciones impuestas) y, *para ese tiempo pasado*, identifica adecuadamente a los infractores. Su aplicación práctica (puesto que nadie tiene interés en predecir el pasado, obviamente) está sometida a la condición de que el futuro se parezca al pasado o reproduzca sus patrones, lo que no siempre sucede (pensemos en rupturas como las que hemos vivido recientemente por el COVID o la guerra en Ucrania) o muchas veces no queremos que suceda (los patrones del pasado pueden responder a una realidad desigual que no responde a la igualdad real que el artículo 9.2 CE quiere promover). Para evitar que las desigualdades del pasado pasen al algoritmo y puedan producir desigualdades en el futuro, hace falta un esfuerzo positivo que consiste en aplicar técnicas matemáticas que permitan compensar las desigualdades de la muestra. Por ejemplo, si analizamos datos de crédito o de actividad empresarial, y existe en los datos pasados una participación femenina muy pequeña, es necesario compensarla mediante diferentes técnicas (que incluyen, por decirlo en términos groseros, “multiplicar” la participación femenina para que “pese” más de lo que se deriva de su escaso número real) a fin de evitar que el modelo produzca resultados sesgados. En definitiva, la creación de modelos o predicciones no es algo ciego o automático, sino una delicada labor que exige hacerse las preguntas adecuadas, convertir aquello que se intenta buscar en una serie de características verificables y preguntarse, también, qué errores pueden afectar al modelo, o qué daños puede provocar, para tratar de corregirlos o compensarlos. Del mismo modo que el vino tiene “química”, en el sentido de que el enólogo puede producir vinos muy distintos a partir de un mismo tipo de uva, el modelo algorítmico refleja muchas opciones tomadas durante su elaboración, lo que impide tomarlo como un oráculo que merezca una confianza total.

Los modelos algorítmicos no son neutros, aunque funcionen automáticamente. Su resultado depende, dramáticamente, de los datos que se analizan e integran en el modelo. Pueden ser datos parciales, anticuados, erróneos y eso lastra el modelo. Incluso un modelo correcto puede producir resultados disfuncionales si se aplica a un lugar diferente. Esto ocurre con claridad en modelos que se aplican a seres humanos y aplican parámetros físicos (estatura, peso, etc.), cuando se aplican a otro país en el que la aplicación de ese modelo puede conducir, por ejemplo, a discriminar a determinadas personas o grupos.

No olvidemos, por otro lado, que los modelos algorítmicos funcionan buscando y hallando correlaciones entre lo que se intenta averiguar (quién es más solvente, quién necesita atención especial, quién va a comprar el producto) y otras circunstancias que se pueden conocer con antelación (los títulos académicos, los ingresos anuales, el historial de relaciones con el banco, etc.). Hay características como el sexo, la raza o la religión, que no pueden utilizarse para dar un trato diferente porque ello violaría el artículo 14 CE, aunque el modelo pueda hallar correlaciones. También puede haber discriminaciones indirectas, cuando no se utiliza, para determinar el tratamiento de las personas, un factor directamente prohibido (por ejemplo, el sexo), pero se utiliza otro que, aunque aparentemente neutro, está muy correlacionado con un factor prohibido. Por ejemplo, cuando no se utiliza la raza, pero sí el código postal, siendo así que puede haber una elevada concentración étnica en determinadas zonas.

Las predicciones algorítmicas sirven para individualizar la actuación de empresas y Administraciones, de forma que, en lugar de actuar de un modo uniforme (y, por la misma razón, “ciego”), traten a cada sujeto del modo más adecuado a sus características, puestas de manifiesto en los datos analizados. No todas las personas que acceden a una determinada *web* ven el mismo contenido ni se les ofrecen las mismas condiciones contractuales. Llevado al ámbito público, también va a ser individualizado el tratamiento que reciban los ciudadanos en relación con diferentes servicios o instituciones. A la vez, esto significa que todo lo que haga o deje de hacer un sujeto, siempre que deje constancia digital, podrá ser utilizado en su contra, puesto que cabe la posibilidad de que un modelo algorítmico detecte correlaciones entre ese hecho y aquello que se intenta averiguar o a lo que aspira el ciudadano. Los ejemplos de Cathy O’Neil en *Armas de destrucción matemática* (2015) pueden quedarse cortos después de la pandemia y el auge de la enseñanza a distancia a través de plataformas. Los estudiantes digitales han dejado gran abundancia de datos sobre su nivel de atención, sus resultados educativos, su forma de escribir, etc. En determinadas plataformas, sobre todo las de uso gratuito, esos datos pueden ser cedidos a terceros (con el consentimiento de los usuarios, por supuesto, que cada vez que entran en una *web* consienten expresamente el uso de *cookies*, y no sólo las necesarias), y en el futuro esos

datos pueden determinar un trato diferente a la hora de acceder a fases sucesivas en la búsqueda de un puesto de trabajo, por ejemplo.¹²

Todo ello al margen de que con frecuencia es imposible saber si la predicción acierta o falla, porque, en el momento en que se utiliza, ya está transformando la realidad y, salvo que se aplique sólo a un número limitado de casos, no puede determinarse el nivel de acierto o compensar sus fallos. Cuando se inspecciona preferentemente a quienes son señalados por una predicción algorítmica, se van a encontrar más infracciones en ellos que en los sujetos no inspeccionados, lo que reforzará la predicción y hará que, aun en el caso de que inicialmente todos los sujetos cometieran infracciones en la misma medida (de modo que el algoritmo sea totalmente erróneo), los datos indicarán que esos sujetos tienden a cometer más infracciones y serán objeto de más inspecciones. Un claro ejemplo de profecía autocumplida. Por otro lado, si utilizamos una aplicación algorítmica para dar créditos, podemos verificar si las personas que reciben un préstamo lo devuelven (y, por tanto, podemos ver si el algoritmo es eficaz en el sentido de que reduce la tasa de morosidad), pero no podemos saber si las personas a las que se deniega financiación la hubieran devuelto el dinero o no.

Se pone de manifiesto, por consiguiente, la necesidad de aplicar controles de calidad a los algoritmos para tratar de reducir estos riesgos (este es el contenido básico del Reglamento europeo en tramitación) y también explica por qué es dudosa su utilización más allá de como soporte de decisiones instrumentales o aproximativas, que contribuyen a dirigir en una u otra dirección los medios de que dispone la Administración pero que no producen efectos directos sobre los ciudadanos.

¹² Una decisión de 22 de agosto de 2022 de un tribunal federal de Ohio (Case: 1:21-cv-00500-JPC) ha establecido, en una resolución sobre el inicio del proceso, que la Universidad del Estado de Cleveland, al exigir a los estudiantes que realizaban exámenes online que enseñaran su habitación a una cámara y estuvieran abiertos a la observación por la misma, violó la Cuarta Enmienda de la Constitución de Estados Unidos, que recoge el derecho a la inviolabilidad del domicilio. El tribunal realiza un test de proporcionalidad y ponderación entre los respectivos intereses jurídicamente protegidos y la proporcionalidad del medio empleado para luchar contra posibles fraudes.

5.- EL ENCAJE JURÍDICO DE LAS PREDICCIONES ALGORÍTMICAS: ALGUNAS TENDENCIAS DEL REGLAMENTO EUROPEO DE INTELIGENCIA ARTIFICIAL

No es posible en el espacio de este artículo sintetizar el Reglamento europeo de Inteligencia Artificial (conocido como AIA, *Artificial Intelligence Act*), que en este momento se encuentra en tramitación y que constituye el esfuerzo normativo más desarrollado y avanzado en el panorama internacional. Me limitaré a algunas pinceladas.

Como en otros ámbitos (me remito aquí a la entrada “Regulation versus Tort”, que publiqué en *Almacén de Derecho* el 12 de mayo de 2019), el enfoque jurídico puede ser preventivo o represivo. Se pueden establecer unas normas sobre los modelos de inteligencia artificial que obliguen a tener en cuenta datos adecuados, a dotar de transparencia al modelo, a dejar registro de toda su actividad para poder verificarla, etc. Eso es lo que hace el Reglamento a través de lo que he denominado genéricamente “normas de calidad”. También cabe (y es lo que ha sucedido hasta ahora en Estados Unidos) aplicar, allí donde se detecta que el “riesgo” se ha convertido en un “daño”, los instrumentos represivos habituales (responsabilidad civil y penal), lo que, en el Derecho norteamericano, suele terminar con una transacción y el pago de una indemnización.¹³ El miedo a esa persecución penal o civil puede producir un importante efecto disuasorio, que incluso lleve a los sujetos a imponer autónomamente reglas similares a las del Reglamento europeo (como ocurre, por ejemplo, con la nueva versión del documento *Microsoft Responsible AI Standard*, que recoge las instrucciones a los empleados de este gigante tecnológico que creen modelos de inteligencia artificial, hecha pública en junio de 2022).

El Reglamento se basa en que muchos modelos de inteligencia artificial son “de alto riesgo”, para imponerles una serie de requisitos o limitaciones (en relación con los datos, la transparencia, el control humano, etc.), cuyo cumplimiento está sujeto a diversos mecanismos de control y cuyo incumplimiento es objeto de sanciones. Hasta aquí nos

¹³ Algunos ejemplos en <https://www.deeplearning.ai/the-batch/the-hammer-drops/> (entrada de 6 de abril de 2022).

movemos en un esquema clásico de Derecho de policía: el riesgo es el título de legitimación que utiliza la Ley para imponer limitaciones y controles a las actividades privadas. Creo que, de todas formas, hace falta un esfuerzo importante para identificar y definir el “daño” que la inteligencia artificial puede producir y que es justamente lo que se intenta evitar. Cuando hablamos de otras actividades de riesgo, como las de tipo de industrial, está claro el tipo de daños que se intenta evitar (fallecimiento, daños corporales). Hace falta un esfuerzo teórico importante, en el campo de los derechos fundamentales, para identificar qué supondría un auténtico daño (que daría lugar a responsabilidad civil). No se trata sólo de los supuestos de discriminación (que el modelo algorítmico dé lugar a un trato diferente basado en características que no puede utilizarse con esa finalidad, como el sexo, la raza o la religión), o de errores derivados del algoritmo, sino que incluso sería discutible hasta qué punto es compatible con el derecho a la intimidad que se dé un trato diferente a las personas en función de actos previos que aparentemente no tienen ninguna relación con el contexto en el que se da ese trato diferente.¹⁴

El Derecho de protección de datos constituye un marco jurídico ineludible y el primero llamado a regular la inteligencia artificial. Sin embargo, como también he dicho en otro lugar, es necesario tener en cuenta que los datos con los que se crean (“entrenan”) los modelos algorítmicos son normalmente datos anonimizados, por lo que el control es limitado. Nada impide a un banco utilizar la masa de datos obtenida a través de su actividad crediticia, una vez “tachados” los nombres de los clientes, para intentar identificar cuáles son las características “objetivas” que permiten identificar a los buenos y malos

¹⁴ El derecho a la intimidad y a la inviolabilidad del domicilio, tal como es protegido por el artículo 8.2 del Convenio Europeo de Derechos Humanos, es la norma que el Tribunal de Distrito de La Haya, en su conocida sentencia de 5 de febrero de 2020 (asunto Syri) consideró incompatible con la utilización del sistema Syri para la detección de fraudes en ayudas sociales. La sentencia se basa en un test de proporcionalidad y ponderación entre la intimidad (en la que incide la utilización de un programa como Syri, que toma datos de distintas fuentes para detectar posibles fraudes) y los intereses públicos protegidos (detección del fraude) y llega a la conclusión de que se produce un desequilibrio, entre otras cosas debido a la falta de transparencia y a que no se especifican debidamente los límites de la utilización de los datos y éstos no se minimizan.

clientes. Una de las previsiones más importantes en esta materia procedentes del Derecho de protección de datos es la prohibición (relativa) de que las personas estén sujetas a decisiones basadas “únicamente en el tratamiento automatizado” (artículo 22 del RGPD), aunque su funcionalidad sea limitada porque muchas veces las predicciones no se aplican automáticamente y, por tanto, influyen en las decisiones más que producirlas en sentido estricto.

Por último, una característica muy destacable del Reglamento es que expresa principios y objetivos más que reglas operativas concretas. En lo que constituye su núcleo, que es la regulación de los sistemas de inteligencia artificial de alto riesgo, el Reglamento establece una serie de principios que pueden ser concretados o aplicados de diferente manera. Decir que en el diseño de un automóvil se deberá garantizar la “seguridad”, o que en la conducción se deberá actuar de modo seguro, es algo genérico, porque puede haber diferentes formas de entender la seguridad. Además, son conceptos que exigen un equilibrio. Es decir, proteger a ultranza la seguridad, o evitar a toda costa el riesgo de errores o discriminación, puede llevar a renunciar al modelo algorítmico, porque es la mejor forma de garantizar que no se produzca ningún fallo. Por eso el Reglamento habla, entre otras cosas, de un “sistema de gestión de riesgos”, en el que se acepten determinados riesgos residuales. Ese equilibrio puede ser diferente en unos casos o en otros (en función, por ejemplo, de los recursos mayores o menores de que disponga el sujeto que crea el algoritmo, de sus beneficios potenciales para el interés público o de que la gravedad de los riesgos). Teniendo en cuenta, además, que el Reglamento se basa, como regla general, en la técnica del “autocontrol” por parte del proveedor del sistema (sometido, por supuesto, a posible verificación posterior por la autoridad responsable), es difícil aplicar los mecanismos de disuasión más habituales (y a los que se remite el propio Reglamento) como las sanciones administrativas, porque la infracción estaría descrita en términos tan genéricos que difícilmente son compatibles con el principio de tipicidad.

Capítulo XIX

Inteligencia artificial y reconocimiento biométrico en el procedimiento administrativo

IGNACIO ALAMILLO DOMINGO¹

Doctor en Derecho.

Grupo de investigación iDerTec, de la Universidad de Murcia

Resumen: En este trabajo presentamos la oportunidad, y problemática asociada, de la aplicación de técnicas de inteligencia artificial para el reconocimiento biométrico, de interés atendiendo a la transformación del modelo de atención a los interesados. Para ello, se realizan precisiones terminológicas técnicas y jurídica, se introducen los principales casos de uso y se realizan propuestas para hacer efectivo el uso del reconocimiento biométrico basado en inteligencia artificial en el procedimiento administrativo.

Abstract: In this paper we introduce the opportunity, and associated problems, of the application of artificial intelligence techniques for biometric recognition, of interest attending to the transformation of the model of attention to interested parties. For this, technical and legal terminological clarifications are made, the main use cases are presented, and proposals are made to make effective the use of biometric recognition based on artificial intelligence in the administrative procedure.

Palabras-clave: inteligencia artificial, reconocimiento biométrico, identificación electrónica, identificación biométrica, verificación biométrica, firma electrónica biométrica

Key-words: artificial intelligence, biometric recognition, electronic identification, biometric identification, biometric verification, biometric signature

SUMARIO.- 1.INTRODUCCIÓN.- 2.PRECISIONES TERMINOLÓGICAS PREVIAS.- 2.1.- Precisiones desde la perspectiva tecnológica.- 2.1.- Precisiones desde la perspectiva jurídica.- 3.ALGUNOS CASOS DE USO DEL RECONOMIENTO BIOMÉTRICO BASADO EN INTELIGENCIA ARTIFICIAL EN EL PROCEDIMIENTO ADMINISTRATIVO.- 3.1.- Reconocimiento biométrico en la asistencia por canal telefónico.- 3.2.- Reconocimiento biométrico en la asistencia por videoconferencia.- 3.3. La captura electrónica de firma manuscrita.- 4.PROPUUESTAS PARA EL USO DEL RECONOCIMIENTO BIOMÉTRICO BASADO EN INTELIGENCIA ARTIFICIAL EN EL PROCEDIMIENTO ADMINISTRATIVO. BIBLIOGRAFÍA CITADA

¹ ORCID: 0000-0003-1426-4093

1.- INTRODUCCIÓN

El reconocimiento biométrico ha recibido, en tiempos recientes, un impulso gracias a la aplicación de nuevas técnicas de inteligencia artificial, habilitando nuevas oportunidades (y riesgos) relativas a la verificación de la identidad personal².

Como es sabido, la LPAC ordena la necesaria identificación del interesado en su artículo 9.1, que exige que “las Administraciones Públicas están obligadas a verificar la identidad de los interesados en el procedimiento administrativo, mediante la comprobación de su nombre y apellidos o denominación o razón social, según corresponda, que consten en el Documento Nacional de Identidad o documento identificativo equivalente”.

Esta obligación se desarrolla en el propio artículo 9, epígrafes 2 a 4, de la LPAC, y en los artículos 15, y 26 a 28 del RAFESP, que esencialmente autorizan el uso de certificados cualificados de firma electrónica y otros sistemas que las Administraciones públicas consideren válido en los términos y condiciones que se establezca, siempre que cuenten con un registro previo como usuario que permita garantizar su identidad (y que cumplan las restantes condiciones legalmente previstas, esencialmente de control previo por parte de un órgano estatal), normativa que, en general, presenta insuficiencias notables³ desde la perspectiva de la innovación tecnológica, especialmente en el caso del RAFESP, que desaprovecha la oportunidad de ampliar los mecanismos previstos en los artículo 9.2.c) y 10.2.c) de la LPAC.

² Sobre la identidad personal y su relación con la identidad humana, cfr. el análisis realizado por DE ASÍS ROIG, R. (2022): “Ética, tecnología y derechos”, en LLANO ALONSO, F. H. (Dir.) y GARRIDO MARTÍN, J. y VALDIVIA JIMÉNEZ, R. (Coords.): *Inteligencia artificial y filosofía del derecho*, Laborum ediciones, Murcia, pp. 91 ss. GONZÁLEZ FUSTER, G. y NADOLNA PEETERS, M. (2021): *Person identification, human rights and ethical principles. Rethinking biometrics in the era of artificial intelligence*, European Parliamentary Research Service, Bruselas.

³ Cfr. COTINO HUESO, L. (2021): “El nuevo reglamento de Administración electrónica, que no innova en tiempos de transformación digital”, *Revista Catalana de Dret Públic*, nº 63, pp. 118 ss., <https://doi.org/10.2436/rcdp.i63.2021>, y ALAMILLO DOMINGO, I. (2021): “La identificación y firma electrónica en el Real Decreto 203/2021: de estos polvos, estos lodos...”, *El Consultor de los Ayuntamientos*, nº 5.

Los nuevos escenarios de atención a distancia –y la correlativa “falta de atención” personal a los interesados– presentan limitaciones importantes dimanantes de la inadecuación de los medios de identificación más ampliamente empleados, conduciendo a carencias en la asistencia a los interesados, situación que puede corregirse mediante el uso de procedimientos de reconocimiento biométrico basados en inteligencia artificial, siempre que se realice con garantía jurídica⁴.

2.- PRECISIONES TERMINOLÓGICAS PREVIAS

Para abordar de forma adecuada nuestro objeto de interés resulta preciso realizar una serie de precisiones terminológicas previas, tanto desde un punto de vista técnico (en cuanto sustrato fáctico) como jurídico (en cuanto institucionalización a efectos del establecimiento del correspondiente régimen jurídico).

2.1.- *Precisiones desde la perspectiva tecnológica*

Desde un punto de vista técnico, la normalización internacional ha producido diversos trabajos relevantes, de los que quizá el más pertinente sea la importante norma ISO/IEC 2382-37, relativa al vocabulario de tecnologías de la información: biometría, actualmente en su edición tercera de 2022, preparada por el Comité Técnico Conjunto 1 de ISO e IEC, subcomité 37, biometría. Esta norma es especialmente útil porque establece una descripción sistemática de los conceptos en el campo de la biometría aplicada al reconocimiento de seres humanos, pero también porque reconcilia términos que venían teniendo usos diversos en normas internacionales.

Por su interés, presentamos a continuación una selección apropiada de entre los 198 términos contenidos en dicha norma, con su correspondiente análisis.

⁴ Cfr. con carácter general, GOMEZ ABEJA, L. (2022): “Inteligencia artificial y derechos fundamentales”, en LLANO ALONSO, F. H. (Dir.) y GARRIDO MARTÍN, J. y VALDIVIA JIMÉNEZ, R. (Coords.): *Inteligencia artificial y filosofía del derecho*, Laborum ediciones, Murcia, pp. 25 ss.

Lo biométrico (37.01.01 *biometric*), que sólo puede usarse como adjetivo, es lo relativo o perteneciente a la biometría, definición coincidente con la del diccionario de la RAE.

La biometría (37.01.03 *biometrics*), sustantivo sinónimo al reconocimiento biométrico (*biometric recognition*), es el reconocimiento automático de individuos basado en sus características biológicas y conductuales. Esta definición de biometría contrasta con la contenida en el diccionario de la RAE, que la caracteriza como el estudio mensurativo o estadístico de los fenómenos o procesos biológicos, en línea con el término inglés *biometry*.

El reconocimiento biométrico incluye la verificación y la identificación biométricas. De especial interés resulta que esta norma internacional considere obsoleto y desaconseje el uso del término “autenticación” como sinónimo de verificación o de identidad biométricas. El término preferible es el de reconocimiento biométrico al que nos estamos refiriendo. La misma norma descalifica el uso del término biometría como sustantivo, considerando incorrecta una expresión como “el rostro es la biometría adecuada para...”, por ejemplo.

La verificación biométrica (37.08.03 *biometric verification*) es el proceso de confirmar una declaración biométrica mediante comparación, definiéndose la declaración biométrica (37.06.04 *biometric claim*) como aquella referida a que un sujeto de captura biométrica es o no la fuente corporal de una referencia biométrica especificada o no especificada, y donde el sujeto de captura biométrica (37.07.03 *biometric capture subject*) es aquel individuo cuyos datos biométricos individualizados se encuentran en el sistema biométrico.

Los datos biométricos (37.03.06 *biometric data*) consisten en una muestra biométrica o una agregación de muestras biométricas en cualquier etapa de procesamiento, siendo una muestra biométrica (37.03.21 *biometric sample*) una representación analógica o digital de características biométricas (37.01.02 *biometric characteristics*) previa a la extracción de rasgos distintivos biométricos (37.03.11 *biometric feature*).

Por su parte, la comparación (37.05.07 *comparison*) es la estimación, cálculo o medida de similitud o disimilitud entre un (o varios) sondeo/s biométrico/s y una (o varias) referencia/s biométrica/s. Una sonda biométrica (37.03.14 *biometric probe(s)*) es una muestra bio-

métrica o un rasgo biométrico a utilizar para la comparación, mientras que una referencia biométrica (37.03.16 *biometric reference*) es una (o varias) muestra biométrica, plantilla biométrica o modelo biométrico atribuido a un sujeto, que se utiliza en procesos de comparación.

Por tanto, podemos hablar de verificación biométrica, como modalidad del reconocimiento biométrico, cuando se ha verificado la corrección de una declaración biométrica concreta mediante el contraste entre una o varias sondas biométricas frente a una o varias referencias biométricas del mismo sujeto.

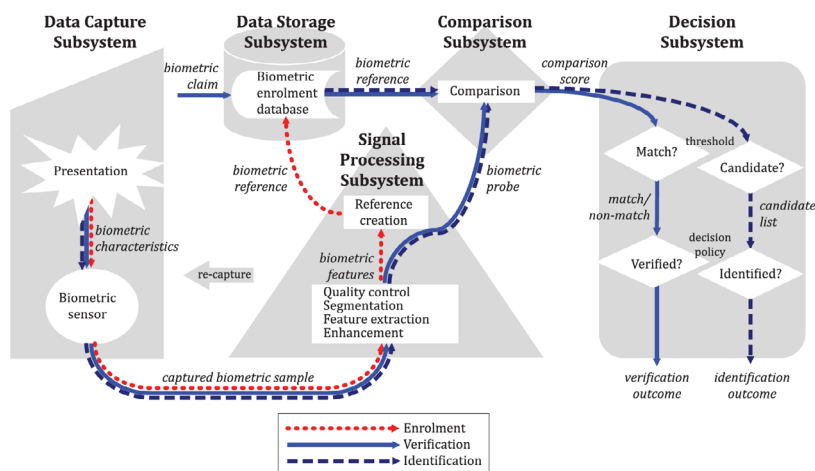
La comparación a que nos estamos refiriendo puede ser uno-a-uno (37.05.10 *one-to-one comparison*) o uno-a-muchos (37.05.11 *one-to-many comparison*). En el primer caso, la comparación se realiza entre una sonda biométrica y un individuo, mientras que, en el segundo caso, la comparación se realiza frente a diversos sujetos. Adicionalmente, la búsqueda biométrica uno-a-muchos permite devolver una lista de sujetos candidatos, o de no inscripción (por ejemplo, cuando se debe verificar que un sujeto no consta en la base de datos en cuestión).

La identificación biométrica (37.08.02 *biometric identification*) es el proceso de realizar una búsqueda en una base de datos de inscripción biométrica (37.03.09 *biometric enrolment database*) para encontrar y devolver un (o varios) identificador de referencia biométrica (37.03.19 *biometric reference identifier(s)*) atribuible a un único individuo.

Como se puede ver, el proceso de identificación exige de la previa inscripción de datos de un individuo en una base de datos específica. Dicha inscripción biométrica (37.05.03 *biometric enrolment*) es el acto de crear y almacenar un registro de datos de inscripción biométrica (37.03.10 *biometric enrolment data record*) atribuido al sujeto, el cual contiene no biométricos y se encuentra asociado a los identificadores de referencias biométricas, que apuntan a registros de datos de referencia biométrica (37.03.18 *biometric reference data record*) en la base de datos de referencias biométricas (37.03.17 *biometric reference database*).

La ilustración 1, extraída del documento ISO/IEC TR 24741:2018, muestra de forma gráfica los tres procesos biométricos que nos interesan: la inscripción, la verificación y la identificación:

**Ilustración 1. Componentes de un sistema biométrico genérico
(ISO/IEC TR 24741:2018)**



Como se puede ver, esencialmente la inscripción biométrica consiste en (1) la captura de muestras biométricas (correspondientes a determinadas características biométricas) mediante sensores biométricos, (2) la extracción de rasgos distintivos biométricos, (3) la creación de referencias biométricas y (4) el almacenamiento de estas informaciones en una base de datos.

Por su parte, la verificación consiste en (1) la recepción de una declaración biométrica (por ejemplo, “soy el interesado con nombre X, apellidos Y y número de DNI Z”), (2) la obtención de una sonda biométrica del supuesto sujeto (la imagen facial), (3) la recuperación de una o varias referencias biométricas de dicho supuesto sujeto (por ejemplo, un patrón de la imagen facial, asociado al interesado con nombre X, apellidos Y y número de DNI Z), (4) la comparación de los anteriores y (5) la adopción de la decisión acerca de la verificación (por ejemplo, “la persona es el interesado con nombre X, apellidos Y y número de DNI Z”).

El proceso de identificación biométrica consiste en (1) la recepción de una declaración biométrica, (2) la obtención de una sonda biométrica del supuesto sujeto, (3) la recuperación de una o varias referencias biométricas varios o de todos los sujetos inscritos en la base de

datos, (4) la comparación de los anteriores, (5) la devolución de una lista de sujetos candidatos, en función de un umbral, y (6) la adopción de la decisión acerca de la identificación.

Hay que decir que, en algunos procesos de verificación biométrica no existe una base de datos de inscripción biométrica, sino que se basan en la comparación de dos muestras biométricas entre ellas, obtenidas en el mismo instante temporal. Por ejemplo, en el caso de la Orden ETD/465/2021, de 6 de mayo, por la que se regulan los métodos de identificación remota por vídeo para la expedición de certificados electrónicos cualificados, se prevé la obtención de una muestra biométrica facial del solicitante, que se compara con la fotografía del documento oficial de identidad.

Por otra parte, la base de datos biométrica no tiene porqué consistir, en todo caso, en un sistema centralizado gestionado por un tercero. Por ejemplo, los documentos oficiales de identidad conforme a la normativa ICAO 9303, como los pasaportes, almacenan la base de datos de referencia biométrica del sujeto dentro de la propia tarjeta soporte. En otros casos, la base de datos biométrica se encuentra distribuida.

Es imperativo, finalmente, tener en cuenta que, como recuerda ISO/IEC TR 24741, las tecnologías biométricas no permiten “identificar” directamente a personas, sino que permiten vincular cuerpos a registros de atributos, que serían las “identidades”, por lo que frecuentemente ocurrirá que el reconocimiento biométricos se convierte en parte de un sistema de gestión de identidad, resultando importante asegurar una correcta asignación de los citados atributos de identidad en el momento de la inscripción del sujeto.

En este sentido, entre los desarrollos recientes del uso de la biometría se encuentra el agrupamiento (*clustering*) de muestras biométricas, como imágenes faciales por parte de redes sociales, a los efectos de atribuirles a un único sujeto sin necesidad de conocer su identidad, lo que elimina la necesidad de inscripción.

2.2.- Precisiones desde la perspectiva jurídica

Frente a esta precisa terminología técnica, el RGPD resulta ciertamente parco, si bien contiene una definición de datos biométricos,

como aquellos “datos personales obtenidos a partir de un tratamiento técnico específico, relativos a las características físicas, fisiológicas o conductuales de una persona física que permitan o confirmen la identificación única de dicha persona, como imágenes faciales o datos dactiloscópicos”.

Se trata de una definición que, sin embargo, que viene a incluir perfectamente todos los conceptos técnicos a que nos acabamos de referir. De hecho, la definición legal viene a poner foco precisamente en los datos que se obtienen mediante un tratamiento específico, lo que coincidiría con el concepto técnico de “datos biométricos”, en el sentido de muestra biométrica, de rasgo distintivo biométrico, de sonda biométrica, de referencia biométrica, etc. que son precisamente resultado de diversos tratamientos técnicos.

En interpretación literal, quedaría fuera de esta noción las propias características físicas, fisiológicas o conductuales que no sean tratadas de forma técnica para permitir o confirmar esta identificación única, en línea con lo establecido en el considerando 51 del RGPD, en cuya virtud “el tratamiento de fotografías no debe considerarse sistemáticamente tratamiento de categorías especiales de datos personales, pues únicamente se encuentran comprendidas en la definición de datos biométricos cuando el hecho de ser tratadas con medios técnicos específicos permita la identificación o la autenticación unívocas de una persona física”. De forma similar, parece que la simple captura de otras características personales sin tratamiento ulterior, como por ejemplo la captura electrónica del trazo, también quedaría fuera de dicha definición.

Hasta la publicación del RGPD han sido muy escasas las referencias al tratamiento de datos biométricos en las normativas europeas e internacionales que han versado sobre el derecho a la protección de datos. Los primeros trabajos se realizaron por Grupo de Trabajo del Artículo 29 en su Documento de trabajo sobre biometría del año 2003, y en el posterior Dictamen 4/2007 sobre el concepto de datos personales, en el que ofreció una primera definición relativa a esta tipología de datos, calificando los mismos como “propiedades biológicas, características fisiológicas, rasgos de la personalidad o tics, que son, al mismo tiempo, atribuibles a una sola persona y mensurables,

incluso si los modelos utilizados en la práctica para medirlos técnicamente implican un cierto grado de probabilidad”.

Posteriormente, el Dictamen 3/2012 sobre la evolución de las tecnologías biométricas mantuvo esta concepción de los datos biométricos, con la matización de que, si bien los datos biométricos fueron clasificados como dato personal, no lo fueron como dato sensible (sin perjuicio de que pudieran inferirse de los mismos otros datos que sí lo fueran, como aquellos relativos a la salud o a la revelación del origen étnico o racial).

El tratamiento de los datos biométricos se ha visto afectado profundamente por nuevo Reglamento, que los equipara plenamente a otros datos tradicionalmente calificados como sensibles, como los datos genéticos o de salud.

La recogida de elementos biométricos, como manifestó en su día el Grupo de trabajo del art. 29 de la Directiva, en definitiva, supone “recoger datos del cuerpo de una persona” según el Dictamen 3/2005 del Grupo de Trabajo del art. 29, sobre la aplicación del Reglamento (CE) n° 2252/2004 del Consejo, de 13 de diciembre de 2004, sobre normas para las medidas de seguridad y datos biométricos en los pasaportes y documentos de viaje expedidos por los Estados miembros (DO L 385 de 29 de diciembre 2004, págs. 1-6), y comporta graves riesgos éticos, pero por otro lado, también es cierto que se utilizan con éxito y eficacia en la investigación científica, son un elemento clave de la ciencia forense y un valioso elemento de los sistemas de control de acceso. Pueden contribuir a aumentar el nivel de seguridad y a facilitar, acelerar y simplificar los procedimientos de identificación y autenticación (según el Dictamen 3/2012 del Grupo de Trabajo del art. 29, sobre la evolución de las tecnologías biométricas, adoptado el 27 de abril de 2012).

Inicialmente, su carácter de dato sensible venía determinado por el juicio de finalidad de su tratamiento; sin embargo, ante la extensión de estos tratamientos en materia de seguridad y la dificultad de garantizar que la información así conseguida no se use para otras finalidades, se ha incidido en que se haya reforzado su carácter de dato sensible en el RGPD.

Por ejemplo, en el caso de la firma manuscrita capturada electrónicamente, es evidente que el acceso a estos datos por terceros no

autorizados puede generar un riesgo importante al interesado, dada la posibilidad que dicho tercero no autorizado tendría de producir documentos falsos aparentemente atribuibles al interesado.

La AEPD en su informe 36/2020 en respuesta a una serie de cuestiones relativas al uso de técnicas de reconocimiento facial en la realización de pruebas de evaluación online planteadas por la Conferencia de Rectores de las Universidades Españolas (CRUE), indica lo siguiente en la consideración III:

“No obstante, hay que adelantar que el RGPD no parece considerar a todo tratamiento de datos biométricos como tratamiento de categorías especiales de datos, ya que el artículo 9.1. se refiere a los “datos biométricos dirigidos a identificar de manera unívoca a una persona física”, por lo que, de una interpretación conjunta de ambos preceptos parece dar a entender que los datos biométricos solo constituirían una categoría especial de datos en el caso de que se sometan a un tratamiento técnico específico dirigido a identificar de manera unívoca a una persona física. En este sentido, parece que igualmente se pronuncia el Considerando 51 al señalar que «El tratamiento de fotografías no debe considerarse sistemáticamente tratamiento de categorías especiales de datos personales, pues únicamente se encuentran comprendidas en la definición de datos biométricos cuando el hecho de ser tratadas con medios técnicos específicos permita la identificación o la autenticación unívocas de una persona física»”.

Asimismo, en apoyo de dicha tesis, la AEPD cita el Protocolo de enmienda al Convenio para la Protección de Individuos con respecto al procesamiento de datos personales, aprobada por el Comité de Ministros en su 128º período de sesiones en Elsinore el 18 de mayo de 2018 (Convenio 108+), el cual incluye únicamente como categorías especiales de datos, en su artículo 6.1 a los datos biométricos dirigidos a la identificación unívoca de una persona (“biometric data uniquely identifying a person”), sin incluir la referencia a la autenticación.

Al objeto de aclarar las dudas interpretativas que surgen respecto a la consideración de los datos biométricos como categorías especiales de datos, la AEPD acude igualmente a la ya citada distinción entre identificación biométrica y verificación/autenticación biométrica que estableció el Grupo del Artículo 29 en su Dictamen 3/2012 sobre la evolución de las tecnologías biométricas:

“1) Identificación biométrica: la identificación de un individuo por un sistema biométrico es normalmente el proceso de comparar sus datos

biométricos (adquiridos en el momento de la identificación) con una serie de plantillas biométricas almacenadas en una base de datos (es decir, un proceso de búsqueda de correspondencias uno-a-varios).

2) Verificación/autenticación biométrica: la verificación de un individuo por un sistema biométrico es normalmente el proceso de comparación entre sus datos biométricos (adquiridos en el momento de la verificación) con una única plantilla biométrica almacenada en un dispositivo (es decir, un proceso de búsqueda de correspondencias uno a- uno)".

Esta misma diferenciación se recoge en el Libro blanco sobre la inteligencia artificial de la Comisión Europea, el cual establece lo siguiente que:

"En lo que se refiere al reconocimiento facial, por «identificación» se entiende que la plantilla de la imagen facial de una persona se compara con otras muchas plantillas almacenadas en una base de datos para averiguar si su imagen está almacenada en ella. La «autenticación» (o «verificación»), por su parte, se refiere habitualmente a la búsqueda de correspondencias entre dos plantillas concretas. Permite la comparación de dos plantillas biométricas que, en principio, se supone que pertenecen a la misma persona; así, las dos plantillas se comparan para determinar si la persona de las dos imágenes es la misma. Este procedimiento se emplea, por ejemplo, en las puertas de control automatizado de fronteras empleadas en los controles fronterizos de los aeropuertos", añadiendo que "la identificación biométrica remota debe distinguirse de la autenticación biométrica (esta última es un procedimiento de seguridad que se basa en las características biológicas exclusivas de una persona para comprobar que es quien dice ser). La identificación biométrica remota consiste en determinar la identidad de varias personas con la ayuda de identificadores biométricos (huellas dactilares, imágenes faciales, iris, patrones vasculares, etc.) a distancia, en un espacio público y de manera continuada o sostenida contrastándolos con datos almacenados en una base de datos".

Finalmente, la AEPD concluye que, atendiendo a la citada distinción, puede interpretarse que, de acuerdo con el artículo 4 del RGPD, el concepto de dato biométrico incluiría ambos supuestos, tanto la identificación como la verificación/autenticación. Sin embargo, y con carácter general, los datos biométricos únicamente tendrán la consideración de categoría especial de datos en los supuestos en que se sometan a tratamiento técnico dirigido a la identificación biométrica (uno-a-varios) y no en el caso de verificación/autenticación biométrica (uno-a-uno).

En definitiva, para la AEPD no todo tratamiento de datos biométricos constituye un tratamiento de categorías especiales de datos, si bien considera que se trata esta de una cuestión compleja, sometida a interpretación, respecto de la cual no se pueden extraer conclusiones generales, debiendo atenderse al caso concreto según los datos tratados, las técnicas empleadas para su tratamiento y la consiguiente inferencia en el derecho a la protección de datos, debiendo, en tanto en cuanto no se pronuncie al respecto el EDPB o los órganos jurisdiccionales, adoptarse, en caso de duda, la interpretación más favorable para la protección de los derechos de los afectados. En este sentido, cabe citar que el dictamen CNS 59/2021 de la Autoritat Catalana de Protecció de Dades considera que identificación y verificación biométrica implican en ambos casos la inclusión en el ámbito del artículo 9.1 del RGPD, conclusión que no podemos compartir.

Más recientemente, la propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión, COM/2021/206 final, ha aportado una definición legal de identificación biométrica remota, a la que caracteriza como:

“un sistema de IA destinado a identificar a personas físicas a distancia comparando sus datos biométricos con los que figuran en una base de datos de referencia, y sin que el usuario del sistema de IA sepa de antemano si la persona en cuestión se encontrará en dicha base de datos y podrá ser identificada”.

La propuesta normativa distingue entre los sistemas “en tiempo real” y “en diferido”, a efectos del establecimiento de un tratamiento diferenciado, dado los riesgos también diferentes que se plantean en cada caso.

Asimismo, la propuesta aporta dos definiciones adicionales relevantes, como son el sistema de reconocimiento de emociones (un sistema de IA destinado a detectar o deducir las emociones o las intenciones de personas físicas a partir de sus datos biométricos) y el sistema de categorización biométrica (un sistema de IA destinado a asignar a personas físicas a categorías concretas, como un sexo, edad, color de pelo, color de ojos, tatuajes, origen étnico u orientación sexual o política, en función de sus datos biométricos), que no constituyen sistemas de reconocimiento biométrico.

Finalmente, a efectos puramente operativos, cabe citar la definición de sistema de inteligencia artificial contenida en la citada propuesta de Reglamento:

“El software que se desarrolla empleando una o varias de las técnicas y estrategias que figuran en el anexo I y que puede, para un conjunto determinado de objetivos definidos por seres humanos, generar información de salida como contenidos, predicciones, recomendaciones o decisiones que influyan en los entornos con los que interactúa”.

Por su parte, el anexo I se refiere a lo siguiente:

“Estrategias de aprendizaje automático, incluidos el aprendizaje supervisado, el no supervisado y el realizado por refuerzo, que emplean una amplia variedad de métodos, entre ellos el aprendizaje profundo.

Estrategias basadas en la lógica y el conocimiento, especialmente la representación del conocimiento, la programación (lógica) inductiva, las bases de conocimiento, los motores de inferencia y deducción, los sistemas expertos y de razonamiento (simbólico).

Estrategias estadísticas, estimación bayesiana, métodos de búsqueda y optimización”.

3.- ALGUNOS CASOS DE USO DEL RECONOCIMIENTO BIOMÉTRICO BASADO EN INTELIGENCIA ARTIFICIAL EN EL PROCEDIMIENTO ADMINISTRATIVO

En este epígrafe presentamos los principales casos de uso del reconocimiento biométrico basado en inteligencia artificial en el procedimiento administrativo. No forma parte de nuestro objeto de interés el análisis de otras aplicaciones en el sector público, significativamente el uso estas técnicas para la identificación remota en el espacio público para el cumplimiento de la ley, que la propuesta de Reglamento europeo prohíbe con carácter general.

El artículo 4 del RAFESP, relativo a los canales de asistencia para el acceso a los servicios públicos establece lo siguiente:

“Las Administraciones Públicas prestarán la asistencia necesaria para facilitar el acceso de las personas interesadas a los servicios electrónicos proporcionados en su ámbito competencial a través de alguno o algunos de los siguientes canales:

- a) Presencial, a través de las oficinas de asistencia que se determinen.
- b) Portales de internet y sedes electrónicas.
- c) Redes sociales.
- d) Telefónico.
- e) Correo electrónico.
- f) Cualquier otro canal que pueda establecerse de acuerdo con lo previsto en el artículo 12 de la Ley 39/2015, de 1 de octubre”.

Esta normativa sitúa en plano de igualdad los diferentes canales de acceso a los servicios públicos, si bien la sede electrónica aporta una serie de garantías específicas que permiten garantizar el cumplimiento de la normativa de protección de datos, entre ellas una suficiente identificación del interesado, que resulta compleja en casos como los que se presentan a continuación.

3.1.- Reconocimiento biométrico en la asistencia por canal telefónico

Una primera posibilidad se centra en la oficina de asistencia telefónica que debe atender a los interesados, en especial, personas sin acceso a medios tecnológicos avanzados, personas mayores o en riesgo de riesgo de exclusión.

A título de ejemplo, la Orden HAP/1949/2014, de 13 de octubre, por la que se regula el Punto de Acceso General de la Administración General del Estado y se crea su sede electrónica, prevé como canal de acceso a los servicios de la sede electrónica el telefónico, mediante el número de teléfono 060, limitado a los servicios de información departamental.

Se trata, como se puede ver, y confirma la Resolución de 30 de noviembre de 2011, de la Secretaría de Estado de Telecomunicaciones y para la Sociedad de la Información, por la que se modifica la atribución del número 060 al servicio de información de la Administración General del Estado, de un servicio de información de carácter administrativa, cuyo objetivo es la integración de todos sus servicios de atención telefónica. Sin embargo, a tenor de la citada Resolución, el citado número 060 puede utilizarse también para la atención de consultas sobre el estado de tramitación de expedientes.

La implementación de esta posibilidad exige la adopción de medios para la verificación de la identidad del interesado, existiendo diversas posibilidades en función de la sensibilidad de la información. Ciertamente, para que la Administración informe al interesado acerca del estado de su trámite, normalmente será suficiente con la adopción de las medidas previstas en el ENS para el nivel medio... que remite a la normativa europea de niveles de seguridad de los medios de identificación electrónica; esto es, el ReIDAS y el Reglamento de Ejecución (UE) 2015/1502 de la Comisión, de 8 de septiembre de 2015, sobre la fijación de especificaciones y procedimientos técnicos mínimos para los niveles de seguridad de medios de identificación electrónica con arreglo a lo dispuesto en el artículo 8, apartado 3, del Reglamento (UE) n° 910/2014 del Parlamento Europeo y del Consejo, relativo a la identificación electrónica y los servicios de confianza para las transacciones electrónicas en el mercado interior.

Ninguna de estas normas prevé el uso de sistemas de reconocimiento biométrico, en general, mucho menos de sistemas basados en la voz, si bien sí consideran dichas normas el uso de mecanismos biométricos como factor de autenticación reforzada. Y, por otra parte, los mecanismos que sí se prevén, como podría ser Cl@ve, no parecen fácilmente aplicables a este canal. Por ello, una aplicación estricta del marco jurídico expuesto conduciría a la imposibilidad de habilitar un canal efectivo de atención en el que, por ejemplo, se pueda dar acceso a datos de carácter personal, o a la búsqueda de alternativas.

Una alternativa sería, por supuesto, acudir a mecanismos de verificación de la identidad mediante “información conocida por ambas partes”, como por ejemplo los datos personales de identidad, o la aportación de información del expediente.

Pero en el caso que estamos planteando, pueden darse situaciones las que resulte necesario o incluso imprescindible ofrecer acceso pleno al procedimiento o a la información del expediente. Por ejemplo, en el caso de asistencia remota a personas con capacidades limitadas de movilidad, y que no dispongan de los dispositivos y medios de acceso a la sede electrónica. Sin duda en determinados casos resulta razonable permitir que el interesado pueda inscribirse en un sistema biométrico a efectos de la posterior verificación de su identidad, especialmente en la

medida en que la Administración despliegue sistemas automatizados de atención que complementen la asistencia personal.

3.2.- *Reconocimiento biométrico en la asistencia por videoconferencia*

Una segunda posibilidad, de aplicación más realista, consiste en una oficina de asistencia en materia de registros mediante videoconferencia, que implemente un proceso de verificación biométrica contra documento oficial.

Se trata de un caso de uso diferente del anterior, puesto que no implica la expedición de un nuevo medio de identificación electrónica, sino la utilización de un medio ya existente. En el caso del DNI electrónico, además, hay que recordar que, conforme al artículo 8.1 de la Ley Orgánica 4/2015, de 30 de marzo, de protección de la seguridad ciudadana, en redacción dada por Real Decreto-ley 14/2019, de 31 de octubre, por el que se adoptan medidas urgentes por razones de seguridad pública en materia de administración digital, contratación del sector público y telecomunicaciones, “es el único documento con suficiente valor por sí solo para la acreditación, a todos los efectos, de la identidad y los datos personales de su titular”.

A diferencia del anterior, nos encontraríamos ante un caso que encajaría con mayor facilidad en el artículo 9.2.c) de la LPAC, en la medida en que existe un registro de identidad creado por la autoridad competente para expedir un documento que, como recuerda la exposición de motivos del Real Decreto-ley 14/2019, es el único que tiene suficiente valor por sí solo para la acreditación, a todos los efectos, de la identidad y los datos personales de su titular.

Lo que se hace en este caso es utilizar técnicas de reconocimiento biométrico basados en inteligencia artificial para determinar que la persona que se conecta a la videoconferencia cuya fotografía consta en el documento que se exhibe, proceso que debe realizarse con elevadas garantías técnicas⁵.

⁵ En relación con las más recientes técnicas de ataque a alguno de estos sistemas, cfr. TSCHIRSICH, M. (2022): *Praktischer Angriff auf Video-Ident. Demonstra-*

Es cierto que, desde la perspectiva de la normativa de protección de datos⁶, parece desproporcionado aplicar este mecanismo por defecto, por lo que quizá deba reservarse a aquellas situaciones en que exista un riesgo significativo, o valorar su uso de forma absolutamente voluntaria por parte del usuario, como mecanismo alternativo.

3.3.- *La captura electrónica de firma manuscrita*

Un tercer caso lo encontramos en la captura electrónica de la firma manuscrita capturada en un dispositivo (como una tableta) en actuaciones presenciales. No solo resulta aplicable en la oficina de asistencia en materia de registros, sino también en inspección, por ejemplo en vía pública, donde no parece razonable exigir otros sistemas de firma, que están pensados para las relaciones a distancia.

En la captura de esta firma manuscrita se obtienen y almacenan las características biométricas del trazo, como presión o velocidad, que son de tipo conductual. Dado que no se comparan con ninguna otra información, resulta dudoso que nos encontremos ante un tratamiento biométrico de verificación ni, por supuesto, de identificación, pero dicho proceso ciertamente se producirá en caso de conflicto judicial.

De ello no se deduce que dicho tratamiento de datos resulte inoecu, dado que, al contrario, un mal diseño de la solución puede conllevar el acceso no autorizado a dichas informaciones, permitiendo a terceros no autorizados la producción de documentos falsos.

Las prácticas de seguridad recomendadas incluyen: (1) el empleo de tabletas digitalizadoras de calidad que capturen, de forma fiable, la máxima cantidad de información de la dinámica de la firma; (2) que la firma manuscrita capturada electrónicamente se encuentre vinculada de forma unívoca con un contenido específico, y no deba poderse vincular a otro contenido, por ejemplo mediante el uso de resúmenes criptográficos; (3) que no se puedan reutilizar los datos de firma capturados (presión, velocidad, inclinación, etc.), lo cual exige que

tion inhärenter Schwächen der videobasierten Echtheitsprüfung physischer ID-Dokumente, Chaos Computer Club.

⁶ Sobre los desafíos del reconocimiento facial, cfr. SANTISTEBAN GALARZA, M. (2021): “Reconocimiento facial y protección de datos: una respuesta provisional a un problema pendiente”, *Revista de Derecho UNED*, nº 28, pp. 499 ss.

las partes no tengan acceso a los mismos (por ejemplo, mediante su cifrado, o su captura por tercero de confianza), lo que hace imprescindible definir e implantar una colección de normas y procedimientos de seguridad con respecto a las claves de cifrado/descifrado de los datos biométricos de la firma manuscrita capturada electrónicamente; (4) garantizar que la clave tiene calidad y ha sido generada y probada correctamente, que nadie ha tenido acceso a la clave de descifrado, y que se han establecido procesos de copia de seguridad para evitar la pérdida de la clave de descifrado (o perdemos el 100% de las pruebas), previéndose la rotación de claves, atendiendo a la seguridad de los algoritmos; (5) que se establezca un proceso de descifrado de los datos de firma manuscrita capturada electrónicamente, con tiempos de respuesta alineados con los plazos de los diferentes procedimientos judiciales, administrativos o de mediación, proceso que debe realizar un tercero al que no se le pueda exigir la entrega de la clave de descifrado; (6) proteger el archivo con el contenido firmado y la firma manuscrita capturada electrónicamente frente a alteraciones, por ejemplo mediante un sello cualificado de tiempo electrónico, que goza de presunción de autenticidad, o un sello electrónico cualificado de empresa, que tiene valor contractual en el caso de condiciones generales de la contratación y facilita la conservación futura; (7) y, finalmente, el establecimiento de guías para la aportación de la prueba al procedimiento, resultando conveniente disponer de informe jurídico y de informe pericial informático sobre el sistema, además de soporte de perito calígrafo en caso necesario.

4.- PROPUESTAS PARA EL USO DEL RECONOCIMIENTO BIOMÉTRICO BASADO EN INTELIGENCIA ARTIFICIAL EN EL PROCEDIMIENTO ADMINISTRATIVO

Los casos que hemos presentado en el epígrafe anterior tienen en común poner de manifiesto las insuficiencias de los medios de identificación electrónica actualmente implementados en determinados tipos de interacciones entre el interesado y la Administración o, dicho de otra forma, que los medios de identificación disponibles se encuentran orientados a la relación a distancia a través de la interfaz web de la sede electrónica, por lo que los mismos no son utilizables en oficinas

de asistencia o de inspección, entre otros escenarios de interacción personal.

Lograr el uso del reconocimiento biométrico basado en inteligencia artificial en el procedimiento administrativo requiere abordar una serie de desafíos jurídicos relevantes, si bien no puede obviarse la importancia de la perspectiva ética en el diseño del uso de estas tecnologías⁷.

En primer lugar, determinados tratamientos ya presentados, como la captura de los datos de la firma manuscrita sin verificación alguna difícilmente parecen encajar en el artículo 9.1 del RGPD, dado que sin datos de contraste no resulta posible la singularización del individuo. Y a pesar de ello, alguna autoridad de control de datos de carácter personal parece considerar que esta misma captura implicaría ya la sujeción a dicho artículo. Quizá la aplicación del modelo de aplicación por fases descrito en la STJUE en el caso *Fashion ID* contribuya a la resolución de esta problemática, en el sentido de que en la fase de captura no resulte aplicable el artículo 9.1 del RGPD, y de que en la fase de prueba judicial, en que se deberá obtener una colección de muestras biométricas (otras firmas manuscritas capturadas electrónicamente) para la práctica de la prueba pericial, sí resulte aplicable dicho artículo y, lo que es más importante, el artículo 9.2.f) del propio RGPD, a efectos de la necesaria base legal para este tratamiento (“el tratamiento es necesario para la formulación, el ejercicio o la defensa de reclamaciones o cuando los tribunales actúen en ejercicio de su función judicial”).

En los restantes casos, también resulta preciso distinguir la identificación biométrica y la verificación biométrica, si bien cabrá considerar que ambos casos quedan eventualmente sujetos al artículo 9.1 del RGPD.

En segundo lugar, resulta preciso delimitar qué casos de uso de reconocimiento biométrico resultan defendibles desde la perspectiva del

⁷ Cfr. KISSELBURGH, L., BEEVER, J. (2022): “The Ethics of Privacy in Research and Design: Principles, Practices, and Potential”, en KNIJNENBURG, B.P., PAGE, X., WISNIEWSKI, P., LIPFORD, H.R., PROFERES, N., ROMANO, J. (Coord.): *Modern Socio-Technical Perspectives on Privacy*. Springer, Cham. https://doi.org/10.1007/978-3-030-82786-1_17.

necesario análisis de proporcionalidad, especialmente cuando los mismos se encuentren sujetos al artículo 9.1 del RGPD (“datos biométricos dirigidos a identificar de manera unívoca a una persona física”), dado el régimen de prohibición general de esta posibilidad. Un ejemplo de escenario donde puede resultar proporcionado habilitar una verificación biométrica sería aquel en que una persona no dispone de ningún medio de identificación, o cuando los mismos no se encuentran disponibles, especialmente cuando en estos casos el interesado pueda sufrir un perjuicio significativo (debido, por ejemplo, a la aplicación de los plazos administrativos).

En tercer lugar, cuando el caso de uso quede sujeto al artículo 9.1 del RGPD, sólo podrá realizarse el tratamiento en alguna de las circunstancias previstas en el epígrafe 2 del citado artículo, lo que, en el ámbito del procedimiento administrativo, apunta singularmente al consentimiento explícito. En este sentido, se deben considerar siempre todas las alternativas al uso de la biometría, significativamente admitir los medios electrónicos de identificación (por ejemplo, en la asistencia a través de videoconferencia).

Resulta deseable, de *lege ferenda*, establecer habilitaciones legales para estos tratamientos, desde la óptica de protección de datos. En el caso de la biometría para la identificación unívoca se requiere normativa con rango de ley, y casi con toda seguridad también para la simple verificación biométrica. La propuesta sería claramente una reforma del artículo 9 de la LPAC para establecer el correspondiente régimen jurídico.

En el caso de videoconferencia, existe el interesante antecedente del artículo 7.2 de la Ley 6/2020, de 11 de noviembre, reguladora de determinados aspectos de los servicios electrónicos de confianza, desarrollado por Orden ETD/465/2021, para expedir certificados cualificados de firma o sello electrónico. Dicha normativa reglamentaria puede constituir un buen ejemplo de qué habría que regular, si bien trasladando a la norma legal el conjunto de medidas y garantías esenciales.

En otro orden de cosas, parece que los sistemas de inteligencia artificial para la identificación deben estar sujetos a las exigencias regulatorias de la actuación administrativa automatizada, incluida la publicidad acerca de los datos para el entrenamiento, novedad importante

del artículo 11.1.i) del RAFESP, conforme al cual la sede electrónica debe contener “la relación actualizada de las actuaciones administrativas automatizadas vinculadas a los servicios, procedimientos y trámites”, con “la descripción de su diseño y funcionamiento, los mecanismos de rendición de cuentas y transparencia, así como los datos utilizados en su configuración y aprendizaje”.

En cuarto lugar, resulta imprescindible implementar procesos fiables para el uso de la tecnología biométrica, en línea con las normas técnicas como ISO/IEC 30107. Resulta de extraordinaria importancia establecer sistemas de certificación y control de los sistemas (ej. NIST o CCN-STIC 140 anexo F.11 para Orden ETD/465/2021), con una orientación a normas de seguridad de producto, sin perjuicio de regular su uso en escenarios concretos. En este sentido, la información de biometría debe ser adecuadamente protegida, en línea con propuestas como ISO/IEC 24745, por ejemplo, empleando referencias biométricas revocables.

Como idea final, y siempre dentro del mayor respeto a los derechos y libertades de los interesados, en un mundo que cada vez plantea mayores desafíos de ciberseguridad, quizá debería analizarse la posibilidad de constituir y ofrecer servicios públicos de verificación de identidad con base en procedimientos biométricos⁸.

La gestión de datos biométricos es inherente a la gestión del ciclo de vida de los documentos oficiales de identidad, existiendo procedimientos de enrolamiento y captura de gran calidad.

Por ello, la regulación legal de un servicio que habilitara el uso de la información biométrica contenida dentro de dichos documentos para procesos de verificación biométrica, posibilidad que técnicamente se encuentra normalizada desde hace años, permitiría incorporar las necesarias garantías y mejorar de forma verdaderamente relevante la asistencia remota, especialmente por videoconferencia.

Dado que, incluso tras la finalización de la pandemia, nos encontramos frente a una extraordinaria e injustificada restricción de la

⁸ Sobre el uso de información biométrica, con base en el sistema público de identidad digital italiano, en el ámbito de la fe pública, cfr. LEVANTE, G. (2020): *Il sistema pubblico d'identità digitale*, Studio n. 1-2020/DI, Consiglio Nazionale del Notariato.

atención presencial de los interesados no obligados a relacionarse electrónicamente con la Administración, resulta imperativo habilitar los mecanismos que garanticen la asistencia a distancia incluso cuando no dispongan o puedan utilizar otros medios.

BIBLIOGRAFÍA CITADA

- ALAMILLO DOMINGO, I. (2021): “La identificación y firma electrónica en el Real Decreto 203/2021: de estos polvos, estos lodos...”, *El Consultor de los Ayuntamientos*, nº 5.
- COMISIÓN EUROPEA (2020): Libro blanco sobre la inteligencia artificial – un enfoque europeo orientado a la excelencia y la confianza, COM(2020) 65 final.
- COTINO HUESO, L. (2021): “El nuevo reglamento de Administración electrónica, que no innova en tiempos de transformación digital”, *Revista Catalana de Dret Públic*, nº 63, pp. 118 ss., <https://doi.org/10.2436/rcdp.i63.2021>.
- DE ASÍS ROIG, R. (2022): “Ética, tecnología y derechos”, en LLANO ALONSO, F. H. (Dir.) y GARRIDO MARTÍN, J. y VALDIVIA JIMÉNEZ, R. (Coords.): *Inteligencia artificial y filosofía del derecho*, Laborum ediciones, Murcia, pp. 91 ss.
- GONZÁLEZ FUSTER, G. y NADOLNA PEETERS, M. (2021): *Person identification, human rights and ethical principles. Rethinking biometrics in the era of artificial intelligence*, European Parliamentary Research Service, Bruselas
- GOMEZ ABEJA, L. (2022): “Inteligencia artificial y derechos fundamentales”, en LLANO ALONSO, F. H. (Dir.) y GARRIDO MARTÍN, J. y VALDIVIA JIMÉNEZ, R. (Coords.): *Inteligencia artificial y filosofía del derecho*, Laborum ediciones, Murcia, pp. 25 ss.
- ISO/IEC 2382-37 (2022): *Information technology — Vocabulary — Part 37: Biometrics*, ISO, Vernier, Geneva.
- ISO/IEC TR 24741 (2018): *Information technology — Biometrics — Overview and application*, ISO, Vernier, Geneva.
- ISO/IEC 30107-1 (2016): *Information technology — Biometric presentation attack detection — Part 1: Framework*, ISO, Vernier, Geneva.
- KISSELBURGH, L., BEEVER, J. (2022): “The Ethics of Privacy in Research and Design: Principles, Practices, and Potential”, en KNIJNENBURG, B.P., PAGE, X., WISNIEWSKI, P., LIPFORD, H.R., PROFERES, N., ROMANO, J. (Coord.): *Modern Socio-Technical Perspectives on Privacy*. Springer, Cham. https://doi.org/10.1007/978-3-030-82786-1_17

- LEVANTE, G. (2020): *Il sistema pubblico d'identità digitale*, Studio n. 1-2020/DI, Consiglio Nazionale del Notariato.
- SANTISTEBAN GALARZA, M. (2021): “Reconocimiento facial y protección de datos: una respuesta provisional a un problema pendiente”, *Revista de Derecho UNED*, n° 28, pp. 499 ss.
- TSCHIRSICH, M. (2022): *Praktischer Angriff auf Video-Ident. Demonstration inhärenter Schwächen der videobasierten Echtheitsprüfung physischer ID-Dokumente*, Chaos Computer Club.

Apuesta por Tirant Online, la base de datos jurídica de la editorial más prestigiosa de España.*



www.tirantonline.com

Suscríbete a nuestro servicio de base de datos jurídica y tendrás acceso a todos los documentos de Legislación, Doctrina, Jurisprudencia, Formularios, Esquemas, Consultas o Voces, y a muchas herramientas útiles para el jurista:

- * Biblioteca Virtual
- * Herramientas Salariales
- * Calculadoras de tasas y pensiones
- * Tirant TV
- * Personalización
- * Foros y Consultoría
- * Revistas Jurídicas
- * Gestión de despachos
- * Biblioteca GPS
- * Ayudas y subvenciones
- * Novedades

* Según ranking del CSIC

 96 369 17 28

 96 369 41 51

 atencionalcliente@tirantonline.com

 www.tirantonline.com