

# INTELIGENCIA ARTIFICIAL, DECISIONES ADMINISTRATIVAS DISCRECIONALES TOTALMENTE AUTOMATIZADAS Y ALCANCE DEL CONTROL JUDICIAL: ¿INDIFERENCIA, INSUFICIENCIA O DEFERENCIA?\*

## ARTIFICIAL INTELLIGENCE, FULLY AUTOMATED DISCRETIONARY ADMINISTRATIVE DECISIONS AND THE SCOPE OF JUDICIAL REVIEW:

---

\* Fecha de finalización del trabajo: 30 de diciembre de 2023 (actualizado a 19 de marzo de 2024). 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 el blog del mencionado proyecto de investigación: [Inteligencia artificial en el sector público: retos, límites y medios. – Nudging aplicado a la Mejora de la Regulación y al Uso de Algoritmos y de Inteligencia Artificial \(wordpress.com\)](#) y en trabajos previos mencionados en la bibliografía. Esas reflexiones iniciales han sido reconsideradas, modificadas y ampliadas específicamente para esta publicación.

Los puntos principales de este análisis fueron debatidos en el marco del cuarto Seminario del ciclo Del Derecho y del Revés Tecnológico, asociado al proyecto de investigación mencionado más arriba, que versó sobre el control judicial de las decisiones administrativas adoptadas con uso de Inteligencia Artificial, realizado en la Facultad de Derecho de la Universidad Pompeu Fabra el día 19 de abril de 2023. Asimismo, una versión de este estudio se presentó en el Seminario de Teoría y Método del Derecho Público, en su XIX sesión, celebrado en la Universidad de Valencia a finales de mayo de 2023. Quiero agradecer a los y las colegas académicos que participaron en ambas actividades sus observaciones y sugerencias. Naturalmente, las opiniones y errores de este estudio sólo son atribuibles a su autor.

Abreviaturas utilizadas: AP= Administración Pública, CE=Constitución Española, EBEP=Estatuto Básico del Empleado Público, LJCA= Ley de la Jurisdicción Contencioso-Administrativa de 1998, IA=inteligencia artificial, LPAC= ley 39/2015, de 1 de octubre, del procedimiento administrativo común, LRJSP=Ley 40/2015, de 1 de octubre, del Régimen Jurídico del Sector Público, RUEIA= Resolución legislativa del Parlamento Europeo, de 13 de marzo de 2024, sobre 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)0206 – C9-0146/2021 – 2021/0106(COD))

## INDIFFERENCE, INSUFFICIENCY OR DEFERENCE?

Juli Ponce Solé

Catedrático de Derecho Administrativo  
Universidad de Barcelona

**Resumen:** El objeto de este artículo se centra en las posibilidades y límites de la adopción totalmente automatizada de decisiones administrativas discrecionales, el papel de la supervisión humana y el alcance e intensidad del control judicial de dichas decisiones. El estudio se detiene brevemente en la explicación de algunos conceptos esenciales vinculados a la IA para clarificar su uso en este análisis. Entre ellos, se realiza un breve análisis del concepto tradicional de discrecionalidad administrativa en España, poniendo de relieve su falta de profundidad y la necesidad de repensarlo a la luz de la buena administración. A partir de aquí, se abordan cuáles son los límites jurídicos de la IA y cómo deberían conducir a evitar su uso para la automatización total del ejercicio de potestades discrecionales, en el estado actual de la técnica. Esa limitación conduce a lo que se denomina una reserva de humanidad, que se desarrolla a continuación. Finaliza el análisis exponiendo la adaptación de las técnicas de control judicial al ejercicio de potestades administrativas con IA, los posibles problemas de indiferencia e insuficiencia que puede tener dicho control judicial y el nivel de intensidad de control judicial que debiera desarrollarse.

**Palabras clave:** inteligencia artificial, discrecionalidad, supervisión humana, control judicial, reserva de humanidad, deferencia.

**Abstract:** The focus of this article is on the possibilities and limits of fully automated discretionary administrative decision-making, the role of human supervision and the scope and intensity of judicial control of such decisions. The study stops briefly to explain some essential concepts linked to AI in order to clarify their use in this analysis. Among them, a brief analysis is made of the traditional concept of administrative discretion in Spain, highlighting its lack of depth and the need to rethink it in the light of good administration. From here, the legal limits of AI are addressed and how they should lead to avoid its use for the total automation of the exercise of discretionary powers, in the current state of the art. This limitation leads to what is called a reservation of humanity, which is discussed below. The analysis ends with the adaptation of judicial control techniques to the exercise of administrative powers with AI, the possible problems of indifference and insufficiency that such judicial control may have, and the level of intensity of judicial control that should be developed.

**Keywords:** artificial intelligence, discretion, human in the loop, judicial review, human reserve, deference.

**Sumario:** 1. INTRODUCCIÓN.— 2. HUMANOS E INTELIGENCIA ARTIFICIAL: EMPATÍA E INFERENCIAS ABDUCTIVAS: 2.1. IA y humanos; 2.2. Clases de IA, tipo de participación de la IA en la toma de decisiones administrativas y clase de potestad a ejercer: buena administración y discrecionalidad; 2.3. La IA simbólica o de reglas preprogramadas: el problema jurídico de la *rigidificación (fettering)* de la discrecionalidad; 2.4. La IA basada en estadística: el problema jurídico de la falta de empatía humana y de ausencia de capacidad de realizar inferencias abductivas: 2.4.1. Empatía humana; 2.4.2. Inferencias abductivas.; 2.5. Derecho a una buena administración:

*audi alteram parte*, participación ciudadana y motivación; 2.6. Consecuencias de la imposibilidad de eliminar totalmente la discrecionalidad, de la falta de empatía, de la incapacidad de abducción de la IA, de la audiencia y la participación y de la motivación en el ámbito administrativo.— 3. EN DEFINITIVA: RESERVA DE HUMANIDAD.— 4. VIGILANCIA HUMANA DE LA IA: *THE HUMAN IN THE LOOP*.— 5. EL CONTROL JUDICIAL DE LAS DECISIONES ADMINISTRATIVAS ADOPTADAS CON INTELIGENCIA ARTIFICIAL, CON ESPECIAL ATENCIÓN A LAS DISCRECIONALES: 5.1. Control judicial, IA y actividad administrativa; 5.2. Adaptación del control judicial del ejercicio de potestades por humanos a su posible ejercicio por máquinas: ¿Indiferencia insuficiencia o deferencia judicial?: 5.2.1. Técnicas jurídicas para el control judicial del ejercicio de potestades administrativas mediante IA: la relevancia del procedimiento administrativo complejo y la invalidez de las decisiones administrativas; 5.2.2. Cuestiones específicas relativas al control judicial del uso de IA por las Administraciones: indiferencia, insuficiencia, deferencia.— 6. CONCLUSIONES.— 7. BIBLIOGRAFÍA.

“Lejos de ser ciencia-ficción, la inteligencia artificial (IA) forma ya parte de nuestras vidas”  
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* {SWD(2018) 137 final}

“En el software, las cosas suelen ir mal”  
SHNEIDERMAN, B., “Algorithmic Accountability: designing for safety through human-centered independent oversight”, *Turing Lecture*, 31<sup>st</sup> May 2017.

“La deferencia es respeto y el respeto fortalece en el ejercicio de sus funciones a quien lo tiene,  
(...), y el respeto, se gana”

RODRÍGUEZ PONTÓN, F., “El control judicial en el ámbito de la regulación económica: Comisión Nacional de los Mercados y la Competencia (CNMC) y Comisión Nacional del Mercado de Valores (CNMV) un relato de la deferencia en el marco del Derecho común europeo”, en MALARET, E., *Autonomía Administrativa, decisiones cualificadas y deferencia judicial*, Thompson Reuters Aranzadi, 2019, pág. 212.

## 1. INTRODUCCIÓN

La adopción de decisiones administrativas a través de máquinas nos obliga a repensar algunas de las categorías y técnicas esenciales del Derecho Administrativo, para reflexionar sobre si las mismas son actualmente idóneas para un escenario en el que de forma creciente las decisiones administrativas no serán adoptadas ya únicamente por humanos, sino por humanos con ayuda de máquinas o por máquinas, directamente, sin decisor humano.

El uso de sistemas algorítmicos con IA creemos que obliga a reflexionar, por ejemplo, sobre el concepto histórico de la discrecionalidad administrativa en España, así como cuál deba ser el papel humano en su ejercicio y, en general, su relación con las máquinas. También nos obliga a considerar si los estándares tradicionales de control judicial contencioso siguen siendo útiles en este contexto o si hay que adaptarlos en relación con la IA.

Por lo que se refiere a la relación entre humanos y máquinas, ya en los años 60 del pasado siglo Lickelider, psicólogo de formación y uno de los padres de la IA, se ocupó

de esta cuestión, en un famoso artículo en el que abogaba por una colaboración en la que las máquinas desarrollaran las labores rutinarias aportando a los humanos sus capacidades singulares (LICKELIDER, 1960).

Sin embargo, fue un militar de la extinta Unión Soviética, el teniente Stanislav Petrov, quien experimentó en los años ochenta, en la práctica, las tensiones que pueden surgir en la relación máquina-humano. La mañana del 26 de septiembre de 1983, las computadoras del centro donde se hallaba como responsable detectaron el lanzamiento de misiles nucleares estadounidenses. Petrov tomó la decisión de que se trataba de una falsa alarma, generada por el sistema de algoritmos existente, y no actuó. Esta decisión, que se dio a conocer años después, posiblemente salvó al mundo de una guerra nuclear<sup>1</sup>.

En las líneas que siguen, de los temas posibles en esa relación entre seres humanos e IA se van a seleccionar los referidos al papel de los humanos respecto a la *limitación y supervisión* del funcionamiento de la IA en relación con las decisiones administrativas discrecionales formalizadas, esto es adoptadas a través de un procedimiento administrativo, por lo que se refiere a su posible automatización total. En la primera, limitación, son los humanos, a través de los mecanismos propios del Estado Democrático y de Derecho los que deciden prohibir el uso de la IA en ciertos ámbitos, por diversas razones jurídicas que analizaremos. En cambio, en el segundo caso, aceptado dicho uso, la supervisión humana deberá de garantizar el correcto funcionamiento del sistema para evitar la violación de derechos o intereses legítimos de los ciudadanos afectados por la IA usada por la Administración.

En ambos casos, el Derecho tiene que hacer frente a una posible situación que hasta hace relativamente poco era inédita: que quien tome la decisión final no sea un humano, un gestor público, un burócrata, sino una IA, un sistema algorítmico que hace uso de *Big Data*, una máquina, pues. Ésta puede servir de apoyo para la toma de la decisión final administrativa por parte de un ser humano (resolución, contrato o convenio, reglamento, plan), caso de decisión semiautomatizada, o bien, prescindiendo totalmente de éste, para adoptarla directamente mediante el correspondiente procedimiento, caso de decisión totalmente automatizada (art. 41.1 LRJSP<sup>2</sup>), supuesto éste que es el que vamos a considerar aquí.

Frente a esa intervención de las máquinas, el Derecho, como decimos, puede decidir, prohibirla en ciertos casos, como veremos ya sucede, debido a determinadas razones, o aceptarla, pero con supervisión humana o sin ella. Una pregunta relevante

---

<sup>1</sup> Véase: [Muere a los 77 años Stanislav Petrov, el hombre que salvó al mundo de una guerra nuclear entre la Unión Soviética y Estados Unidos —BBC News Mundo](#). Existe un documental al respecto: [This Man Single Handedly Stopped All Out Nuclear War— The Man Who Saved The World —War Documentary— Bing video](#), así como una entrevista con Petrov, quien murió en 2017, a la edad de 77 años: [Stanislav Petrov, who averted possible nuclear war, dies at 77 —BBC News— YouTube](#)

<sup>2</sup> “Se entiende por actuación administrativa automatizada, 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.”

entonces puede ser la siguiente: ¿por qué el Derecho se debe ocupar de prohibir o imponer una supervisión humana respecto al uso de la IA por las AAPP?

A partir de estas ideas iniciales expuestas, el objeto de este artículo se centra en las posibilidades y límites de la adopción totalmente automatizada de decisiones administrativas discrecionales, el papel de la supervisión humana en el uso de la IA, así como el alcance e intensidad del control judicial de las decisiones automatizadas en general, con especial atención a las discrecionales. Si bien el objeto del trabajo puede sostenerse que es doble, creemos que la exposición conjunta de las limitaciones y obligaciones en el desarrollo de decisiones automatizadas discrecionales y de su control judicial tiene todo el sentido, pues sólo tras la exposición de las primeras se está en condiciones de entender el alcance del segundo, existiendo una conexión entre ambos aspectos que se pretende resaltar aquí.

La metodología empleada es la jurídica, aunque también se tienen en cuenta aportaciones de otras ciencias, puesto que ello deviene inevitable e importante<sup>3</sup>. Se emplean, asimismo, elementos de comparación con otros sistemas jurídicos, si bien éste no es un estudio de Derecho comparado.

Algunas de las principales preguntas que pretende abordar este trabajo son las siguientes, tomando como referencia el estado *actual* de desarrollo tecnológico de la IA:

¿Es posible y deseable emplear IA para ejercer potestades discrecionales? ¿Con qué requisitos y límites, en su caso? ¿Deben quedar reservadas a los seres humanos gestores públicos ciertas funciones públicas? ¿Qué relevancia tiene la supervisión humana de la IA respecto de la protección de las personas que se van a ver afectadas por decisiones administrativas adoptadas mediante IA y de su derecho a una buena administración?

Y, como cierre del sistema en un Estado de Derecho:

¿Debe ser el juez o la jueza del contencioso-administrativo más deferente, menos deferente o igual de deferente de lo que lo es ante decisiones humanas cuando nos hallamos ante decisiones adoptadas por máquinas o con ayuda de éstas? En otras palabras, ¿la intervención de la IA en la toma de una decisión administrativa debe alterar el alcance e intensidad de su control judicial de algún modo?

A partir de estas reflexiones y cuestiones, el estudio se estructura de la siguiente manera. Tras esta introducción, me detengo brevemente en la explicación de algunos conceptos esenciales vinculados a la IA para clarificar su uso en este análisis. Realizo luego un breve análisis del concepto tradicional de discrecionalidad administrativa

---

<sup>3</sup> Ello va en la línea, por ejemplo, de los trabajos de la Liga de Universidades Europeas de Investigación (LERU, que en 2017 publicó un documento titulado (traducción nuestra) “La interdisciplinariedad y la universidad intensiva en investigación del siglo XXI” (DARBELLAY, WERNLI, MAES, 2017). El documento considera que la colaboración entre disciplinas académicas (interdisciplinariedad) es esencial y puede ir más allá incluyendo la asociación entre el mundo académico y la sociedad en general (enfoque transdisciplinar).

en España, poniendo de relieve su falta de profundidad. Por ello, intento incorporar una visión más sofisticada de lo que significa ejercer la discrecionalidad por humanos/as, de la mano de la buena administración, para plantearme a continuación si mi enfoque tiene alguna relevancia respecto al posible ejercicio de discrecionalidad administrativa por máquinas.

Tras este previo recorrido por algunos conceptos claves me ocupo en los apartados siguientes de cuáles son los límites jurídicos de la IA y cómo dichos límites deberían conducir a evitar su uso para la automatización total del ejercicio de potestades discrecionales, en el estado actual de la técnica y de acuerdo con el principio jurídico de precaución. Esa prohibición conduce a lo que hemos denominado una reserva de humanidad, que se expone a continuación. Finalizo mi análisis exponiendo la adaptación de las técnicas de control judicial al ejercicio de potestades administrativas con IA, los posibles problemas de indiferencia e insuficiencia que puede tener dicho control judicial y concluyo con algunas reflexiones respecto al nivel de intensidad de control judicial que debiera desarrollarse en estos momentos respecto a ese tipo de decisiones.

Mi hipótesis de partida es que IA y humanos adoptan decisiones de forma distinta en el momento actual y que ello debería conducir a la limitación del uso de la primera, de acuerdo con el principio de precaución, en relación con el ejercicio de la discrecionalidad, en el que sólo debería cumplir funciones de apoyo al juicio volitivo humano. Además, el control judicial del uso de IA en el ejercicio de potestades administrativas en general no debe dar lugar a la deferencia —entendiendo ésta aquí como una modulación de su intensidad en función del contexto administrativo organizativo y procedimental y elementos democráticos y/o de *expertise* administrativa específica— si se busca una protección adecuada de los derechos de los ciudadanos, incluido el derecho a la buena administración, frente al uso de la IA. Por el contrario, ante el panorama actual que describiremos, los jueces y las juezas deben estar muy atentas a los posibles impactos del uso de la IA en los derechos de los ciudadanos, puesto que, si bien la IA puede redundar en una mejora de la gestión pública, también puede generar sistemas opresivos y con profundos impactos negativos, lo que exige un escrutinio judicial adecuado de la actuación administrativa.

## 2. HUMANOS E INTELIGENCIA ARTIFICIAL: EMPATÍA E INFERENCIAS ABDUCTIVAS

### 2.1. IA y humanos

Las ventajas de la IA frente a nosotros los humanos parecen indudables: una capacidad de procesamiento mucho mayor (la memoria humana 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 administración (COGLIANESE y LAI, 2022), la posibilidad de evitar los sesgos cognitivos de los humanos (PONCE, 2022b) y de acabar con el *ruido* humano, esto es la existencia de

decisiones distintas no queridas en condiciones similares (KAHNEMAN, SIBONY, SUNSTEIN, 2021). Asimismo, como elemento positivo añadido, el uso de la IA tiene 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 privadas y públicas con el consiguiente desempleo (para el sector público, PASTOR, A. y NOGALES, P. 2019), pasando por el hecho de que las correlaciones generadas por la IA pueden dar lugar a errores y que dichas correlaciones tienen una tendencia *conservadora*, al primar el *statu 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 humanos que interaccionan con los algoritmos y los datos, (PONCE, 2022), apareciendo sesgos como el de *confirmación*, por ejemplo, al seleccionar dichos datos (MARTÍNEZ, AGUDO y MATUTE, 2022).

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 navegador (BRIDLE, 2020).

Por otro lado, 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 procedimiento, en términos de interdicción de la arbitrariedad (recogida en el art. 9.3 de la Constitución española), es decir la prohibición de decisiones irracionales o no motivadas, y también al respeto de una buena administración, de la que luego nos ocuparemos, esto es, el cumplimiento de 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.

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 Reglamento de la Unión Europea de IA



aprobado por el Parlamento el 13 de marzo de 2024 (en adelante, RUEIA<sup>4</sup>), para justificar la “desconfianza” humana en la IA, debido, como dice su art. 14, a “...los riesgos para la salud, la seguridad o los derechos fundamentales que pueden surgir cuando se utiliza un sistema de IA de alto riesgo conforme a su finalidad prevista o cuando se le da un uso indebido razonablemente previsible, en particular cuando dichos riesgos persistan a pesar de la aplicación de otros requisitos establecidos en la presente sección”<sup>5</sup>.

## **2.2. Clases de IA, tipo de participación de la IA en la toma de decisiones administrativas y clase de potestad a ejercer: buena administración y discrecionalidad.**

Para abordar el objeto de este estudio es necesario realizar algunas distinciones que serán importantes para extraer conclusiones. En concreto, es interesante utilizar una tríada de conceptos que, como veremos, se interrelacionan entre sí:

a) Según la clase de IA que se emplea: frente a la simbólica, que desarrolla inferencias deductivas (*Good Old Fashioned IA* o *GOFAD*), se sitúa la no simbólica, conexionista o basada en estadística (aprendizaje automático, aprendizaje profundo), que realiza inferencias inductivas.

b) Según el tipo de participación de la IA en la toma de la decisión administrativa formalizada: obviando ahora matices, la participación puede consistir en ser soporte y ayuda para la toma de una decisión administrativa humana, que será semiautomatizada, frente a una participación de la IA que supone la adopción de la decisión final que termina el procedimiento, esto es una decisión automatizada (para distinciones más detalladas, que creemos se pueden reconducir a estas dos variantes, véanse los seis tipos posibles de interacción entre humanos/as y máquina identificados por Ulrik Rohel y expuestos por MORAL, 2022).

c) Según la clase de potestades administrativas que se ejercen: que por lo que nos interesa aquí, la IA puede aplicarse al ejercicio de potestades regladas o discrecionales.

A continuación, vamos a analizar cómo se combinan estas tres perspectivas y que conclusiones podemos extraer de todo ello. Para ello, debemos partir de tres conceptos cruciales, uno técnico y los otros dos jurídicos: la IA, la buena administración y la discrecionalidad administrativa.

---

<sup>4</sup> Resolución legislativa del Parlamento Europeo, de 13 de marzo de 2024, sobre 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)0206 – C9-0146/2021 – 2021/0106(COD)): [Textos aprobados - Reglamento de Inteligencia Artificial - Miércoles 13 de marzo de 2024 \(europa.eu\)](#)

<sup>5</sup> Los sistemas de IA de alto riesgo vienen definidos en el art. 6 en combinación con los Anexos I y III del Reglamento.



## a) Tipos de IA

La Comisión Europea se ha referido a que la expresión IA se usa para hablar de un grupo de tecnologías que están automatizadas, es decir, que pueden decidir por sí solas. Se trata de programas informáticos que son diseñados para resolver de forma autónoma. Son 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 objetivos específicos<sup>6</sup>. Estos sistemas, basados en el uso de algoritmos y de datos masivos<sup>7</sup>, pueden consistir en un simple programa informático (por ejemplo, motores de búsqueda o sistemas de reconocimiento facial o de voz) pero también pueden estar incorporados en dispositivos *de hardware* (como robots o automóviles autónomos).

La IA incluiría tanto la IA basada en reglas (o simbólica, o determinista o *GOFAI*, *good old-fashioned AI*) en desarrollo desde los años 50 del pasado siglo hasta los 90 (por ejemplo, un programa básico del juego de ajedrez), basado en razonamientos lógicos del tipo de inferencia deductiva, basada en premisas que llevan a una conclusión (si a y b, entonces c), como la IA que con uso de la estadística para inferencias inductivas conduce a la AI de aprendizaje automático o profundo (*machine o deep learning*), un tipo de IA entrenada para absorber una gran cantidad de datos, encontrar patrones entre ellos y arrojar ciertas conclusiones basadas en el análisis.

Este segundo tipo de IA es capaz de proporcionar más conclusiones diferentes al captar más datos, mediante un entrenamiento y un *autoaprendizaje*, sea este supervisado (esto es con datos etiquetados por humanos, lo que puede ayudar a minimizar los errores) o no supervisado (o auto supervisado o por refuerzo). Este segundo tipo de IA (también llamado conexionista o subsimbólico) también incluiría el llamado aprendizaje profundo (*deep learning*), un tipo de IA con aprendizaje autónomo con modelos matemáticos de gran complejidad<sup>8</sup>.

Estos dos tipos de IA no tienen por qué considerarse separadamente: así, por ejemplo, un *chatbot* usa ambos tipos de IA<sup>9</sup>.

---

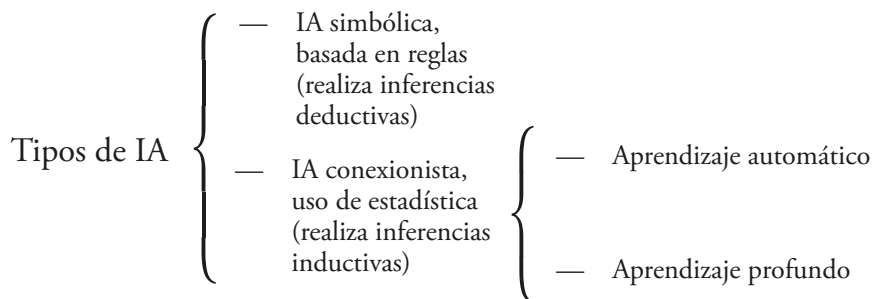
<sup>6</sup> Comunicación de la Comisión al Parlamento Europeo, al Consejo Europeo, al Consejo, al Comité económico y social europeo y al Comité de las regiones, *Inteligencia artificial para Europa* {SWD(2018) 137 final}: <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52018DC0237>

<sup>7</sup> Como es sabido, un algoritmo es una secuencia de operaciones y reglas, como una ecuación matemática, diseñada por una persona o grupo de personas para resolver un problema. Por su parte, los datos masivos, o *big data*, son ingentes cantidades de datos caracterizados por su volumen, variedad y velocidad, las llamadas tres v.

<sup>8</sup> Comunicación citada en nota anterior.

<sup>9</sup> El RUEIA define a los sistemas de IA como “un sistema basado en una máquina diseñado para funcionar con distintos niveles de autonomía, que puede mostrar capacidad de adaptación tras el despliegue y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar información de salida, como predicciones, contenidos, recomendaciones o decisiones, que puede influir en entornos físicos o virtuales” (art. 3. 1).

Lo dicho puede esquematizarse así:



#### b) Buena administración y discrecionalidad

Las últimas décadas han conocido una creciente relevancia de la buena administración, sobre la que ya existen numerosos estudios a nivel nacional e internacional y abundante jurisprudencia internacional, europea y española (PONCE, 2019b). A los limitados efectos de este análisis, simplemente queremos destacar que, como es sabido, el derecho a una buena administración puede hallarse en el art. 41 de la Carta de Derechos Fundamentales de la Unión Europea y también en ordenamientos nacionales como el español (entre otros, para Italia, véase art. 97 de su Constitución), sea implícitamente en la Constitución española (como repetidamente ha sostenido el Tribunal Supremo), sea en diversos Estatutos de Autonomía aprobados en las últimas décadas, sea en legislación de procedimiento administrativo, régimen jurídico, contratación o acceso a la información y buen gobierno (PONCE, 2023b).

Referencias a la buena administración se encuentra en la jurisprudencia del Tribunal Europeo de Derechos Humanos (que lo extrae del Convenio Europeo de Derechos Humanos de 1950, pese a no estar reconocido en éste explícitamente, PONCE y VILLORIA, 2020, 2021, 2022, 2023), del Tribunal de Justicia de la Unión Europea, en literalmente miles de sentencias desde 1955, en las que ha reconocido a la buena administración como una de las tradiciones constitucionales comunes a los Estados miembros de la Unión Europea) y también en los tribunales nacionales (HOFMANN, 2020).

Por ejemplo, en España, la jurisprudencia del Tribunal Supremo, como dijimos, ha establecido la existencia de un derecho a una buena administración implícito en la Constitución española, en numerosas sentencias, por ejemplo, STS de 4 de noviembre de 2021, ECLI:ES:TS:2021:4117. Asimismo los Tribunales Superiores de Justicia de las Comunidades autónomas hacen un uso frecuente de la buena administración en el control de las AAPP (PONCE, 2023b).

Un elemento relevante y propio de la buena administración, que actúa como principio y derecho, que la distingue y la aleja de ser un concepto paraguas sin substantividad propia, el cual destaca al analizar toda la normativa y la jurisprudencia aludidas, es el de que los cargos electos, los altos cargos y los empleados públicos en

cumplimiento de sus obligaciones de buen gobierno y buena administración deberán actuar con diligencia debida y debido cuidado, un elemento enfatizado hasta la saciedad por el TJUE y nuestro TS, lo que implica para las AAPP (PONCE, 2019b):

- actuar solo en ausencia de conflictos de intereses,
- con la información necesaria,
- teniendo en cuenta los factores relevantes y
- otorgando a cada uno de los mismos su debida importancia en la decisión, excluyendo de su consideración todo elemento irrelevante,
- motivando suficiente y congruentemente la misma en relación con la fundamentación existente en el expediente.

En definitiva, lo que supone la buena administración, pues, es un replanteamiento radical de la tradicional idea de discrecionalidad administrativa empleada en la doctrina y jurisprudencia española, como libertad de elección entre indiferentes jurídicos. La discrecionalidad administrativa en el siglo XXI no implica ni libertad ni indiferencia para el Derecho, sino la existencia de un poder entregado para ser ejercido con la debida diligencia y el debido cuidado en la toma en consideración de los elementos relevantes para adoptar la mejor decisión posible en servicio a los intereses generales. Como hemos expuesto con más detalle en otro momento, la discrecionalidad no puede ser arbitrariedad, como es sobradamente sabido, pero ello no es suficiente. Debe ser, además, buena administración (PONCE, 2016).

Sin embargo, conviene antes de seguir precisar el concepto de discrecionalidad administrativa que aquí se maneja, evitando darlo por supuesto. Cuando aquí hablamos de discrecionalidad administrativa nos referimos a todos aquellos casos en que las administraciones toman decisiones entre alternativas jurídicamente relevantes, bajo la obligación de debida diligencia y debido cuidado, disponiendo de un margen de apreciación para elegir lo más conveniente a los intereses generales. Ello supone la existencia de márgenes de apreciación a desarrollar por la Administración en base a criterios no jurídicos, sean estos de tipo político, económico, estético, técnico<sup>10</sup>, etc., incluyendo, por tanto, la determinación de ciertos conceptos jurídicos indeterminados con un componente valorativo. Se trata en todos estos supuestos de ámbitos, como señala Bacigalupo, en los que estamos ante la adopción administrativa de “una

---

<sup>10</sup> Incluyendo por tanto la escurridiza noción de discrecionalidad técnica, que entendemos, siguiendo a LAGUNA DE PAZ (2017), como la existencia de un margen de apreciación que el uso de la técnica no elimina, pues “existe un margen de apreciación que no puede ser controlado con arreglo a criterios jurídicos, ni con la realización de pruebas o la solicitud de otros informes, que —por regla general— no añadirían mayor certeza al análisis, sino que simplemente ofrecerían una valoración alternativa. A este respecto, no puede dejar de reconocerse que las respuestas que ofrece la técnica no siempre son unívocas. La solución puede variar en función del método aplicado o, en último término, de la escuela de pensamiento a la que se adscriba su autor, sin que exista un consenso científico que permita reputar unos u otros resultados como equivocados. Esto explica que el juez deba respetar el inevitable margen de discrepancia que existe en todos los sectores de saber, siempre que la solución técnica, aunque discutible, resulte aceptable para un porcentaje significativo de la comunidad científica de que se trate”.

decisión volitiva que en modo alguno emana de un juicio meramente cognitivo” (BACIGALUPO, 2023, 150).

### 2.3. La IA simbólica o de reglas preprogramadas: el problema jurídico de la rigidificación (*fettering*) de la discrecionalidad.

El correcto ejercicio de la discrecionalidad de acuerdo con la buena administración exige, pues, aplicarla a cada caso en función de las circunstancias específicas del mismo, no siendo posible su cristalización de antemano en decisiones ya pre adoptadas, lo que en el ámbito angloamericano se denomina *fettering*, debiendo el decisor administrativo cumplir la obligación de tener en cuenta las circunstancias del caso<sup>11</sup>, obligación jurídica que diversos ordenamientos jurídicos destacan, especialmente los de aquellos países donde la visión de la discrecionalidad no es esencialmente negativa, probablemente por no haber experimentado períodos prolongados de autoritarismo político ni dictaduras y haber contado con regímenes democráticos durante siglos<sup>12</sup>.

CRAIG (2021) señala al respecto que el principio general en Derecho británico es que un organismo público dotado de poderes discrecionales concedidos por ley no puede adoptar una política o norma que le permita resolver un caso sin ninguna consideración de los méritos del solicitante individual. La línea de autoridad dominante permite al organismo aplicar su norma, siempre que se conceda al individuo la oportunidad de impugnar su aplicación al caso concreto. La política debe ser legítima dado el marco legal en el que se ejerce la discrecionalidad. Debe basarse en consideraciones pertinentes y no perseguir fines indebidos. Estos controles son necesarios ya que, de lo contrario, una autoridad pública podría eludir las limitaciones normales al ejercicio de la discrecionalidad mediante la elaboración de políticas generales. Sin embargo, señala CRAIG también existe un principio de coherencia, que crea la presunción de que un organismo público seguirá su propia política. Si pretende apartarse de esa política, debe haber buenas razones para ello y éstas deben comunicarse al solicitante.

---

<sup>11</sup> “Where a decision maker has a discretionary power, he should take individual circumstances into account when exercising it, they should make each decision on its merits rather than taking a one size fits all approach, and they should be prepared to depart from policies or guidelines where appropriate. Otherwise they may have acted illegally by fettering their discretion. This will particularly be where decisions involve human right issues, with a balance of rights and interests, and thus necessarily require discretionary powers to be exercised with due consideration” (COBBE, 2018)

<sup>12</sup> Existe, por ejemplo, en el *Common Law* un deber de la Administración de ejercer su discrecionalidad previo análisis específico del supuesto concreto. Este deber ha sido afirmado en distintas ocasiones por la jurisprudencia, la cual ha declarado gráficamente en el caso *British Oxygen v. Board of Trade* que “*the general rule is that anyone who has to exercise a statutory discretion must not “shut his ears” to an application*”, afirmando la necesidad de un análisis administrativo de las circunstancias del caso antes de adoptar la decisión discrecional. Un análisis de estos casos con cita específica de la jurisprudencia y de la doctrina de diversos países, también de la tradición del *Civil Law*, puede hallarse en PONCE, 2001.

Estos preceptos doctrinales expuestos pueden resultar problemáticos cuando las decisiones se toman mediante sistemas de inteligencia artificial en la opinión de CRAIG. Esto se debe a que dichos sistemas suelen estar orientados a la uniformidad de los resultados, en lugar de tener en cuenta consideraciones específicas de cada caso concreto. Esto es especialmente cierto cuando un organismo público se basa exclusiva o irreflexivamente en un algoritmo a la hora de ejercer su poder discrecional<sup>13</sup>.

En la misma línea se pronuncia PILVING a propósito de una iniciativa legislativa que tuvo entrada a mediados de 2022 en el Parlamento estonio (PILVING, 2023). El proyecto de ley permitiría excepcionalmente el ejercicio de potestades discrecionales mediante ley específica o en el caso en que la discrecionalidad pueda especificarse de forma adecuada en una directriz administrativa interna, un instrumento de *soft law*, que parametrizaría la discrecionalidad.

Para este autor, esta estrategia “puede parecer problemática y conducir a deficiencias en el uso de la discrecionalidad”, dado que “las autoridades disponen de potestades discrecionales precisamente porque la naturaleza de muchas funciones públicas exige un juicio humano que no puede programarse de antemano y que las máquinas no pueden substituir”, por lo que la solución diseñada por la máquina sólo puede ser “provisional y condicional: la autoridad debe tener la posibilidad y el deber de desviarse de la propuesta basada en el algoritmo siempre que ello parezca justo y necesario a la luz de circunstancias no consideradas en la fase en que se elaboró el algoritmo sobre la base de las directrices”.

Puede existir aquí una controversia basada en diversas tradiciones jurídicas y maneras de afrontar la discrecionalidad. Así, en el caso continental europeo, incluyendo a España, dado que la discrecionalidad es vista con desconfianza, debido a la historia autoritaria preexistente en diversos casos, a la *path dependency* y a la consecuente cultura política y administrativa existente cabría afirmar que la obsesión es eliminar la discrecionalidad, mediante instrucciones, circulares, protocolos, directrices, etc., mientras que, en cambio en el caso británico, los jueces son muy conscientes del valor positivo de la discrecionalidad en la gestión pública y no permiten su eliminación: así lo sostuvo, precisamente GIANNINI en 1965 en la presentación a la traducción al italiano del libro de WADE sobre Derecho administrativo inglés, como destaca DELLA CANANEA (2022).

Quizás el enfoque de GIANNINI era algo excesivo, por cuanto los jueces franceses y de otros países de la Europa continental también protegen la discrecionalidad, aunque no utilicen la doctrina británica del *fettering*. Incluso en el caso español, donde este tipo de reflexiones son inexistentes aún, en conexión con un análisis de la discrecionalidad que ha sido históricamente limitado, podemos encontrar algún caso judicial. Así, por ejemplo, cabe citar la STS de 28 de octubre de 2015, ECLI:ES:TS:2015:4576, por la que se anula la Disp. Adicional Segunda del Real Decreto 177/2014, por el que se regula el sistema de precios de referencia y agrupa-

---

<sup>13</sup> CRAIG, 2021, párrafo 10-025.

ciones homogéneas de medicamentos en el Sistema Nacional de Salud. Tal anulación deriva de que el reglamento fija un procedimiento de revisión de los precios de comercialización de determinadas presentaciones de los medicamentos que, en contra de lo establecido legalmente que implicaba otorgamiento de discrecionalidad administrativa, se articula por el Real Decreto como una decisión meramente aritmética que se adoptará de forma automática en función de un único parámetro que se crea.

En todo caso, esta obligación de respeto del ejercicio individualizado de la discrecionalidad, que en el sistema español se desprende sin esfuerzo del deber de diligencia debida o debido cuidado asociado al derecho a una buena administración citado, impediría una eliminación completa de la discrecionalidad prefigurándola rígidamente mediante su parametrización previa sin análisis del caso concreto.

## **2.4. La IA basada en estadística: el problema jurídico de la falta de empatía humana y de ausencia de capacidad de realizar inferencias abductivas.**

### *2.4.1. Empatía humana*

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, manteniendo el debido respeto al 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. 10.1 CE). Podemos definir la empatía, un concepto muy poco conocido entre nosotros en el ámbito jurídico a diferencia de otros países, muy básicamente como un componente cognitivo humano que permite comprender a los demás, incluyendo sus dificultades, vulnerabilidades y problemas, y adaptar nuestro comportamiento de forma correspondiente. Se trata del sentimiento de identificación con algo o alguien y la capacidad de identificarse con alguien y compartir sus sentimientos<sup>14</sup>. De hecho, formar parte de la humanidad implica una sensibilidad, una compasión por las desgracias de otras personas, pues es humano aquél que es comprensivo, sensible a los infortunios ajenos, como señala el Diccionario de la Lengua Española de la Real Academia Española<sup>15</sup>.

---

<sup>14</sup> La bibliografía sobre empatía es enorme. Desde una perspectiva divulgativa pero completa, puede verse MATTHIEU, 2016.

<sup>15</sup> Éste define a la empatía 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 por cien 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 (!) (PULLMAN, REFAIE, LALUMIÈRE y KRUPP, 2021), en vez de como lo que es: un trastorno grave de la personalidad, denominado por el DSM-5 trastorno de la personalidad antisocial<sup>16</sup>.

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 asimilar a la empatía con una emoción que va a conducir en el ámbito jurídico a la adopción de decisiones parciales y no objetivas, algo que normalmente viene a la mente de forma rápida: 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 IA 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 datos específicamente negativos, al cual se le ha denominado el primer sistema de IA psicópata<sup>17</sup>.

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...

En ese sentido, es destacable 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 comuni-

---

<sup>16</sup> El DSM-5 es el *Diagnostic and statistical Manual of Mental Disorders*, editado por la *American Psychiatric Association*.

<sup>17</sup> La BBC, por ejemplo, habla de algoritmo psicópata. Véase [Así funciona “la mente” de Norman, el algoritmo psicópata del MIT que solo ve lo más tenebroso de la red - BBC News Mundo](#)



cación 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*<sup>18</sup>. En 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 o jueza u oficial de policía<sup>19</sup>.

Mucho más recientemente se ha cuestionado la utilización de la automatización total en relación con la IA y las armas, aludiendo al riesgo de deshumanización de la guerra (TRUMBULL, 2020). Asimismo, en el ámbito del Derecho mercantil, CHAMORRO en España ha señalado como la sustitución de las decisiones de los administradores en las sociedades de capital por decisiones adoptadas por máquinas no es “conveniente”, puesto que la IA “carece del sentido común y las emociones que caracterizan a los humanos” (CHAMORRO, 2020).

Pasando de estas reflexiones generales y jurídicas del ámbito privado al ámbito del Derecho administrativo y de la gestión pública, hay que destacar que, en diversos 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 Derecho<sup>20</sup>, que cristaliza, por ejemplo, en el principio jurídico de equidad, que “habrá de ponderarse en la aplicación de las normas”, como manda el art. 3 del Código Civil español, principio cuya importancia ha sido destacada por ejemplo por NIETO (NIETO, 2021, NIETO y SÁNCHEZ, 2022).

De hecho, la empatía debe entenderse 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.

Así lo destacan tanto el Consejo de Europa como el *ombudsman* federal belga, por ejemplo. En el primer caso, el Consejo de Europa destaca su relevancia, porque nos lleva a comprender los intereses, las necesidades y los puntos de vista de los demás. La empatía es un importante determinante del comportamiento moral y un elemento

---

<sup>18</sup> Del que, por cierto, en 2022 hemos tenido un nuevo ejemplo con la IA LaMDA desarrollada por Google, un sistema mucho más sofisticado. 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.

<sup>19</sup> WEIZENBAUM, J. (1976) traducido al español un año más tarde con el título *La frontera entre el ordenador y la mente*, Ed. Pirámide, S.A. Madrid. 1977.

<sup>20</sup> RANCHORDAS, S. (2022).

necesario para construir comunidades morales<sup>21</sup>. En el segundo caso, el ombudsman belga afirma con rotundidad que la empatía es la clave en la base de una prestación de buenos servicios públicos y una condición necesaria para un servicio público humano que se adapta a la situación individual de cada ciudadano. Favorece los cambios de perspectiva, determina la forma en que miramos a una persona o a un grupo, y tiene un gran impacto en la forma en que percibimos y respondemos a los demás y a los acontecimientos de la vida (MÉDIATEUR FÉDÉRAL, 2021).

Para adoptar buenas decisiones (también administrativas) hace falta racionalidad, pero no sólo, como ha destacado el neurocientífico portugués que investiga en los EE. UU. Antonio DAMASIO, por ejemplo, en su conocido libro *El error de Descartes* (DAMASIO, 1994). En el ámbito jurídico, como en otros ámbitos profesionales, persiste la creencia arraigada, heredera 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. Usar la *fría racionalidad* conduce, en realidad, a malas decisiones (sobre el papel de las emociones en el correcto desempeño de la actividad administrativa por sus agentes y la necesidad de combinar la objetividad con la subjetividad, VAQUER, 2023, 41).

En este sentido, es muy destacable el trabajo de los profesores brasileños de Derecho REIS DE ALBURQUERQUE Y DE BRITO MACHADO SEGUNDO (2023), quienes han puesto de relieve como en la toma de decisiones razón y emociones, cuerpo y mente se entrelazan en los humanos, algo que la IA está lejos de conseguir, si lo consigue algún día.

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.

Nada de todo eso pueden hacer, hoy en día, y está por ver si lo logran alguna vez<sup>22</sup>, los sistemas algorítmicos y la IA. En consecuencia, esa falta de empatía de las

---

<sup>21</sup> <https://www.coe.int/fr/web/digital-citizenship-education/ethics-and-empathy>

<sup>22</sup> ¿Podrían adquirir empatía real en el futuro las máquinas y por tanto replantearnos algunas de las reflexiones que vamos a efectuar aquí? En un interesante artículo en coautoría con Man, aparecido en 2019 en *Nature* (MAN, K.; DAMASIO, A. (2019) y en su libro *Sentir y Saber* (2021, Destino), 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

máquinas es la que las hace inadecuadas para adoptar decisiones administrativas totalmente automatizadas que afectan a seres humanos.

#### 2.4.2. *Inferencias abductivas*

Además de emplear racionalidad y emociones, adoptar buenas decisiones (administrativas) significa pensar como lo hacen los humanos. Las máquinas no deciden como nosotros, que utilizamos inferencias denominadas abductivas, teorizadas por PIERCE en el siglo XIX, capaces de formular conjeturas e integrar el sentido común en un mundo cambiante (BRENT, 1993, 349).

Las máquinas realizan inferencias deductivas (IA basada en reglas) o inductivas (IA estadística) pero no inferencias abductivas. Las máquinas, a diferencia de nosotros no pueden realizar conjeturas vinculadas a una comprensión global del mundo y a las experiencias vividas, con sentido común, mediante este tipo de inferencias que, de acuerdo con Larson, “impregnan nuestra cultura” y “son en un sentido amplio lo que nos convierte en humanos. Es posible que el sueño de una IA capture algún día esos saltos de manera automática, pero mientras tanto deberíamos admitir que la verdad es que no sabemos cómo mecanizar nuestra experiencia” (LARSON, 2022, p. 226).

Esta imposibilidad de las máquinas de realizar abducciones puede tener conexión con su falta de empatía, ya señalada antes, puesto que como se ha dicho (la traducción es nuestra): “La abducción requiere la capacidad de considerar un punto de vista singular implicado en el razonamiento. En otras palabras, el modelo abductivo presupone una forma de empatía. (...) La empatía es, por tanto, uno de los motores cognitivos cruciales del sistema abductivo. Sin empatía, el razonamiento abductivo es imposible” (HUYS, 2018).

#### 2.5. **Derecho a una buena administración: *audi alteram parte*, participación y motivación.**

Una regla clásica del Estado de Derecho, que forma parte del derecho a una buena administración (por ejemplo, art. 41 de la Carta de Derechos Fundamentales de la Unión Europea), como es sabido, es la imposibilidad de adoptar decisiones que afecten a una persona sin escuchar a ésta antes y considerar sus aportaciones (en Derecho español, art. 82 LPAC). Ello impide automatizar totalmente la toma de decisiones, puesto que exige una previa intervención humana en el procedimiento administrativo de adopción de la decisión que considere diligentemente, de acuerdo con la

---

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.

En todo caso, se trata de una hipótesis. El presente estudio elaborado en 2023 se basa en la situación existente en estos momentos, que no parece vaya a cambiar ni a corto ni a medio plazo.

buena administración, las alegaciones del afectado. Además, en caso de existencia de información pública en la toma de la decisión, “quienes presenten alegaciones u observaciones en este trámite tienen derecho a obtener de la Administración una respuesta razonada” (art. 83 LPAC), por lo que la automatización total de todas las fases de la decisión no es posible jurídicamente.

Por otro lado, exigencias de racionalidad, buena administración y transparencia explican la necesidad de motivar las decisiones administrativas, incluidas las adoptadas con IA (art. 9.3 CE prohibiendo la arbitrariedad, art. 35 LPAC). Ahora, el art. 86 RUEIA, titulado “derecho a explicación de decisiones tomadas individualmente” señala, por su parte, que toda persona que se vea afectada por una decisión que se adopte basándose en los resultados de un sistema de IA de alto riesgo que figure en el anexo III, con excepción de los sistemas enumerados en su punto 2, y que produzca efectos jurídicos o le afecte considerablemente del mismo modo, de manera que considere que tiene un efecto perjudicial para su salud, su seguridad o sus derechos fundamentales, tendrá derecho a obtener del responsable del despliegue explicaciones claras y significativas acerca del papel que el sistema de IA ha tenido en el proceso de toma de decisiones y los principales elementos de la decisión adoptada.

Sin embargo, las decisiones discrecionales que pudieran adoptarse con IA estadística (aprendizaje automático o profundo) no podrían ser justificadas conforme a hechos, fundamentos de derecho y criterio de oportunidad manejado, al encontrarnos ante cajas negras. Si esta es la situación actual, debería concluirse entonces que si se pretendiera adoptar decisiones administrativas discrecionales mediante sistemas que no pueden justificarlas (más allá de la explanabilidad del funcionamiento del propio sistema, esto es ofreciendo razones basadas en criterios no jurídicos) tal pretensión resultaría en decisiones arbitrarias contrarias al ordenamiento jurídico vigente. En definitiva, no es posible emplear por las Administraciones sistemas que no permitan motivar sus decisiones.

## **2.6. Consecuencias de la imposibilidad de eliminar totalmente la discrecionalidad, de la falta de empatía, de la incapacidad de abducción de la IA, de la audiencia y participación y de la motivación en el ámbito administrativo.**

En resumen, el uso de sistemas algorítmicos e IA para adoptar decisiones administrativas discrecionales de forma totalmente automatizada puede presentar problemas en relación con el adecuado uso de la discrecionalidad administrativa en un Estado de Derecho y violar el derecho a una buena administración, en nuestra opinión, frente a cualificadas opiniones que no ven mayor problema en ello (por ejemplo, GAMERO, 2023 o HUERGO, 2023<sup>23</sup>), debido a las razones expuestas.

---

<sup>23</sup> GAMERO, 2023, sostiene en la pág. 450: “En mi opinión, fuera de los ámbitos expresamente vetados por el Derecho positivo no existe una objeción de carácter dogmático que impida adoptar

McCANN, quien se muestra en contra del uso de IA para adoptar decisiones discrecionales, añade una interesante consideración complementaria en contra del mismo: el ejercicio por máquinas de la discrecionalidad conllevaría el desplazamiento del compromiso moral humano que el ejercicio de la misma debe conllevar, de tal manera que los servidores públicos tendrían la (muy humana) tentación de escudarse tras las decisiones maquinales, no afrontando su responsabilidad última por ellas (McCANN, 2023).

Por consiguiente, la IA para el ejercicio totalmente automatizado de potestades administrativas discrecionales (esto es, todos aquellos casos en que exista un margen de valoración) debería quedar, en principio, descartada (y prohibida legalmente de forma expresa), en aplicación del principio/derecho de buena administración, en conexión además con el principio de precaución relacionado con el uso de IA (SOM, C., HILTY, L.M and KÖHLER, A.R. 2009, PONCE, 2022c), salvo decisión legal expresa distinta, que deberá justificar la ausencia del violación del principio constitucional de buena administración.

Por tanto, el uso de IA para una total automatización debería quedar restringida al ejercicio de potestades regladas y aún en estas no será posible una automatización total si hay obligación de apertura de trámite de audiencia.

Eso sí, en todos los casos puede emplearse la IA como apoyo a la final decisión administrativa humana.

Todo lo expuesto respecto a las decisiones administrativas discrecionales automatizadas conduce, entonces, al uso de dos vías (la reserva de humanidad y la supervisión humana), que debemos entender basadas en el principio de precaución aplicado al uso de la AI en el ejercicio de potestades administrativas, para limitarla y vigilarla en aquellos ámbitos que una sociedad considere especialmente sensibles, dadas las limitaciones intrínsecas de la tecnología ya expuestas. Veamos ahora ya brevemente, a la luz de todo lo antedicho, estos dos conceptos.

---

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”. Por su parte, HUERGO, 2023, 49 y 58, tampoco ve problema alguno, afirmando que la IA conexionista basada en la estadística “influye” en el contenido de la decisión y tiene sentido “en decisiones...en las que la Administración tiene un margen de discrecionalidad o bien hay conceptos legales indeterminados...”, siendo su uso “posible”, pues “Teóricamente sí” puede emplearse donde exista discrecionalidad administrativa.

### 3. EN DEFINITIVA: RESERVA DE HUMANIDAD

En cuanto al concepto de reserva de humanidad, se trata de la decisión humana de excluir el uso de la IA de ciertos ámbitos porque se considera inadecuado el uso de ésta de acuerdo con el Estado de Derecho de conformidad con todo lo expuesto anteriormente<sup>24</sup>.

La idea de reserva de humanidad es comparable a la idea de reserva del ejercicio de potestades para los funcionarios de carrera<sup>25</sup>. 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 fundamento similar creemos que se encuentra implícitamente en las normas jurídicas que ya requieren que sólo sean humanos los que tomen ciertas decisiones (impidiendo pues, decisiones administrativas completamente automatizadas) debido a diversas razones, entre las que se encuentran las limitaciones que hemos expuesto con detalle en el epígrafe anterior.

Así, el art. 22 del Reglamento General de Protección de Datos de la Unión Europea (en adelante, RGPD<sup>26</sup>) configura, de hecho, una reserva de humanidad como

---

<sup>24</sup> La expresión Reserva de Humanidad fue escuchada por mí por primera vez en boca de la profesora Silvia Díez en el marco de las conversaciones sostenidas en uno de los Seminarios de Teoría y Metodología del Derecho que son organizados cada seis meses desde hace más de diez años. Introducida en el ámbito de la IA en España en 2019 (PONCE, 2019), la expresión española ha conocido fortuna más allá de nuestras fronteras (GALLONE, 2023).

<sup>25</sup> Reserva prevista en España, como es sabido, en el art. 9.2 del Real Decreto Legislativo 5/2015, de 30 de octubre, por el que se aprueba el texto refundido de la Ley del Estatuto Básico del Empleado Público.

<sup>26</sup> “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.

principio, que puede ser excepcionado en diversos casos, como es sabido, entre los que se encuentra que una norma específicamente prevea y regule el uso de la IA. Norma que en el ámbito español entendemos que debería ser una norma con rango de ley aprobada por los representantes democráticamente elegidos, no un simple reglamento, por aplicación del art. 18.4 de la Constitución española (“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.”). No consideramos que la reserva de ley deba ser entendida como una habilitación específica para cada uso de la IA en el sector público (en la misma línea, TORCHIA, 115), sino como una reserva de ley relativa (similar a la que establece la Constitución española para el ejercicio de la potestad organizativa, que no supone necesidad de una ley en la creación de cada órgano concreto<sup>27</sup>).

En definitiva, y por lo que se refiere al papel de la ley en el ámbito de las decisiones automatizadas, podríamos, decir, parafraseando la conocida expresión acuñada en la Revolución norteamericana: *No Automation Without Representation*. Reserva de ley que impone una regulación marco con rango de ley con la suficiente densidad, evitando remisiones en blanco a reglamentos, que establezca los límites (incluyendo la posible reserva de humanidad que aquí se propugna respecto a la total automatización de la discrecionalidad), el procedimiento, la competencia y los límites en relación con el uso de la IA.

En cuanto a las regulaciones nacionales referidas a esa reserva de humanidad, diversas leyes prohíben ya hoy en día el uso de IA, salvo que se trate del ejercicio de potestades regladas. Así, la legislación alemana (*Verwaltungsverfahrensgesetz, VwVfG, § 35<sup>a</sup> Vollständig automatisierter Erlass eines Verwaltungsaktes*, la traducción es nuestra): “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”, 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”. En la misma línea en España, si bien con mayor flexibilidad, y sin referencia específica a autorización por norma con rango de ley, aludido en el art. 18 de la Constitución española ya mencionado) se sitúa el art. XVIII.6.d de la Carta española de derechos digitales, norma de *soft law* aprobada en 2021 por el gobierno español (“Que la adopción de decisiones discrecionales quede reservada a personas, salvo que normativamente se prevea la adopción de decisiones automatizadas con garantías adecuadas”)<sup>28</sup>.

---

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”.

<sup>27</sup> Art. 103.2 CE: “Los órganos de la Administración del Estado son creados, regidos y coordinados de acuerdo con la ley”

<sup>28</sup> [140721-Carta Derechos Digitales RedEs.pdf \(lamoncloa.gob.es\)](#)



La lógica detrás de estas regulaciones, creemos, debe hallarse en las ya mencionadas limitaciones intrínsecas de la IA en relación la discrecionalidad, esto es, cuando exista un margen de apreciación, una necesaria elección basada en elementos no jurídicamente predeterminados que incorporan una valoración, un juicio volitivo, que debe basarse con la empatía y la sabiduría humanas, por decirlo en los términos ya referidos de Weizenbaum<sup>29</sup>. En este sentido, creemos que la legislación estatal debería incorporar un límite semejante, por ejemplo, en el actual art. 41 LRJSP. Cuestión diferente es la aplicación automatizada de potestades regladas sin ningún tipo de margen de apreciación o el uso como apoyo de la IA en la adopción de decisiones discrecionales por humanos, posibles jurídicamente.

En el nivel europeo, el RUEIA, en su art. 5, toma la decisión de excluir diversos supuestos del uso de la IA, entre ellos, salvo autorización específica, ciertos sistemas de identificación biométrica, pero no utiliza la distinción reglado-discrecional para prohibir el ejercicio totalmente automatizado de potestades administrativas discrecionales ni siquiera en los supuestos de alto riesgo de uso de la IA por administraciones, supuestos en los que creemos que debería evitarse que una máquina adopte decisiones que incidan en personas, por los motivos expuestos.

#### 4. VIGILANCIA 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 mencionado art. 22 RGPD (en caso de que se exceptione la reserva de humanidad), 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) y legislación española, como el ya citado art. 41.2 de la ley 40/2015, de 1 de octubre, por ejemplo.

En las Directrices éticas para una IA fiable, de 2019, elaboradas por el Grupo Independiente de Expertos de Alto Nivel sobre IA creado por la Comisión Europea se desarrollaron siete principios éticos no vinculantes para la IA, que tienen por objeto contribuir a garantizar la fiabilidad y el fundamento ético de la IA, entre ellos el de acción y supervisión humanas. De acuerdo con las directrices del Grupo de Expertos, por «acción y supervisión humanas» se entiende que los sistemas de IA se desarrollan y utilizan como herramienta al servicio de las personas, que respeta la dignidad humana y la autonomía personal, y que funciona de manera que pueda ser controlada y vigilada adecuadamente por seres humanos (véase ahora el Considerando 27 del RUEIA).

En aplicación de lo expuesto, el ya citado art. 14 del RUEIA se refiere ahora precisamente a esta cuestión, bajo el título “Vigilancia humana” y respecto, eso sí,

---

<sup>29</sup> Sobre esta cuestión, tempranamente, el clásico trabajo de MARTÍN DELGADO, I (2009)

sólo a los usos de IA calificados como de alto riesgo (los detallados en el art. 6, en conexión con los Anexos I y III), no existiendo dicha obligación jurídica de vigilancia, pues, en caso de inexistencia de ese tipo de riesgo<sup>30</sup>. Las líneas maestras de este precepto del RUEIA, que hay que interpretar sistemáticamente con otros, son las siguientes:

Para qué se vigila a los sistemas de IA de alto riesgo:

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 se utiliza un sistema de IA de alto riesgo conforme a su finalidad prevista o cuando se le da un uso indebido razonablemente previsible<sup>31</sup>, en particular cuando dichos riesgos persistan a pesar de la aplicación de otros requisitos establecidos en el RUEIA.

*Quien vigila:*

— La vigilancia debe llevarse a cabo por el responsable del despliegue, bajo las indicaciones del proveedor.

El primero es definido en el art. 3.4 RUEIA como una persona física o jurídica o autoridad, órgano u organismo de otra índole públicos que utilice un sistema de IA bajo su propia autoridad, salvo cuando su uso se enmarque en una actividad personal de carácter no profesional. Por su parte, el art. 3. define al proveedor como una persona física o jurídica o autoridad, órgano u organismo de otra índole públicos que desarrolle un sistema de IA o un modelo de IA de uso general o para el que se desarrolle un sistema de IA o un modelo de IA de uso general y lo introduzca en el mercado o ponga en servicio el sistema de IA con su propio nombre o marca comercial, previo pago o gratuitamente.

Si bien quien debe supervisar es quien usa el sistema, el responsable del despliegue, el proveedor tiene determinadas obligaciones jurídicas al respecto también. Así, los sistemas de IA deben ser diseñados y desarrollados de modo que puedan ser vigilados de manera efectiva por personas físicas durante su período de uso, lo que incluye dotarlos de herramientas de interfaz humano-máquina adecuados. Dicha supervisión del responsable del despliegue del sistema de alto riesgo de IA debe

---

<sup>30</sup> Sobre la propuesta realizada en 2021 por la Comisión, el Consejo formuló en 2022 su posición común, con muy pocos cambios relevantes en este aspecto, mientras que el Parlamento europeo aprobó algunas enmiendas al texto en junio de 2023 pero que en relación con este art. 14 no introducían grandes cambios, insistiendo sobre todo en el conocimiento de la IA por parte de la persona que supervise. Analizamos aquí el texto final del art. 14 del RUEIA aprobado por el Parlamento europeo el 13 de marzo de 2024.

<sup>31</sup> Un uso indebido razonablemente previsible es, de conformidad con el art. 3.13 del RUEIA, un supuesto en el que se produce la utilización de un sistema de IA de un modo que no corresponde a su finalidad prevista, pero que puede derivarse de un comportamiento humano o una interacción con otros sistemas, incluidos otros sistemas de IA, razonablemente previsible.

realizarse de conformidad con las especificaciones que el proveedor incluya en sus instrucciones de uso (art. 13, especialmente apartado d)<sup>32</sup>.

— Los seres humanos que vigilen a la IA deben tener determinadas capacidades para poder realizar bien su labor, pues en las palabras del art. 14 deben estar en condiciones de:

a) entender adecuadamente las capacidades y limitaciones pertinentes del sistema de IA de alto riesgo y poder vigilar debidamente su funcionamiento, por ejemplo,

---

<sup>32</sup> Considerando 72 del RUEIA: “A fin de abordar las preocupaciones relacionadas con la opacidad y complejidad de determinados sistemas de IA y ayudar a los responsables del despliegue a cumplir sus obligaciones en virtud del presente Reglamento, debe exigirse transparencia respecto de los sistemas de IA de alto riesgo antes de su introducción en el mercado o su puesta en servicio. Los sistemas de IA de alto riesgo deben diseñarse de modo que permitan a los responsables del despliegue comprender la manera en que el sistema de IA funciona, evaluar su funcionalidad y comprender sus fortalezas y limitaciones. Los sistemas de IA de alto riesgo deben ir acompañados de la información adecuada en forma de instrucciones de uso. Dicha información debe incluir las características, las capacidades y las limitaciones del funcionamiento del sistema de IA. Estos elementos abarcarían la información sobre las posibles circunstancias conocidas o previsibles relacionadas con el uso del sistema de IA de alto riesgo, incluida la actuación del responsable del despliegue capaz de influir en el comportamiento y el funcionamiento del sistema, en cuyo marco el sistema de IA puede dar lugar a riesgos para la salud, la seguridad y los derechos fundamentales, sobre los cambios que el proveedor haya predeterminado y evaluado para comprobar su conformidad y sobre las medidas pertinentes de supervisión humana, incluidas las medidas para facilitar la interpretación de la información de salida del sistema de IA por parte de los responsables del despliegue. La transparencia, incluidas las instrucciones de uso que acompañan a los sistemas de IA, debe ayudar a los responsables del despliegue a utilizar el sistema y tomar decisiones con conocimiento de causa. Entre otros, los responsables del despliegue deben estar en mejores condiciones para elegir correctamente el sistema que pretenden utilizar a la luz de las obligaciones que les son aplicables, estar informados sobre los usos previstos y excluidos y utilizar el sistema de IA correctamente y según proceda. A fin de mejorar la legibilidad y la accesibilidad de la información incluida en las instrucciones de uso, cuando proceda, deben incluirse ejemplos ilustrativos, por ejemplo sobre las limitaciones y sobre los usos previstos y excluidos del sistema de IA. Los proveedores deben garantizar que toda la documentación, incluidas las instrucciones de uso, contenga información significativa, exhaustiva, accesible y comprensible, que tenga en cuenta las necesidades y los conocimientos previsibles de los responsables del despliegue destinatarios. Las instrucciones de uso deben estar disponibles en una lengua fácilmente comprensible para los responsables del despliegue destinatarios, según lo que decida el Estado miembro de que se trate”

Considerando 73 del RUEIA: “Los sistemas de IA de alto riesgo deben diseñarse y desarrollarse de tal modo que las personas físicas puedan supervisar su funcionamiento, así como asegurarse de que se usan según lo previsto y de que sus repercusiones se abordan a lo largo del ciclo de vida del sistema. A tal fin, el proveedor del sistema debe definir las medidas adecuadas de supervisión humana antes de su introducción en el mercado o puesta en servicio. Cuando proceda, dichas medidas deben garantizar, en concreto, que el sistema esté sujeto a limitaciones operativas incorporadas en el propio sistema que este no pueda desactivar, que responda al operador humano y que las personas físicas a quienes se haya encomendado la supervisión humana posean las competencias, la formación y la autoridad necesarias para desempeñar esa función. También es esencial, según proceda, garantizar que los sistemas de IA de alto riesgo incluyan mecanismos destinados a orientar e informar a las personas físicas a las que se haya asignado la supervisión humana para que tomen decisiones con conocimiento de causa acerca de si intervenir, cuándo hacerlo y de qué manera, a fin de evitar consecuencias negativas o riesgos, o de detener el sistema si no funciona según lo previsto.”

con vistas a detectar y resolver anomalías, problemas de funcionamiento y comportamientos inesperados;

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, por ejemplo, los métodos y herramientas 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 de este género;

e) intervenir en el funcionamiento del sistema de IA de alto riesgo o interrumpir el sistema pulsando un botón de parada o mediante un procedimiento similar que permita que el sistema se detenga de forma segura.

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, lo que debe relacionarse con el art. 4 del RUEIA, que señala que los proveedores y responsables del despliegue de sistemas de IA adoptarán medidas para garantizar que, en la mayor medida posible, su personal y demás personas que se encarguen en su nombre del funcionamiento y la utilización de sistemas de IA tengan un nivel suficiente de alfabetización en materia de IA, teniendo en cuenta sus conocimientos técnicos, su experiencia, su educación y su formación, así como el contexto previsto de uso de los sistemas de IA y las personas o los grupos de personas en que se utilizarán dichos sistemas<sup>33</sup>. Asimismo, en el art. 26.2 el RUEIA insiste en que se encomendará la supervisión humana a personas físicas que tengan la competencia, la formación y la autoridad necesarias.

---

<sup>33</sup> Considerando 91 del RUEIA: “Habida cuenta de las características de los sistemas de IA y de los riesgos que su uso puede conllevar para la seguridad y los derechos fundamentales, también en lo que respecta a la necesidad de garantizar la correcta vigilancia del funcionamiento de un sistema de IA en un entorno real, conviene definir las responsabilidades específicas de los responsables del despliegue. En particular, los responsables del despliegue deben adoptar las medidas técnicas y organizativas adecuadas para garantizar que utilizan los sistemas de IA de alto riesgo conforme a las instrucciones de uso. Además, es preciso definir otras obligaciones en relación con la vigilancia del funcionamiento de los sistemas de IA y la conservación de registros, según proceda. Asimismo, los responsables del despliegue deben garantizar que las personas encargadas de poner en práctica las instrucciones de uso y la supervisión humana establecidas en el presente Reglamento tengan las competencias necesarias, en particular un nivel adecuado de alfabetización, formación y autoridad en materia de IA para desempeñar adecuadamente dichas tareas. Estas obligaciones deben entenderse sin perjuicio de otras obligaciones que tenga el responsable del despliegue en relación con los sistemas de IA de alto riesgo con arreglo al Derecho nacional o de la Unión”

### Cuándo se vigila:

El art. 14 RUEIA quiere que exista vigilancia humana en todo el ciclo de funcionamiento de los sistemas de IA, incluyendo el servicio poscomercial, entendido como todas las actividades realizadas por los proveedores de sistemas de IA destinadas a recoger y examinar la experiencia obtenida con el uso de sistemas de IA que introducen en el mercado o ponen en servicio, con objeto de detectar la posible necesidad de aplicar inmediatamente cualquier tipo de medida correctora o preventiva que resulte necesaria (art. 3. 25 RUEIA). Por ello, los responsables del despliegue deben informar a los proveedores de incidencias (art. 26.5, 72, suspendiendo el sistema si existe un riesgo en el sentido del art. 79.1) y en caso de incidencias graves también al importador o distribuidor y a la autoridad de vigilancia del mercado (art. 73, si bien no abarca a los datos operativos sensibles cuando quien despliega el sistema sea una autoridad encargada de aplicar la ley).

### Cómo se vigila:

- Las medidas a adoptar pueden ser de uno o dos de estos tipos:
- Las definidas e integradas por el proveedor cuando sea técnicamente viable en el sistema de IA de alto riesgo antes de su introducción en el mercado o su puesta en servicio,
- Las definidas por el proveedor y que sean adecuadas para que las ponga en práctica el responsable del despliegue.
- Recuérdese el papel ya apuntado de las instrucciones de uso del proveedor al respecto (art. 13). Antes del primer uso de un sistema de IA de alto riesgo, el responsable del despliegue debe realizar una evaluación de impacto relativa a los derechos fundamentales, diferente y concurrente con la evaluación de impacto relativa a la protección de datos, indicando en la misma que medidas de supervisión humana prevé de acuerdo con las instrucciones de uso del proveedor (art. 27.1e RUEIA).
- Para el caso de los sistemas de alto riesgo mencionados en el punto 1, letra a del Anexo III del RUEIA, esto es, sistemas de identificación biométrica remota, con exclusión de los utilizados con fines de verificación cuya única finalidad sea confirmar que una persona física concreta es la persona que afirma ser (y sólo para éste supuesto del Anexo III, no para otros casos que pueden llegar a ser igual de sensibles, por cierto), se aplica la técnica de los *cuatro ojos*: una doble supervisión humana, consistente en que las medidas de vigilancia garantizarán que no se actúe ni se tome ninguna decisión basándose en la identificación generada por el sistema, salvo si al menos dos personas físicas con la competencia, formación y autoridad necesaria verifican y confirman por separado dicha identificación (este requisito podrá no aplicarse si el Derecho nacional o de la UE considera que es desproporcionado respecto

a sistemas de IA empleados con fines de aplicación de la ley, de migración, de control fronterizo o de asilo)<sup>34</sup>.

— Conforme a lo que desde hace tiempo se viene señalando (así, recuérdense las *guidelines* de *Article 29 Working Party* en relación con el RGPD<sup>35</sup>) la supervisión debe ser significativa (*meaningful*) esto es, debe ser capaz de intervenir en el sistema, decidiendo, 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. Incluso más aún, como vimos: el vigilante debe poder interrumpir el sistema pulsando un botón de parada o mediante un procedimiento similar que permita que el sistema se detenga de forma segura, mediante lo que podríamos definir como el *botón del pánico*: la capacidad de activar un mecanismo de parada de un sistema que funcione deficientemente.

Qué grado de intensidad en la vigilancia:

— No se establece por el RUEIA un estándar determinado de diligencia en la supervisión de la IA, indicándose que las medidas que se adopten deben ser proporcionales a tres variables: los riesgos, el nivel de autonomía del sistema de IA y el contexto del uso del sistema.

En lo que se refiere a la Administración, cuando ésta sea responsable del despliegue y por tanto de organizar la vigilancia humana, 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 © 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 administrativa (PONCE, 2023b).

<sup>34</sup> Considerando 73 del RUEIA: “Teniendo en cuenta las enormes consecuencias para las personas en caso de una correspondencia incorrecta efectuada por determinados sistemas de identificación biométrica, conviene establecer un requisito de supervisión humana reforzada para dichos sistemas, de modo que el responsable del despliegue no pueda actuar ni tomar ninguna decisión basándose en la identificación generada por el sistema, salvo si al menos dos personas físicas la han verificado y confirmado por separado. Dichas personas podrían proceder de una o varias entidades e incluir a la persona que maneja o utiliza el sistema. Este requisito no debe suponer una carga ni retrasos innecesarios y podría bastar con que las verificaciones que las distintas personas efectúen por separado se registren automáticamente en los registros generados por el sistema. Dadas las especificidades de los ámbitos de la aplicación de la ley, la migración, el control fronterizo y el asilo, este requisito no debe aplicarse cuando el Derecho nacional o de la Unión considere que su aplicación es desproporcionada.”

<sup>35</sup> 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 tenga la autoridad y competencia para cambiar la decisión. Como parte del análisis, debe considerar todos los datos de datos de entrada y salida disponibles”.

Finalmente, habrá que esperar a ver, el papel que asumirá la Agencia Española de Supervisión de Inteligencia Artificial al respecto, pues su regulación aprobada en 2023 parece abrirle un campo importante respecto a lo dicho, incluyendo la formación de los supervisores públicos humanos<sup>36</sup>.

## 5. EL CONTROL JUDICIAL DE LAS DECISIONES ADMINISTRATIVAS AUTOMATIZADAS ADOPTADAS CON INTELIGENCIA ARTIFICIAL, CON ESPECIAL ATENCIÓN A LAS DISCRECIONALES.

### 5.1. Control judicial, IA y actividad administrativa

En primer lugar, cabe aclarar que no es objeto de esta sección el papel que la IA vaya a tener, a su vez, en la actividad judicial en general y, en particular, en el control que ejerce la jurisdicción contenciosa de las decisiones administrativas<sup>37</sup> (al respecto, por ejemplo, DE LA SIERRA, 2020). El interés de esta sección se centra, en cambio, en el control por parte de la jurisdicción contenciosa de decisiones administrativas adoptadas de forma automatizada. En todo caso, simplemente quiero apuntar aquí que existen debates abiertos relativos a la posibilidad de un juez o jueza (también en el contencioso-administrativo) “robot” y al papel humano y la discrecionalidad judicial, algunos de los cuales guardan semejanza con los planteados y vistos en el caso del poder ejecutivo, como la posible existencia de una reserva de humanidad también en el ámbito judicial<sup>38</sup>.

---

<sup>36</sup> Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial.

<sup>37</sup> Así, por ejemplo, en el caso argentino la prensa informa de la emisión de miles de sentencias ya con el uso de IA: “La justicia de Río Negro dictó seis mil fallos con inteligencia artificial”, *Infobae*, 2 de octubre de 2023.

<sup>38</sup> Es el caso de la conocida sentencia norteamericana *State v. Loomis*, 881 N.W.2d 749 (2016), del Tribunal Supremo del estado de Wisconsin, se conoció del caso de Eric L. Loomis, quien fue arrestado en Wisconsin por resistirse a una autoridad pública después de ser encontrado conduciendo un automóvil robado. Para determinar la sentencia, el tribunal utilizó COMPAS (*Correctional Offenders Management Profiling for Alternative Sanctions*), un sistema algorítmico capaz de estimar el riesgo de reincidencia del delincuente sobre la base de una entrevista (137 preguntas con respuestas restringidas) así como la inclusión en el programa de datos relacionados con los *antecedentes* penales y las condiciones socioeconómicas y personales del imputado. La sentencia se planteó si hubo un erróneo uso de la discrecionalidad judicial. La conclusión del Tribunal Supremo de Wisconsin fue que no, pues en este caso la discrecionalidad judicial fue ejercida tomando en consideración todos los elementos relevantes del caso, entre ellos COMPAS, como uno más, pero sin substituir la ponderación judicial. Un voto particular recogió alguna de las cuestiones cruciales del caso (traducción nuestra):

“Al dictar sentencia, el tribunal de circuito debe considerar tres factores principales: la gravedad del delito, el carácter del delincuente y la necesidad de proteger al público (...) El ejercicio adecuado de la discreción de un tribunal de circuito para dictar sentencia incluye una sentencia individualizada basada en los hechos del caso y puede incluir la explicación de cómo la sentencia impuesta fomenta los objetivos del tribunal de circuito. Id. (...) Un tribunal sentenciador debe articular los factores que consideró



Así, el Anexo III del RUEIA aprobado por el Parlamento europeo en 2024 señala que entre los sistemas de IA de alto riesgo se encuentran:

“8. Administración de justicia y procesos democráticos:

a) Sistemas de IA destinados a ser utilizados por una autoridad judicial, o en su nombre, para *ayudar* a una autoridad judicial en la investigación e interpretación de hechos y de la ley, así como en la aplicación de la ley a un conjunto concreto de hechos, o a ser utilizados de forma similar en una resolución alternativa de litigios”.

El texto emplea el verbo *ayudar*. La cuestión por dilucidar aquí sería si el uso de este verbo excluye la toma de decisiones judiciales con margen de apreciación de forma totalmente automatizada. En esta línea, en las enmiendas presentadas por el Parlamento Europeo al proyecto de reglamento de IA, se modificó el considerando 40 mediante la enmienda 71, señalando lo siguiente: “La utilización de herramientas de inteligencia artificial puede apoyar la toma de decisiones, pero no debe substituir el poder de toma de decisiones de los jueces o la independencia judicial, puesto que la toma de decisiones finales debe seguir siendo una actividad y una decisión de origen humano”. Las ideas esenciales de esa enmienda han sido finalmente incorporadas con el número 61 al texto definitivamente aprobado por el Parlamento en marzo de 2024, con este texto: “La utilización de herramientas de IA puede apoyar el poder de decisión de los jueces o la independencia judicial, pero no debe sustituirlas: la toma de decisiones finales debe seguir siendo una actividad humana”. Consideramos que las mismas razones que utilizamos en el caso de los decisores del poder ejecutivo tienen aplicación en el ámbito judicial y que los límites explicitados por el RUEIA para éste último deberían extenderse al primero conforme lo ya expuesto<sup>39</sup>.

No cabe duda, en todo caso, que el uso adecuado de la IA en el ámbito judicial puede contribuir a hacer realidad el principio de una buena administración de justicia, afirmado por el TS, el TJUE y el TEDH (PONCE, 2021)<sup>40</sup>. Buena adminis-

---

al dictar la sentencia y cómo afectaron a la sentencia que impuso. (...) Es a través de esta articulación que determinamos si el tribunal de circuito ejerció adecuadamente su discreción en la sentencia. Los acusados tienen un derecho al debido proceso (*due process*) de no ser sentenciados en base a factores impropios como la raza o el género”.

<sup>39</sup> Por su parte, el Real Decreto-ley 6/2023, de 19 de diciembre, por el que se aprueban medidas urgentes para la ejecución del Plan de Recuperación, Transformación y Resiliencia en materia de servicio público de justicia, función pública, régimen local y mecenazgo, en sus arts. 56 y 57, en una dirección similar, señala que entre las actuaciones automatizadas procesales sin intervención humana no puede hallarse un documento complejo basado en datos y producido por algoritmos que sea una resolución judicial o procesal “sin validación de la autoridad competente”: este documento sólo puede ser un borrador parcial o total, una ayuda para la decisión humana final (“actuación asistida”), el cual ha de poder ser “libremente modificado” y validado como texto definitivo por el juez/a, magistrado/a, fiscal o letrado/a de la Administración de justicia.

<sup>40</sup> El Poder Judicial debe, a su vez, desarrollarse también conforme a las exigencias de *una buena administración de justicia*, a la que se refiere la Ley Orgánica 6/1985, de 1 de julio, del Poder Judicial, en distintos artículos (81.3, 152.1. 3º, 165, 167.1, 268, 269, 292, 293, 296, pero sin definir el concepto). La misma es citada no infrecuentemente también por nuestro Tribunal Supremo (por todas, STS, sala de lo contencioso, de 11 de junio de 2020, aludiendo al “principio de una buena administración de justici-

tración de justicia que, a su vez, en analogía con lo ya expuesto sobre las decisiones administrativas, exigiría una ponderación cuidadosa de todos los elementos relevantes por parte del juez o jueza (ser humano) en ejercicio de los posibles márgenes de apreciación.

Entrando ahora ya en el objeto de nuestra atención, el control judicial de decisiones administrativas, cabe señalar que existen ya decisiones judiciales en todo el mundo controlando la toma de decisiones automatizadas<sup>41</sup>. En el caso español, la primera litigación asociada al uso de sistemas algorítmicos ha llegado en relación al código fuente del programa BOSCO de otorgamiento del bono social previsto en la legislación energética, centrándose en la necesidad jurídica de conceder el acceso al mismo cuando es solicitado ejerciéndose el derecho de acceso a la información pública. Se trata de la sentencia del Juzgado Central de lo Contencioso Administrativo núm. 8 núm. 143/2021, de 30 de diciembre de 2021, que ha denegado dicho acceso y se encuentra en fase de recurso al escribir estas líneas<sup>42</sup>.

---

cia”), por el TJUE (en numerosas sentencias como la STJUE de 6 de octubre de 2009, asunto C-40/08, *Asturcom Telecomunicaciones, S.L.*, y *Cristina Rodríguez Nogueira*, apartado 36) y por el TEDH, quien ha declarado que el derecho de acceso a un tribunal se ve menoscabado cuando las normas dejan de servir a los objetivos de seguridad jurídica y de buena administración de justicia y constituyen una especie de barrera que impide al litigante que el tribunal competente se pronuncie sobre el fondo de su asunto (véase la STEDH 14 de septiembre de 2021, caso de *Inmovilizados y Gestiones S.L. v. Spain*, asunto en el que España se vio condenada a indemnizar con 16.600 euros a una empresa por una disparidad de criterios de la sala de lo contencioso del Tribunal Supremo ante cinco casos idénticos).

<sup>41</sup> Por ejemplo, en el caso australiano de *Robodebt*, determinadas tareas administrativas desarrolladas por humanos fueron automatizadas en el ámbito social. El resultado de los errores de la automatización ha sido 400,000 beneficiarios de asistencia social en Australia que fueron acusados erróneamente de informar de manera equivocada sus ingresos a la agencia de bienestar y, en consecuencia, fueron multados. Desde julio de 2015 hasta noviembre de 2019, el sistema utilizó algoritmos para calcular los pagos excesivos, recaudando más de 1.2 mil millones de dólares, que el gobierno se verá obligado a devolver, como consecuencia de que un tribunal decidió que el esquema era ilegal en 2019.

En Francia, la sentencia de 7 de diciembre de 2022 del Tribunal de lo Contencioso-Administrativo de París anuló una decisión de asignar al alumno a un centro de enseñanza secundaria que solo cumplía la sexta opción registrada para el alumno en la solicitud *Affelnet*, debido a que los métodos de cálculo de la escala de puntos de dicho alumno eran incorrectos y que esos «errores materiales» influían en el sentido del dictamen emitido por la comisión preparatoria para el encargo y, en consecuencia, sobre la decisión de asignación tomada por el director académico de los Servicios Nacionales de Educación.

En Italia, el Consejo de Estado italiano, se ha pronunciado en diversas ocasiones sobre el uso de IA por la Administración en Italia, por ejemplo, en su decisión de 13 de diciembre de 2019, No 8472 o en la Sentencia de 4.2.2020, N° 881. Más recientemente, la decisión judicial italiana del Tribunal Administrativo Regional de Campania de 14 de noviembre de 2022 ha anulado una decisión administrativa adoptada con uso de algoritmos y con mención también al art. 41 de la CDFUE y la buena administración.

<sup>42</sup> La sentencia ha sido objeto de apelación no resuelta en el momento de escribir estas líneas. Puede consultarse un comentario a la misma del autor de estas líneas en: [El bono social energético y el programa Bosco: sobre algoritmos, errores y código fuente. A propósito de la primera decisión judicial recaída en 2021: una mala sentencia que esperamos sea corregida pronto – Nudging aplicado a la Mejora de la Regulación y al Uso de Algoritmos y de Inteligencia Artificial \(wordpress.com\)](#).

## 5.2. Adaptación del control judicial del ejercicio de potestades por humanos a su posible ejercicio por máquinas: ¿Indiferencia insuficiencia o deferencia judicial?

### 5.2.1. *Técnicas jurídicas para el control judicial del ejercicio de potestades administrativas mediante IA. La relevancia del procedimiento administrativo complejo y la invalidez de las decisiones administrativas.*

Más allá de decisiones específicas ya recaídas, nuestra reflexión se refiere al ejercicio de la función judicial de lo contencioso-administrativo de control jurídico de las decisiones administrativas adoptadas con uso de IA. En este sentido, el control judicial clásico, mediante la revisión de los elementos reglados de las decisiones y del respeto de los principios generales del Derecho<sup>43</sup>, debería adaptarse, como argumentaremos, a las nuevas perspectivas abiertas por la IA. En lo que sigue, se utiliza un orden sistemático basado en dicho control clásico, analizando primero los elementos reglados para terminar con el control mediante los principios generales del Derecho.

Así, respecto de los *elementos reglados*, el control judicial deberá tener en cuenta la legislación vigente en materia de IA, tanto la estatal de aplicación a las CCAA y entes locales<sup>44</sup>, como la autonómica que exista (dado que el poder legislativo autonómico dispone de competencias para mejorar el sistema actualmente vigente, derivadas de sus competencias en materia de desarrollo de las bases estatales de régimen jurídico y del procedimiento administrativo común), así como el RUEAI.

Asimismo, la Carta española de Derechos digitales de 2021, aunque es *soft law*, tiene también efectos jurídicos a tener en cuenta en el control judicial<sup>45</sup>. Téngase en

<sup>43</sup> Como señala la STS de 9 de julio de 2021, ECLI: ES:TS:2001:5939:

“Al menos, desde la década de los sesenta forman parte de nuestra jurisprudencia, como mecanismos ordinarios de control judicial de la discrecionalidad administrativa: los presupuestos de hecho, los elementos reglados (existencia de la potestad y su extensión, la competencia, el procedimiento y el contenido parcialmente reglado, en su caso), la desviación de poder, los conceptos jurídicos indeterminados y los principios generales del derecho”

<sup>44</sup> Recuérdese, art. 41 LRJSP, arts. equivalentes en los sectores tributarios y de seguridad social, ley estatal Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación, cuyo art. 23 ya aludido se considera dictado al amparo del art. 149.1.1 CE, de acuerdo con la Disposición Final octava de esta ley y legislación de protección de datos.

<sup>45</sup> En este sentido, la jurisprudencia del TS se ha referido a los efectos interpretativos que despliega este Derecho débil o dúctil, como en la STS de 27 de marzo de 2023, ECLI:ES:TS:2023:990, que alude a sentencias anteriores:

“En el auto de 8 de marzo de 2017 y en la ulterior sentencia de 31 de mayo de 2017 (recurso 88/2017), hacíamos mención al «soft law». Tradicionalmente en el ámbito del Derecho Internacional Público, y más recientemente en Derecho Comunitario (sentencia de 13 de diciembre de 1989, asunto 322/1988, *Grimaldi*), actualmente de la Unión Europea se han venido reputando como «soft law» las recomendaciones, por su carencia de efectos jurídicos vinculantes. Carecen de naturaleza imperativa. No obstante, lo cual afirmó el Tribunal de Justicia que «las recomendaciones» no están privadas de efectos jurídicos desde el momento en que los jueces nacionales han de tenerlas en consideración para resolver las controversias sirviendo como criterio interpretativo”

cuenta, que esta Carta puede pasar a tener la vinculación propia del Derecho si una norma jurídica la incorpora, cosa que ya se ha hecho en el caso de Extremadura<sup>46</sup>.

En cuanto a la *atribución normativa de potestad y competencia*, el juez o la jueza del contencioso debe comprobar, en primer lugar, si es aplicable una reserva de humanidad, bien esté establecida explícitamente por el ordenamiento (como es el caso del actual art. 44 de la ley catalana 26/2010, ya mencionado), bien derivada del principio constitucional de buena administración, conforme a lo expuesto más arriba, lo que impediría el uso de IA para la adopción de decisiones automatizadas discrecionales, tanto mediante IA de reglas predeterminadas (por rigidificación ilegal de la discrecionalidad, como dijimos) como mediante IA estadística, esto es aprendizaje automático o profundo (por ausencia de empatía y de razonamiento abductivo, como se expuso).

Recuérdese asimismo la Carta de Derechos Digitales, que exige una norma específica para que se pueda ejercer la discrecionalidad de forma automatizada. Si el juez o la jueza o Tribunal cree que la regulación legal existente (art. 41.1 LRJSP) no cubre la reserva de ley del art.18.4 CE podría presentar una cuestión de inconstitucionalidad. En todo caso, véase más abajo lo que se indica sobre posible violación del principio de buena administración por el ejercicio automatizado de discrecionalidad.

Asimismo, caso de no existir reserva de humanidad, en cuanto a la decisión finalmente adoptada, debe en todo caso controlarse judicialmente, como es habitual, la atribución de potestad y competencia al órgano específico para adoptar la decisión, que la habrá asumido mediante la firma de esta, como indica MARTÍN DELGADO (2009).

En cuanto al *ejercicio de la competencia con imparcialidad y objetividad*, el juez o la jueza puede detectar sesgos cognitivos de los programadores traspasados al sistema, sesgos contenidos en los datos manejados que han afectado a la decisión, sesgos de la persona supervisora del sistema o sesgos de quien decide (si la decisión no fue automatizada), ya descritos anteriormente. Si existen dichos sesgos, se podrá judicialmente declarar bien una violación de la imparcialidad de la Administración, bien de su objetividad, al tener en cuenta elementos irrelevantes y, por tanto, de la buena administración.

---

<sup>46</sup> Decreto-ley 2/2023, de 8 de marzo, de medidas urgentes de impulso a la inteligencia artificial en Extremadura:

“Artículo 11. Sistemas de inteligencia artificial en la toma de decisiones.

1. La Administración pública autonómica podrá adoptar actos administrativos mediante sistemas de inteligencia artificial en el marco de un procedimiento administrativo, de acuerdo con los Derechos digitales de la ciudadanía en sus relaciones con las Administraciones Públicas, descritos en la Carta de Derechos Digitales del Gobierno de España y la Declaración Europea sobre los Derechos y Principios Digitales para la Década Digital (2023/C 23/01)”.

Recuérdese, art. 41 LRJSP, arts. equivalentes en los sectores tributarios y de seguridad social, ley estatal Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación, cuyo art. 23 ya aludido se considera dictado al amparo del art. 149.1.1 CE, de acuerdo con la Disposición Final octava de esta ley y legislación de protección de datos.

Asimismo, es de gran relevancia el *control procedimental*. En concreto, en cuanto al procedimiento administrativo como relevante institución para la adopción de adecuadas decisiones automatizadas, hay que distinguir entre:

a) el procedimiento de evaluación, aprobación, supervisión y revisión de los algoritmos, de acuerdo con las exiguas exigencias actuales del art. 41.2 LRJSP<sup>47</sup>, algoritmos que serán empleados en muchos procedimientos específicos sectoriales posteriores, desarrollados de forma asíncrona, formando el primer procedimiento administrativo y los segundos un procedimiento complejo asíncrono.

b) a la evaluación en cada uno de estos posteriores procedimientos específicos de la oportunidad y necesidad del uso de IA y de su supervisión humana ya aludida.

Efectivamente, de acuerdo con lo expuesto es importante retener que en la adopción de decisiones automatizadas estamos ante dos procedimientos distintos pero conexos, configurándose así un procedimiento complejo, formado por estos dos procedimientos administrativos integrados (cada uno de los cuales concluye con su propia resolución) que se concatenan entre sí para lograr, en suma, una finalidad de interés público que no sería posible sin esa integración. Procedimiento complejo (en este caso para la adopción de decisiones automatizadas) que coloquialmente es conocido como *procedimiento de procedimientos* (GONZÁLEZ NAVARRO, 1981, RAE *et al.*, 2023).

Sobre el procedimiento administrativo inicial de preparación del uso del sistema algorítmico, el control judicial deberá comprobar la realización de las exiguas exigencias legales actuales, contenidas en el art. 41.2 LRJSP, insuficientes en nuestra opinión. La regulación actual no exige específicamente una ponderación de aspectos relevantes para decidir si automatizar o no el ejercicio de una potestad, considerando costes y beneficios, si bien el juez o la jueza del contencioso podría utilizar diversos principios constitucionales para controlar este momento procedimental, como enseguida argumentaremos, e incluso deducir del art. 23 de la Ley estatal 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación, la necesidad implícita de la existencia de dicha valoración, al menos en cuanto a los posibles sesgos<sup>48</sup>.

<sup>47</sup> En el caso español, como es sabido existe una norma con rango de ley que alude a la automatización, el art. 41 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público. Sin embargo, su densidad normativa es muy ligera y no incluye el uso de la IA en los procedimientos específicos sectoriales tras la aprobación de las especificaciones generales, por lo que creemos que existen dudas sobre el correcto cumplimiento de la reserva de ley del art. 18 CE a la que aludimos más arriba. Su texto es el siguiente:

“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.”

<sup>48</sup> “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 favorece-

En concreto, en relación con la primera fase del procedimiento complejo descrito, el control judicial podría articularse como sigue:

— El control judicial debe garantizar el cumplimiento del art. 41.2 LRJSP (previo establecimiento “del 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”). Dicho cumplimiento no debe ser puramente formal, como parece ocurrir en ocasiones en la realidad, sino comprobando la existencia real de todas esas exigencias y su cumplimiento diligente<sup>49</sup>.

— El juez o la jueza debe tener en cuenta, además, el art. 35 del RGPD y comprobar el cumplimiento de la evaluación de impacto relativa a la protección de datos, cuando sea preciso.

— Sobre el análisis a desarrollar con la diligencia debida y el debido cuidado propio del derecho a una buena administración de los costes y beneficios existentes en el empleo de los sistemas de IA, nada exige actualmente la legislación española vigente, de forma explícita, pero cabría preguntarse si es posible jurídicamente tomar la decisión de usar IA sin esa previa ponderación, lo que podría considerarse irracional y contrario a la buena administración, que, recordemos, es un principio general constitucional conforme afirma reiteradamente la jurisprudencia del TS, complementando así el ordenamiento jurídico (art. 1.6 del Código Civil). Enseguida volveremos sobre el tema.

— Esta ponderación, no sólo referida, pues, a la protección de datos, deberá respetar el principio de proporcionalidad, que puede ser violado por la generación de costes o riesgos excesivos, por ejemplo, que afecten al derecho a la intimidad, aunque el sistema sea meramente de apoyo. El conocido caso holandés *SyRI* que dio lugar a la sentencia del Tribunal de Distrito de La Haya de 13 de febrero de 2020 supone un ejemplo del uso de este principio (COTINO, 2020). Véase más abajo lo que se indica sobre el principio de proporcionalidad como canon de control.

— Respecto a la decisión sobre el papel del control humano del sistema, el control judicial debería comprobar no sólo si se ha nombrado a una persona supervisora (cumpliendo la obligación jurídica establecida en el art. 41 LRJSP al respecto)

---

rán 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.”

<sup>49</sup> En aplicación del art. 41 LRJSP, pueden detectarse publicaciones de decisiones administrativas que no incluyen, por ejemplo, las especificaciones del sistema. Véanse los siguientes ejemplos: para el nivel estatal: [Disposición 2076 del BOE núm. 52 de 2016](#) [Resolucion nuevas aplicac informaticas.pdf](#) ([agenciatributaria.gob.es](#)); para el autonómico: [BOJA22-229-00002-19405-01\\_00272609.pdf](#) ([juntadeandalucia.es](#)); para el local: [0\\_4535\\_1.pdf](#) ([teruel.es](#))

sino si la persona y el sistema de supervisión son adecuados: si el humano tiene el suficiente conocimiento técnico y es capaz de supervisar el sistema de forma significativa, como vimos (aspecto sobre el que el art. 41 nada dice en la actualidad). Cabría plantear si la mera posibilidad de presentar un recurso administrativo o judicial ante un humano contra la decisión automatizada adoptada satisface la necesidad de supervisión humana. Creemos insuficiente una supervisión humana sólo *ex post*, aunque sea necesaria, si no existe también una supervisión *ex ante*, durante el procedimiento administrativo de elaboración de la específica decisión.

— En cuanto a la participación ciudadana, entendemos que, en el estado actual de la legislación, el juez o la jueza debería comprobar el adecuado seguimiento bien del procedimiento reglamentario, si se acepta que el sistema algorítmico supone la existencia de un reglamento, bien, en otro caso, del art. 83 de la LPAC<sup>50</sup> y comprobar, siempre, que la Administración ha cumplido su obligación de apertura de información pública cuando traduce a código la norma y prepara los algoritmos en este primer procedimiento, al concurrir las circunstancias legales descritas en la LPAC, que respecto a actos habla de cuando su naturaleza “lo requiera”, como sucede en este supuesto, puesto que el sistema de IA afectará en el futuro a un conjunto indeterminado de personas.

Asimismo, la decisión adoptada en este primer procedimiento, que debe incluir, como exige el art. 41 LRJSP “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”, no podrá ser eficaz si no es publicada (bien bajo su consideración de reglamento, bien bajo su entendimiento como acto, en este segundo caso, de acuerdo con el art. 45 LPAC).

Entrando ahora en el *segundo tipo de procedimiento administrativo*, se trata del procedimiento administrativo conexo de adopción de la concreta decisión automatizada, que va precedido del previo procedimiento acabado de analizar para el cumplimiento de las exigencias del art. 41.2 LRJSP, del que depende, formando ambos el procedimiento complejo expuesto, con dos decisiones independientes pero íntimamente vinculadas. Aquí el juez o la jueza debe realizar el control habitual sobre el respeto del procedimiento específico establecido sectorialmente. Reviste especial importancia *la audiencia a los interesados*, que en caso de tener que otorgarse, lo que es regla general, conforme al art. 82 LPAC, implica la suspensión de la automatización completa del procedimiento y la introducción de una necesaria ponderación humana de las alegaciones, como se expuso.

En cuanto a este segundo tipo de procedimiento y en relación con la determinación administrativa de los *hechos determinantes*, los datos manejados por el sistema deberán ser suficientemente representativos, reflejar la realidad, ser fiables y ciertos y

---

<sup>50</sup> Desde un punto de vista jurídico, se ha planteado en la doctrina española reciente la naturaleza jurídica de los algoritmos, sosteniéndose su naturaleza reglamentaria (BOIX, 2019) o negándola (como en el caso de, por ejemplo, HUERGO, 2023).



estar exentos de sesgos. Asimismo, la detección de errores de programación también será tarea judicial.

El control judicial deberá analizar también el *elemento causal* de la decisión automatizada concreta que se adopte mediante el procedimiento administrativo. Si se adopta con IA conexionista, esto es aprendizaje automático o profundo u otras técnicas que puedan surgir, deberá controlar que la correlación establecida por el sistema no sea incorrecta, por ilusoria.

En cuanto a la *motivación*, el art. 86 RUEIA, el art. 9.3 CE y el art. 35 de la LPAC, como ya dijimos antes, exige la motivación de las decisiones administrativas adoptadas con el uso de IA. Motivación que deberá ser expresada en lenguaje natural comprensible para cualquier persona, puesto que el derecho a una buena administración impone el derecho a comprender lo que la Administración le explica al ciudadano. Por supuesto, explicaciones y justificaciones estereotipadas y sin contenido deben ser consideradas como inexistentes<sup>51</sup>.

El juez o la jueza debe comprobar que la motivación de la decisión finalmente adoptada, en caso de que en ésta se haya empleado IA, explique qué impacto ha tenido el uso del sistema algorítmico en ella y demostrar que no se han producido los sesgos cognitivos de anclaje, *ilusión correlacional* o *alucinación estadística* o automatización ya expuestos. Para ello, no debería bastar en ningún caso una motivación *in aliuende*, con remisión al sistema algorítmico, con una referencia similar a “de acuerdo”, como en el caso de la empleada para la intervención del Consejo de Estado, por ejemplo, sino que la motivación debe justificar en qué medida el sistema o el ser humano que decida ha considerado los *inputs* proporcionados y por qué, si hay decisión final humana, se ha aceptado, o no, las sugerencias de la IA en la decisión final.

En cuanto, en segundo lugar, al control de la decisión con IA mediante *los principios generales del Derecho*, tendrán especial relevancia los siguientes:

— Principio de interdicción de la arbitrariedad: permitirá el control judicial, al implicar que la decisión debe ser racional, lo que también enlaza con las posibles correlaciones erróneas que establezca el sistema, ya aludidas, que podrían violar este principio. Asimismo, en cuanto a la motivación, violaría también este principio la ausencia de motivación, o la motivación ininteligible para una persona, de acuerdo con lo ya expuesto antes.

— Principio de Proporcionalidad: que exigirá que la decisión automatizada discrecional además de perseguir el interés general sea la menos restrictiva de los derechos de los ciudadanos y que sus beneficios superen a sus costes, como es sabido.

---

<sup>51</sup> Téngase en cuenta que el art. XVIII de la Carta de Derechos Digitales española establece el derecho de las personas 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 estas 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”.

— Principio de igualdad: evidentemente, se violará el mismo si se produce una discriminación causada por sesgos, que, como hemos visto, pueden afectar especialmente a cuestiones de género debido a los datos que se manejen. Recuérdese, asimismo, el apartado 1 del artículo 23 de la Ley 15/2022, de 12 de julio, que establece que las AAPP favorecerán la puesta en marcha de mecanismos para que los algoritmos involucrados en la toma de decisiones que se utilicen 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<sup>52</sup>.

— Principio de buena administración: el art. XVIII 6 de la Carta de Derechos Digitales española prevé que “las decisiones y actividades en el entorno digital respeten los principios de buen gobierno y el derecho a una buena Administración digital”. El ejercicio de la potestad de forma automatizada sin un análisis completo y diligente de los elementos relevantes y descartando los irrelevantes (como los sesgos cognitivos, de los datos o estadísticos) supondría una violación del derecho a la buena administración.

También es una violación de este principio *rigidificar* el ejercicio de la discrecionalidad, procediendo a usar IA simbólica o con reglas preprogramadas como dijimos.

De igual modo, como se dijo también, el ejercicio totalmente automatizado de potestades discrecionales mediante IA estadística violaría también el principio de buena administración, debido a la falta de empatía y de razonamiento abductivo, que impide tener en cuenta todos los factores relevantes en el ejercicio de la discrecionalidad, salvo que una ley permita dicho uso y aun así con comprobación del respeto de la buena administración.

En fin, como es obvio, la detección judicial de una violación del ordenamiento jurídico puede dar lugar a la declaración de invalidez de la decisión automatizada. De conformidad con las categorías clásicas en nuestro Derecho, como es sabido, deberá

---

<sup>52</sup> En relación con el control de la discriminación, téngase en cuenta la facilidad de la carga de la prueba para el ciudadano que se establece ahora en la ley estatal de igualdad 15/2022, en su art. 30:

“Artículo 30. Reglas relativas a la carga de la prueba.

1. De acuerdo con lo previsto en las leyes procesales y reguladoras de los procedimientos administrativos, cuando la parte actora o el interesado alegue discriminación y aporte indicios fundados sobre su existencia, corresponderá a la parte demandada o a quien se impute la situación discriminatoria la aportación de una justificación objetiva y razonable, suficientemente probada, de las medidas adoptadas y de su proporcionalidad.

2. A los efectos de lo dispuesto en el párrafo primero, el órgano judicial o administrativo, de oficio o a instancia de parte, podrá recabar informe de los organismos públicos competentes en materia de igualdad.

3. Lo establecido en el apartado primero no será de aplicación a los 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”.

analizarse caso por caso si se ha producido anulabilidad (art. 48.1 LPAC), nulidad (art. 47.1 y 2 LPAC) o una mera irregularidad no invalidante (art. 48.2 LPAC) (con carácter general, ALAMILLO y URIÓS, 2011). La peculiaridad más destacada en el caso de las decisiones totalmente automatizadas, y en especial, de las discrecionales que son las que más nos interesan especialmente, reside en el incumplimiento de la previsión del actual art. 41.2 LRJSP, en el sentido de no realizar la aprobación del 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, así como el órgano que debe ser considerado responsable a efectos de impugnación, o aprobar sólo unos aspectos y no otros o aprobarlos todos pero con defectos jurídicos.

La ausencia total de aprobación de dichas garantías supone, en nuestra opinión, un supuesto de invalidez de la decisión totalmente automatizada que se adopte, puesto que, como dijimos, estamos ante un procedimiento complejo asíncrono: si no se ha producido dicha aprobación, la decisión automatizada posteriormente adoptada debería ser declarada nula de pleno derecho de acuerdo con el art. 47.1, apartado e, al prescindirse totalmente de uno de los dos procedimientos administrativos que forman el procedimiento complejo (en este sentido, MENÉNDEZ, 2017, 106). O bien, si se considera, como sabemos hace parte de la doctrina, que la naturaleza jurídica de los algoritmos es reglamentaria, la invalidez sobrevendría por inexistencia del reglamento previo que la decisión automatizada debe ejecutar, art. 48.1 LPAC.

Si por el contrario, dicha aprobación existe pero es defectuosa, deberá estarse al tipo concreto de violación del ordenamiento jurídico caso por caso: la aprobación de algunas garantías, pero no de otras, entendemos que debería equipararse a la ya mencionada nulidad de pleno derecho, mientras que la existencia formal de todas pero con comprobación judicial de su insuficiencia (por ejemplo, se designa órgano para supervisión, pero se detecta su incapacidad para realizar adecuadamente la tarea o existe ponderación defectuosa de los costes y beneficios de la automatización, por citar ejemplos aludidos más arriba) debería conducir a que la aprobación inicial de las garantías fuera declarada, en principio, anulable y su invalidez debiera arrastrar la de la decisión automatizada adoptada mediante el segundo procedimiento de ejercicio de la potestad de que se trate, que constituye parte del procedimiento administrativo complejo descrito<sup>53</sup>.

---

<sup>53</sup> Si la aprobación de las garantías del art. 41 LRJSP es entendida como una disposición de naturaleza reglamentaria, entonces se abre temporalmente su control, ya que mediante el recurso indirecto ante la jurisdicción contenciosa puede discutirse al impugnar la decisión totalmente automatizada adoptada posteriormente (art. 26 LJCA). El límite a ello vendría marcado por la jurisprudencia que impide alegar defectos *formales* por la vía del recurso indirecto, si bien en ésta se matiza que “aunque la jurisprudencia relativa al recurso indirecto se ha mantenido después de la entrada en vigor de la Ley Jurisdiccional 29/1998, de 13 de julio, de lo que son exponente, entre otras, las Sentencias de esta Sala de 11 de octubre de 2005 (recurso de casación 6822/2002), 21 de abril de 2008 (recurso de casación 5038/2003), 11 de mayo de 2009 (recurso de casación 1871/2006) y 11 de febrero de 2010 (recurso de casación 4581/2008), sin embargo, se viene admitiendo la impugnación indirecta de disposiciones generales por defectos formales cuando se hubiese incurrido en una omisión clamorosa, total y absoluta

### 5.2.2. *Cuestiones específicas relativas al control judicial del uso de IA por las Administraciones. Indiferencia, insuficiencia, deferencia.*

El control desarrollado por jueces/zas y magistrados/as del uso de IA por las AAPP puede encontrarse con diversos de problemas y riesgos específicos.

Por un lado, si se aceptara que es posible ejercer discrecionalidad mediante IA, cosa que aquí se rechaza, y se mantuviera la concepción tradicional de la discrecionalidad, expuesta más arriba y discutida allí, de la *indiferencia* del Derecho frente a la *libre* elección administrativa, ahora realizada por máquinas, podría provocarse la ausencia de control judicial adecuado.

Además, los jueces/zas y magistrados/as, como seres humanos que son, pueden sufrir sesgos cognitivos específicos, esto es, errores sistemáticos predecibles, que podrían conducir a un control *insuficiente* de las potestades administrativas. Pensemos en el sesgo de automatización (que también pueden sufrir los gestores públicos, como dijimos, es decir, una confianza excesiva en la tecnología), en la combinación de este sesgo con el sesgo de anclaje en referencia a la decisión administrativa adoptada (el sesgo cognitivo de anclaje lleva a decidir en función de la primera información recibida, de la que es difícil deshacernos, fijando el marco de la futura decisión) y en otros posibles sesgos cognitivos, como el de grupo. Así respecto a éste, tomemos el ejemplo del papel de las pruebas periciales en los procesos contenciosos que puedan suscitarse en torno a la IA, las cuales pueden ser cruciales y, en el peor de los casos, llevar a substituir la decisión judicial por la de los expertos<sup>54</sup>.

---

del procedimiento establecido para su aprobación, en perjuicio del recurrente” (STS 19 de abril de 2012, E CLI:ES:TS:2012:2721). Pero es que, además, en nuestra opinión, los problemas procedimentales que se susciten durante la aprobación no son un mero defecto de forma, sino vicios que afectan al fondo tanto de dicha aprobación como de la decisión automatizada que se adopte posteriormente. Con carácter general, hemos insistido en diversas ocasiones en el error de entender los vicios procedimentales como meros defectos de forma, PONCE (2001, 2019c).

Si por el contrario se considera que la aprobación de las garantías del art. 41 LRJSP es un acto administrativo, sin naturaleza reglamentaria, la exigencia de su nulidad no tendría fecha de exigencia, dada la vía abierta por la revisión de oficio (art. 106 LPAC), como es sabido, si bien su impugnación mediante recursos administrativos y judiciales tendrá el conocido límite temporal, de meses, establecido por la LRJSP para los recursos administrativos ordinarios y por la LJCA para los recursos judiciales contra resoluciones expresas. Entonces, al discutir, dentro de plazo, la decisión automatizada adoptada procedente de una ausencia de aprobación de garantías o de aprobación defectuosa (nulidad o anulabilidad, en principio, respectivamente), al alegar que su invalidez trae causa de los defectos del primer procedimiento, nos podríamos encontrar ante un acto previo firme consentido, lo que obligaría entonces a obtener la revisión de oficio primero, para poder discutir luego la decisión automatizada (art. 125 b LPAC). Una manera de salvar esto sería entender que el procedimiento complejo no termina, en realidad, hasta la adopción de la decisión automatizada, por lo que al adoptar ésta podría recurrirse también la defectuosa aprobación de garantías violando el art. 41 LRJSP, aprobación que es, en definitiva, un acto trámite del procedimiento complejo que termina con la decisión automatizada.

<sup>54</sup> En este ámbito de las pruebas judiciales, puede darse, y de hecho se ha dado históricamente, un sesgo cognitivo judicial específico: el sesgo de grupo. Este sesgo supone la existencia de un error al valorar de forma homogénea, sea positiva o negativa, las actitudes, actos y opiniones de las personas de

Hay que tener en cuenta que, el uso de IA y su control es un tema complejo. Esta complejidad dificulta el control jurídico y puede hacer aparecer el riesgo de un insuficiente control judicial en garantía de los derechos de los ciudadanos. Volviendo al tema de las pruebas periciales, se ha señalado como “El juez ha dejado de ser mero receptor pasivo de la ciencia y de la tecnología para situarse en una posición en la que debe establecer qué datos técnicos o científicos son social y jurídicamente relevantes, como deben ser interpretados los datos científicos y qué parámetros (jurídico técnicos o científicos) han de fijar la credibilidad de los peritos” (DE LARRUCEA, 2010). Sin embargo, “el Juez o Sala pueden tener dificultades para comprender los informes técnicos y para resolver las dudas de contenido que estos le plantean (...) para valorar los informes, hay que entenderlos” (TRIANA, 2016)<sup>55</sup>.

Por ello, entendemos que frente a esta complejidad es imprescindible trabajar intensamente en la formación de los jueces en estos ámbitos, con componentes interdisciplinarios, e incidir en esa formación en la evitación (*debiasing*) de los sesgos descritos como más frecuentes en el ámbito de la IA, es decir los sesgos de automatización, anclaje y grupo, para evitar la insuficiencia del control judicial.

Aunque la situación real del uso de la IA en la actividad administrativa es todavía inicial, podríamos decir que se encuentra en su infancia, su despliegue es rápido y no se espera que decaiga, sino todo lo contrario. Además, aunque actualmente no son frecuentes supuestos de automatización total mediante uso de IA, no cabe descartar su incremento futuro. En este contexto, cabe plantearse, en tercer lugar, si el control judicial debe o puede ser deferente. Cuestión que nos introduce en el concepto de deferencia judicial, una expresión no demasiado utilizada en nuestra tradición jurídica y que cuyo significado, cuando se emplea, se da por supuesto<sup>56</sup>.

Respecto a la deferencia judicial, un análisis del *Diccionario Panhispánico del Español Jurídico* (RAE et al, 2023) arroja la inexistencia de una entrada con esta palabra, a diferencia de otras voces relacionadas que sí constan. Sí es posible hallar el término *Self-Restraint*, que se define por este diccionario como “Posición de auto restricción de los tribunales; tendencia conservadora en la interpretación de las leyes que se

---

un mismo grupo. En el ámbito de la prueba pericial, un sesgo de grupo se ha dado históricamente en el sentido de que los jueces de lo contencioso han valorado más positivamente de forma automática los informes periciales elaborados por funcionarios públicos (MUÑOZ, 2011), lo que afortunadamente está dejando paso a pronunciamientos del Tribunal Supremo en el sentido de negar ese mayor peso por sistema de pericias aportadas por la Administración controlada (STS de 17 de febrero de 2022 (rec. 5631/2019)). Pues bien, habrá que evitar que en el futuro se produzca otro sesgo judicial en favor de la delegación del ejercicio de la función judicial en expertos informáticos.

<sup>55</sup> Por ello se han efectuado diversas propuestas para mejorar la situación en general, entre ellas, la posibilidad de crear un cuerpo de peritos judiciales (imitando el cuerpo de médicos forenses) para asesorar a los órganos judiciales. Pese al coste y dificultades que ello pudiera generar, no es absurdo pensar en el futuro en un cuerpo de peritos judiciales profesionales con conocimientos multidisciplinarios que incluyeran la IA para poder ayudar a los jueces del contencioso, al menos respecto a los asuntos de mayor relevancia, que pudieran seleccionarse (por ejemplo, por afectación de derechos fundamentales, por cuantía, etc.).

<sup>56</sup> En la doctrina, véase el interesante análisis de RODRÍGUEZ PONTÓN, 2019.

atiene de modo estricto a su letra, al criterio seguido en actuaciones anteriores o a la adopción de las soluciones menos innovadoras del orden establecido” o el antónimo *activismo judicial*: “Posición fuertemente creativa de los jueces y tribunales, llevada a cabo mediante interpretaciones de la legalidad existente o a través de la cobertura de sus lagunas”<sup>57</sup>.

El TS español alude a la “doctrina legal sobre la naturaleza la extensión y los límites del control jurisdiccional de la potestad reglamentaria no supone reconocer un espacio de inmunidad a la actuación normativa del Gobierno, *que se soporte en el controvertido principio de deferencia del Poder Judicial al Poder Ejecutivo, que no se reconoce en el art. 106 de la Constitución*, sino someter la disposición general a un escrutinio riguroso basado en la aplicación de cánones de estricta legalidad, para no cercenar arbitrariamente el margen de decisión que le corresponde legítimamente al Gobierno” (STS de 23 de febrero de 2010)<sup>58</sup>.

El poco uso tradicional de la expresión en nuestro Derecho administrativo y el riesgo de malinterpretación del concepto aconsejan manejar el término deferencia con mucha cautela. En todo caso, en nuestra opinión, la actividad judicial respecto a las decisiones administrativas automatizadas no debe ser ni indiferente, ni insuficiente, ni tampoco timorata en la defensa de los derechos de las personas en relación con la IA. Es cierto que la legitimidad de las Administraciones obtenida mediante una composición organizativa vinculada estrechamente a elementos democráticos (ej. ámbito local) o procedimientos administrativos participativos son factores que llaman a la prudencia judicial para evitar invadir ámbitos constitucionalmente vedados. Como también lo es la complejidad técnica o mejor dicho “la diferencia de capacidad o brecha cognoscitiva entre la Administración y los Tribunales” (DOMÈNECH, 2021, 919). Pero dicha prudencia no puede convertirse en un incumplimiento de las funciones otorgadas por el art. 106.1 CE para la tutela de los derechos, conforme al art. 24 CE, puesto que como ha sido puesto de relieve, existe el riesgo siempre de que la Administración “se desvíe de la solución más conveniente para el interés público y presumiblemente querida por el legislador. Cuanto mayor sea ese riesgo, menor habrá de ser la deferencia judicial” (DOMÈNECH, 2021, 915).

<sup>57</sup> Sobre el activismo judicial, puede consultarse ATIENZA, M. (2018),

<sup>58</sup> Las referencias judiciales se extraen de RODRÍGUEZ PONTÓN, 2019. En votos particulares de diversos ponentes a diversas SSTs de 2011, como la de 7 de febrero, realizado en este caso por el magistrado Garzón Herrero, se mantiene esa visión negativa de la deferencia judicial: “Por lo que hace al “principio de deferencia”, no puedo por menos que expresar mi frontal rechazo a tal principio. En primer lugar, porque aparece en ordenamientos jurídicos que nada tiene que ver con el nuestro en los que los derechos individuales llevan siglos consagrados, y en donde la misión de los órganos constitucionales está, generalmente, muy elaborada, lo que no sucede en nuestro caso. En segundo término, porque el principio constitucional de igualdad procesal de las partes tiene difícil acomodo con el enunciado expuesto. Finalmente, porque la función de la jurisdicción contenciosa no es “observar” “deferencia con el ejecutivo” sino “controlar la legalidad de la actividad administrativa”, “garantizar derechos e intereses de los ciudadanos” y “hacer frente a las extralimitaciones de la Administración” (Conforme al párrafo primero de la Exposición de Motivos de nuestra Ley)”.



Pues bien, *en el contexto actual expuesto*, de creciente empleo de IA por las Administraciones, de insuficiente regulación en España de las garantías para el ciudadano, como demuestra el art. 41.2 LRJSP aquí comentado, de falta de suficientes garantías organizativas — con una Agencia Española de Supervisión de Inteligencia Artificial en fase de constitución (cuyo Estatuto fue aprobado por Real Decreto 729/2023, de 22 de agosto)— y de riesgo de impactos negativos en el uso de la IA, de carácter masivo, además, que pueden afectar a miles de ciudadanos, un elemento que diferencia el uso de máquinas frente a la decisión humana, creemos que *en la actualidad* el control judicial debería consistir en un control riguroso (*hard look*, en la expresión clásica norteamericana empleada desde los años 70 del siglo pasado<sup>59</sup>), especialmente del cumplimiento del procedimiento administrativo complejo de preparación del uso de los algoritmos y de adopción de la decisión automatizada con estos, eso sí, sin violar el principio constitucional de distinción funcional entre poderes.

En línea similar, la negación de la deferencia judicial en este ámbito es expuesta por diversos especialistas angloamericanos, pese a provenir de sistemas que, por su historia y tradición son menos autoritarios y más favorables al uso de la discrecionalidad y que han manejado históricamente este concepto de deferencia. Así, entre otros, Coglianese (COGLIANESE, 2021, pp. 109 y 110), tras considerar diversos aspectos de las decisiones automatizadas administrativas desde la óptica norteamericana señala que (la traducción del inglés y la cursiva es nuestra):

“Por supuesto, *todo esto no quiere decir que las agencias vayan a recibir o deban recibir siempre deferencia por la forma en que diseñan o gestionan sus sistemas*. En virtud de la norma articulada en el asunto *Motor Vehicle Manufacturers Association v. State Farm Insurance Co.* deberán proporcionar información básica sobre los objetivos

---

<sup>59</sup> Las reflexiones sobre la deferencia en el ámbito del Derecho administrativo no pueden desligarse de la toma en consideración de otros desarrollos simultáneos. Así, en EE. UU. de la doctrina judicial norteamericana del *hard look*, vinculada, de hecho, al llamado segundo paso de la sentencia *Chevron* de 1984 del Tribunal Supremo de los EEUU (*Chevron U.S.A., Inc. v. Natural Resources Defense Council, Inc.*), epítome de la deferencia en aquel país (dictada en relación con el ejercicio de la potestad reglamentaria de las agencias y la interpretación jurídica de las leyes que desarrollan).

La doctrina jurisprudencial norteamericana de *hard look*, también conocida como el estándar del “procedimiento razonado” (*reasoned decision making*) es ampliamente empleada por la jurisprudencia norteamericana y muestra como en el moderno Derecho Público de los Estados Unidos “el énfasis en la revisión de la arbitrariedad se ha movido hacia el escrutinio de la calidad del razonamiento administrativo” (LEVIN Y GELLHORN, 2007), que es también lo que ha ocurrido en el ámbito de la Unión Europea y de sus estados miembros al calor del derecho a una buena administración, por cierto. Desde la década de 1970, los tribunales, e incluso el Tribunal Supremo de los EE. UU., han exigido que los decisores administrativos actúen con cautela, prestando especial atención a los factores e intereses relevantes en juego. La jurisprudencia ha establecido el deber de escuchar a los ciudadanos, el deber de responder a sus comentarios mediante una exposición de motivos, y el deber de estudiar detenidamente la regulación antes de decidir, como marco para garantizar unas normas sólidas (MATHEWS, 2015). Los tribunales exigen a las agencias que incluyan en sus declaraciones de bases y propósitos discusiones detalladas de sus razones que apoyan el ejercicio de la discrecionalidad y muestran la existencia de la debida atención al considerar los hallazgos y comentarios presentados durante la audiencia pública prevista en el artículo 553 de la APA (AUBY & PERRAUD, 2014, 27).



de sus sistemas automatizados y su funcionamiento general<sup>60</sup>. Deberán demostrar que han estudiado detenidamente las principales opciones de diseño. Y es probable que tengan que demostrar, mediante auditorías y validaciones aceptadas, que estos sistemas funcionan para producir los resultados previstos. Pero todo esto equivale a decir que es casi seguro que las agencias podrán proporcionar la información necesaria para justificar resultados que producen sus sistemas. En otras palabras, los antiguos principios del derecho administrativo parecen estar preparados para la era de la automatización.”<sup>61</sup> .

Recogiendo estas preocupaciones, el CONSEJO DE EUROPA a finales de 2022 (CONSEJO DE EUROPA, 2022) publicó el estudio comparativo *Artificial Intelligence and Administrative Law* , preparado por WOLSWINKEL<sup>62</sup>, el cual señala en su página 14 que (traducción del inglés nuestra): “hay el riesgo de que debido a una falta de conocimiento técnico los tribunales avanzarán hacia la deferencia aún más”, por comparación al tradicional estándar de control judicial, “cuando complejos ADM” (iniciales en inglés de sistemas de decisión automatizada) “sean usados por las autoridades públicas. Al mismo tiempo, se sostiene que el Derecho administrativo debería moverse lejos de la tradicional asunción sobre deferencia hacia más fuertes formas de rendición de cuentas en términos de transparencia y comunicación, doctrinas de relevancia y revisión substantiva en particular respecto a resultados que provienen de ADM”.

## 6. CONCLUSIONES

Que exista IA disponible técnicamente, ahora o en el futuro, para ser aplicada en cualquier ámbito de funcionamiento administrativo no significa que sea ni deseable ni posible jurídicamente aplicarla y obtener decisiones discrecionales totalmente automatizadas.

<sup>60</sup> Mientras que en el ámbito de la adjudicación, el caso judicial principal referido al *hard look* en EE.UU. es *Citizens to Preserve Overton Park v. Volpe*, 401 US 402 (1971), en el ámbito de la elaboración de reglamentos, el asunto principal es *Motor Vehicle Manufacturers Ass'n v. State Farm Mutual Automobile Insurance Co.*, 463, U.S. 29 (1983), que se refería a la impugnación de una decisión de la *National Highway Traffic Safety Administration* de rescindir una norma en virtud de la cual la agencia había exigido a todos los fabricantes de automóviles que incluyeran en todos los coches comercializados en EE.UU. a partir de una fecha determinada uno o dos dispositivos de seguridad (airbags o cinturones de seguridad automáticos). En esta decisión, el Tribunal Supremo enunció el criterio de arbitrariedad y capricho de la siguiente manera (traducción nuestra):

“Normalmente, una norma de una agencia sería arbitraria y caprichosa si la agencia se ha basado en factores que el Congreso no pretendía que tuviera en cuenta, ha omitido por completo considerar un aspecto importante del problema, ha ofrecido una explicación de su decisión contraria a las pruebas de que disponía la agencia o es tan inverosímil que no puede atribuirse a una diferencia de opinión o al producto de la experiencia”.

<sup>61</sup> En este mismo sentido, desde la perspectiva británica WILLIAMS (2019).

<sup>62</sup> Consultable en: [4cb20e4b-3da9-d4d4-2da0-65c11cd16116 \(coe.int\)](https://www.coe.int/t/t02/ai_law/ai_law.asp)

En el caso de las potestades discrecionales, desde la perspectiva no de lo deseable sino de lo jurídicamente posible, consideramos que el derecho y principio a una buena administración impide que la IA simbólica elimine la discrecionalidad y que la IA no simbólica o conexionista vaya más allá de ser un apoyo al decisor humano, puesto que, en el primer caso, no es posible jurídicamente preprogramar agotadoramente la discrecionalidad y eliminar ésta y en el caso de la IA estadística la ausencia de empatía y de capacidad de formular inferencias abductivas recomiendan la ausencia de automatización total. Por otra parte, la misma no será posible cuando el principio contradictorio en el procedimiento administrativo exija un trámite de audiencia o cuando exista participación ciudadana en trámites de información pública, sea en este caso la potestad discrecional o reglada, puesto que será preciso sopesar en la decisión las aportaciones de la ciudadanía.

Los límites vistos de la IA y sus diferencias con los humanos pueden aconsejar prohibir normativamente su uso en los ámbitos expuestos, prohibición que entendemos debería desarrollarse aplicando una opción legal por defecto (*opt out*) en favor del no uso de la IA, de acuerdo con el principio de precaución aludido, salvo que una norma específica la permita. Norma que deberá ser aprobada por los representantes democráticos en aplicación de la reserva de ley existente en el art. 18.4 CE, garantizando que no se viola el principio y derecho de buena administración que tiene relevancia constitucional, de acuerdo con reiterada jurisprudencia del Tribunal Supremo.

Como hemos sostenido aquí, argumentamos que esta prohibición debiera adoptarse *en estos momentos, de conformidad con el estado actual de la tecnología y el principio de precaución*. Nada impediría modificar esta conclusión, y la opinión de quien esto escribe, si la situación evoluciona de otro modo en el futuro y se avanza hacia una IA general o fuerte, equiparable a la humana y dotada de empatía, inexistente a fecha de hoy, lo que no parece que vaya a ocurrir ni a corto ni a medio plazo.

Aceptando el uso de la IA, ahora ya en el marco del procedimiento administrativo complejo para la toma de decisiones automatizadas, cuya primera fase se refiere a la evaluación previa mediante un primer procedimiento administrativo de aprobación de las características del sistema, la Administración debiera considerar, caso por caso la calidad, necesidad, idoneidad y eficiencia del uso de IA y establecer las garantías precisas para su empleo.

Superado este primer procedimiento administrativo y en fase ya de adopción de la concreta decisión administrativa en el específico procedimiento administrativo sectorial mediante uso de la IA, esta debería acompañarse de mecanismos de supervisión o vigilancia humana que sean efectivos, como se dijo. Ello implica la necesidad de una formación y entrenamientos específicos de los empleados públicos (pues tal supervisión puede afirmarse que constituye una función pública a ellos reservada) que van a vigilar los sistemas de IA en materia de *desesgamiento* 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 diseñada (HAUSER, 2020).

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. En este sentido 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 práctica, como en el ámbito jurídico destaca Marchetti (MARCHETTI, 2021; muy ilustrativo al respecto, desde una perspectiva técnica, es el punto de vista de LEINS, K y KASPERSEN, A., 2021). 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*, ineficaz en realidad.

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, a incluir en el art. 53 LPAC, generador de correlativas obligaciones jurídicas para la Administración, (obligaciones cuanto menos de medios, sometidas al estándar de debido cuidado propio de la buena administración; véase la interesante sentencia del Tribunal Supremo español de 15 de febrero de 2022<sup>63</sup>).

En cuanto al control judicial de la IA, hemos expuesto que esa tarea judicial se encuentra con distintas dificultades que no la hacen fácil, pero que deben ser afrontadas, en garantía del control de legalidad de la actividad administrativa y de la protección de los derechos de las personas. Para ello, parece imprescindible ofrecer a los jueces y juezas formación específica en esta materia.

La automatización con IA tiene una serie de riesgos ya descritos, con un impacto posible en una escala inimaginable desde el punto de vista tradicional de la adopción de decisiones por humanos. En la actualidad, su regulación legal es insuficiente y la formalización del necesario procedimiento de aprobación de garantías conforme al art. 41.2 LRJSP inexistente. En la medida que en el futuro se afine el marco regulatorio y se potencie la organización administrativa en torno a la IA, *la importancia del atento control judicial puede decrecer*. Pero hoy en día, necesitamos que el control judicial del contencioso administrativo en España de la IA no sea indiferente ni insuficiente para la correcta protección de los derechos de las personas, puesto que no creemos que existan razones *per se* y automáticas de mayor objetividad e imparciali-

---

<sup>63</sup> STS de 15 de febrero de 2022, ECLI:ES:TS:2022:543, 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””.

dad administrativa (recuérdense los sesgos cognitivos posibles de los programadores —por ejemplo, disponibilidad y confirmación— y de los supervisores —así, automatización— en el ámbito de la IA) ni de mejor cognoscibilidad (ténganse en cuenta los problemas de las bases de datos que se manejan actualmente<sup>64</sup>. HUERGO, 2023, sostiene la aplicación del “mismo grado de control” a las decisiones automatizadas con IA que tendría una decisión no automatizada que se “tomara sin IA”<sup>65</sup>.

Para todo ello necesitamos jueces/zas y magistrados/as formados y activos, pero con conciencia de su limitado, aunque importante, y en parte por explorar, papel constitucional en garantía de los derechos de los ciudadanos y de la buena administración. Esta labor judicial se vería facilitada, desde luego, mediante una mejora de la regulación estatal, autonómica y local sobre régimen jurídico y procedimiento administrativo en este campo, que permita orientar y limitar el ejercicio de potestades discrecionales mediante el uso de IA, como hemos ido exponiendo en este trabajo, más allá del mero uso de códigos éticos o de la aprobación de meras normas de *soft law*, como cartas o protocolos (PONCE, 2022c).

En el desarrollo del RUEIA aprobado por el Parlamento europeo el 13 de marzo de 2024 y en las nuevas regulaciones españolas que se aprueben en el futuro en su marco deberá hacerse un esfuerzo para aunar lo mejor de lo humano con lo mejor de la IA, para evitar así combinar lo peor de ambos mundos, esto es, decisiones con la (inexistente) empatía y capacidad abductiva de la IA junto con los límites y errores cognitivos humanos.

## 7. BIBLIOGRAFÍA

ALAMILLO, I. y URIÓS, F. X., *La actuación administrativa automatizada en el ámbito de las administraciones públicas. Análisis jurídico y metodológico para la construcción y la explotación de trámites automáticos*, Generalitat de Catalunya, 2011. Disponible en: [L'actuació administrativa automatitzada en l'àmbit de les Administracions Públiques \(gencat.cat\)](#)

ATIENZA, M., *Sobre la dignidad humana*, Trotta, Madrid, 2022.

---

<sup>64</sup> Por ejemplo, un estudio publicado a finales de 2021 y titulado “*Reduced, Reused and Recycled: The Life of a Dataset in Machine Learning Research*” elaborado por Bernard Koch y Jacob G. Foster de la UCLA y Emily Denton y Alex Hanna de Google Research ha detectado que existe una creciente desigualdad en el uso de los conjuntos de datos a nivel mundial, y que más del 50 por cien de todos los conjuntos de datos utilizados en la muestra empleada correspondían a conjuntos de datos introducidos por doce instituciones de élite, principalmente occidentales (y añadimos nosotros, tradicionalmente dominadas por hombres...). Entre las instituciones dominantes se nos informa que se encuentran la Universidad de Stanford, Microsoft, Princeton, Facebook, Google, el Instituto Max Planck y AT&T. Cuatro de las diez principales fuentes de conjuntos de datos son instituciones corporativas. El estudio es consultable aquí: [2112.01716.pdf \(arxiv.org\)](#) y puede hallarse un comentario al mismo en esta página web: El estudio es consultable aquí: [A Cartel of Influential Datasets Is Dominating Machine Learning Research, New Study Suggests - Unite.AI](#).

<sup>65</sup> Como señala GALLONE (2023, 222), el uso de inteligencia artificial en los procedimientos administrativos debe venir compensado con la garantía de un control judicial humano inspirado en el criterio de *full jurisdiction*, establecido por el TEDH en relación con el art. 6.1 del CEDH.

- ATIENZA, M., “Siete Tesis sobre el Activismo Judicial”, *Grand place: pensamiento y cultura*, núm. 10, 2018.
- AUBY, J-B and PERROUD, T., “Introduction”, en AUBY, J-B and PERROUD, T. (Eds.), *Regulatory Impact Assessment*, INAP, GLP, 2013.
- BACIGALUPO, M., “La vinculación de la Administración Pública a la Ley y al Derecho”, en VELASCO, F. y DARNACULLETA, M., *Manual de Derecho Administrativo*, Marcial Pons, 2023.
- BOIX PALOP, A., “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*, núm. 1, 2020, pp. 223-269.
- BRENT, J., *Charles Sanders Pierce. A life*, Indiana University Press, 1993.
- BRIDLE, J., *La nueva edad oscura: la tecnología y el fin del futuro*, Debate, 2020.
- COBBE, J., “Administrative law and the machines of government: Judicial review of automated public-sector decision-making” *Legal Studies*, 39(4), 636-655, 2019.
- COGLIANESE, C., “Administrative Law in the Automated State”, 150 *Daedalus* 104, 2021.
- COGLIANESE, C. y LAI, A., “Algorithm vs. Algorithm”, *Faculty Scholarship at Penn Law*. 2795, 2022.
- CONSEJO DE EUROPA, *Artificial Intelligence and Administrative Law*, 2022, disponible en: [4cb20e4b-3da9-d4d4-2da0-65c11cd16116 \(coe.int\)](https://www.coe.int/t/e/treaties/4cb20e4b-3da9-d4d4-2da0-65c11cd16116).
- COTINO HUESO, L., ““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*, Wolters Kluwer nº 4, mayo, 2020.
- CRAIG, P., *Administrative Law*, Sweet and Maxwell, Thompson Reuters, 9th edition, 2021.
- CHAMORRO DOMÍNGUEZ, M.C., “La aplicación de sistemas de inteligencia artificial en el seno del órgano de administración de las sociedades de capital”, *Revista de Derecho de Sociedades*, núm. 59, mayo-agosto, 2020.
- DAMASIO, A., *Sentir y Saber*, Destino, 2021.
- DAMASIO, A., *El error de Descartes*, Andrés Bello, 1994.
- DE LARRUECA, R., “La contradicción en el dictamen pericial: tratamiento conceptual de la sana crítica y efectos prácticos”, *Diario LA LEY*, núm. 7494 13413/2010, 2010.
- DELLA CANANEA, G., “The ‘Core’ of Administrative Law: an Outsider’s View”, *BACL*, 2022. Disponible en: <https://british-association-comparative-law.org/?p=4543>.
- DE LA SIERRA, S., “Inteligencia artificial y justicia administrativa: una aproximación desde la teoría del control de la Administración Pública”, *Revista General de Derecho Administrativo*, enero, 2020.
- DOMÈNECH, G., “El juego del legislador, la Administración y el Juez”, en Coordinadores Rodríguez De Santiago, José María, Domènech Pascual, Gabriel y Arroyo Jiménez, Luis, *Tratado de Derecho Administrativo*, Marcial Pons, 2021.
- GALLONE, G., *Riserva di Umanità e Funzione Amministrative*, Wolters Kluwers, CEDAM, 2023.
- GAMERO CASADO, E., “Las garantías de régimen jurídico del sector público y del procedimiento administrativo común frente a la actividad automatizada y la inteligencia artificial”, en *Inteligencia artificial y sector público: retos, límites y medios*, coord. por Eduardo Gamero Casado, Francisco Luis Pérez Guerrero, 2023, págs. 397-461
- GONZÁLEZ NAVARRO, F., “Procedimiento administrativo común, procedimientos triangulares y procedimientos complejos”, *Revista de Estudios de La Administración Local y Autonómica*, (211), 1981, pp. 407-454
- HAUSER, C., “From Preaching to Behavioral Change: Fostering Ethics and Compliance Learning in the Workplace,” *Journal of Business Ethics*, Springer, vol. 162(4), 2020, pp. 835-855.
- HOFMANN, HCH., “The Duty of Care in EU Public Law—A Principle Between Discretion and Proportionality”, *Review of European Administrative Law*, vol. 13, no 2, 2020, p. 87-112.
- HUERGO LORA, “El derecho de transparencia en el acceso a los códigos fuente”, en *Anuario de Transparencia Local*, 5/2022, Fundación Democracia y Gobierno Local, 2023, pp. 35 y ss.
- HUYS, V., “L’Abduction ou le risque interprétatif », *Revue Recherches sur la Philosophie et le Langage*, núm. 34 - *Abduction*, sous la direction de Rémi Clot-Goudard, Viviane Huys et Denis Vernant, 2018.
- KAHNEMAN, D., SIBONY, O., SUNSTEIN, C., *Ruido: Un fallo en el juicio humano*, Debate, 2021.

- LAGUNA DE PAZ, J.C., “El control judicial de la discrecionalidad administrativa”, *Revista española de Derecho Administrativo*, núm. 186, 2017.
- LARSON, E.J., *The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do*, Belknap Press: An Imprint of Harvard University Press, 2021.
- LEINS, K y KASPERSEN, A., “7 Myths of Using the Term “Human on the Loop”: “Just What Do You Think You Are Doing, Dave?”: Seven Myths of Using the Term “Human on the Loop”: “Just What Do You Think You Are Doing, Dave?” | Carnegie Council for Ethics in International Affairs, 2021.
- LEVIN, R. M. & GELLHORN, E., *Administrative law and process in a nutshell*, 5th Edition, 2007.
- LICKELIDER J.C.R., “Man-Computer symbiosis”, *IRE Transactions on Human Factors in Electronics*, volume HFE-1, pages 4-11, March, 1960.
- MATTHIEU, R., *En defensa del altruismo*, Urano, 2016.
- McCANN, S., “Discretion in the Automated State”, *The Canadian Journal of Law & Jurisprudence*, 2023, pp. 1 y ss.
- MAN, K.; DAMASIO, A., “Homeostasis and soft robotics in the design of feeling machines”, *Nature Machine Intelligence*, vol. 1, núm. 10, 2019, p. 446-452.
- MARCHETTI, B., “La garanzia dello human in the loop alla prova della decisione amministrativa algoritmica”, *BioLaw Journal – Rivista di BioDiritto*, núm. 2/2021, 2021.
- MARTÍN DELGADO, I., “Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada”, *Revista de administración pública*, núm. 180, 2009, págs. 353-386.
- MARTÍNEZ, N., AGUDO, U. y MATUTE, H., “Human cognitive biases present in Artificial Intelligence”, *Revista Internacional de los Estudios Vascos*, 67, 2, 2022. Disponible en: [Human cognitive biases present in Artificial Intelligence | Riev, International Journal on Basque Studies \(eusko-ikas-kuntza.eus\)](https://www.eusko-ikas-kuntza.eus/).
- MATHEWS, J., “Searching for Proportionality in US Administrative Law”, *Legal Studies Research Paper number 1-2015*, *Penn State Law*, 2015. Disponible en : [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2561583](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2561583).
- MÉDIATEUR FÉDÉRAL, *Rapport d'activités*, 2021. Disponible en : *Rapport annuel 2021.pdf (federalombudsman.be)*.
- MENÉNDEZ, E., *Las garantías del interesado en el procedimiento administrativo electrónico*, Tirant Lo Blanch, 2017.
- MORAL SORIANO, L., “Decisiones automatizadas, Derecho administrativo y argumentación jurídica”, en LLANO, F.H. (Dr.), GARRIDO, J. y VALDIVIA, R., (Coords.), *Inteligencia artificial y filosofía del Derecho*, Laborum, 2022.
- MUÑOZ ARANGUREN, A., “La influencia de los sesgos cognitivos en las decisiones jurisdiccionales: el factor humano. Una aproximación”, *InDret*, abril, 2011.
- NIETO, A., “El arcano de la discrecionalidad”, estudio preliminar a REBOLLO, M, HUERGO, A, GUILLÉN, J, CAMPO, T., *Anuario del Derecho Administrativo Sancionador 2021*, Civitas Thompson Reuters, 2021, pp. 27 y ss.
- NIETO, A. y SÁNCHEZ, R.M., *La equidad judicial*, Colex, 2022.
- PASTOR, A. y NOGALES, P., “El futuro del trabajo en la administración pública ¿estamos preparados?” *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, núm. extra 3, 2019, págs. 34-51.
- PONCE SOLÉ, J., *Deber de buena administración y derecho al procedimiento administrativo debido. Las bases constitucionales del procedimiento administrativo y del ejercicio de la discrecionalidad*, Lex Nova, 2001.
- PONCE SOLÉ, J., “La discrecionalidad no puede ser arbitrariedad y debe ser buena administración”, *Revista española de derecho administrativo*, núm. 175, 2016, págs. 57-84.
- PONCE SOLÉ, J., “Inteligencia Artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico”, *Revista General de Derecho Administrativo*, núm. 1, enero, 2019.
- PONCE SOLÉ, J., *La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida*, UAH-Defensor del Pueblo, 2019b.
- PONCE SOLÉ, J., *Mejora de la Regulación, Lobbies y Huella Normativa*, Tirant Lo Blanch, 2019c.



- PONCE SOLÉ, J., "La buena administración y el buen gobierno: la crisis de la Covid-19 y las nuevas tendencias en el Derecho administrativo y en el control judicial del poder ejecutivo", *Jueces para la democracia*, nº núm., 102, 2021, págs. 5-20.
- PONCE SOLÉ, J., "Reserva de humanidad y supervisión humana de la Inteligencia artificial", *El Cronista del Estado Social y Democrático de Derecho*, núm. 100, 2022.
- PONCE SOLÉ, J., *Acicates (Nudges), buen gobierno y buena administración. Aportaciones de las ciencias conductuales, nudging y sectores público y privado*, Marcial Pons, Madrid, 2022 b.
- PONCE SOLÉ, J., "Las relaciones entre inteligencia artificial, regulación y ética, con especial atención al sector público", *Revista General de Derecho Administrativo*, núm. 61, 2022c.
- PONCE SOLÉ, J. (2023): "Seres humanos e inteligencia artificial: discrecionalidad artificial, reserva de humanidad y supervisión", en Eduardo Gamero Casado (coord.), Francisco Luis Pérez Guerrero (coord.) *Inteligencia artificial y sector público: retos, límites y medios*, Tirant Lo Blanch.
- PONCE SOLÉ, J. (2023b): "El derecho a una buena administración, su exigencia judicial y el privilegio de ejecutoriedad de los actos administrativos. A propósito de la Sentencia de la Sala 3ª del Tribunal Supremo 1421/2020, de 28 de mayo de 2020, recurso de casación 5751/2017", *Revista de Administración Pública*, núm. 221, 163-182.
- PONCE SOLÉ, J., *Manual de Fonaments del Dret Administratiu i de la Gestió Pública*, Tirant Lo Blanch, València, 4ª edición, 2023c.
- PONCE SOLÉ, J. y VILLORIA MENDIETA, M., *Anuario del Buen Gobierno y de la Calidad de la Regulación*, Fundación Democracia y Gobierno Local, 2020, 2021, 2022 y 2023.
- PULLMAN LE, REFAIE N, LALUMIÈRE ML, KRUPP D., "Is Psychopathy a Mental Disorder or an Adaptation? Evidence From a Meta-Analysis of the Association Between Psychopathy and Handedness". *Evolutionary Psychology*, October, 2021.
- RANCHORDAS, S., "Experimental Regulations for AI: Sandboxes for Morals and Mores", *University of Groningen Faculty of Law*, Research Paper núm. 7/2021, 2021.
- RANCHORDAS, S., "Empathy in the digital administrative state", *Duke Law Journal*, Forthcoming, *University of Groningen Faculty of Law, Research Paper*, 2022, núm. 13.
- REAL ACADEMIA ESPAÑOLA et al. *Diccionario Panhispánico del Español Jurídico*. Disponible en: [Diccionario panhispánico del español jurídico - Real Academia Española \(rae.es\), 2023.](#)
- REIS DE ALBUQUERQUE, A. y DE BRITO MACHADO SEGUNDO, H., "A perspectiva evolutiva e o problema da relação corpo-mente: qual o impacto da ausência do instinto, das emoções e do Nao consciente no processo decisório?", en PEDRO, R. y CALIENDO, P. (Coords), *Inteligência artificial no contexto do Direito público: Portugal e Brasil*, Almedina, 2023.
- RODRÍGUEZ PONTÓN, F., "El control judicial en el ámbito de la regulación económica: Comisión Nacional de los Mercados y la Competencia (CNMC) y Comisión Nacional del Mercado de Valores (CNMV) un relato de la deferencia en el marco del Derecho común europeo", en MALARET, E. (Coord.), *Autonomía Administrativa, decisiones cualificadas y deferencia judicial*, Thompson Reuters Aranzadi, 2019.
- SOM, C., HILTY, L.M and KÖHLER, A.R., "The Precautionary Principle as a Framework for a Sustainable Information Society", *Journal of Business Ethics*, 85, 2009 :493-505.
- TORCHIA, L., *Lo Stato Digitale*, Il Mulino, 2023.
- TRIANA REYES, B., "¿Vale de algo la prueba pericial en la jurisdicción contencioso-administrativa?", *Actualidad administrativa*, núm. 6, junio, 2016.
- TRUMBULL IV, C.P., "Autonomous Weapons: How Existing Law Can Regulate Future Weapons", *Emory International Law Review*, vol. 34, núm. 2, 2020, pp. 539 y ss.
- VAQUER, M., "La Humanidad del Derecho Administrativo de nuestro tiempo", *Revista de Administración Pública*, núm. 222, 2023, pp. 33 y ss.
- WEIZENBAUM, J., *Computer power and human reason: From judgment to calculation*, W. H. Freeman & Co, 1976, traducido al español un año más tarde con el título *La frontera entre el ordenador y la mente*, Ed. Pirámide, S.A. Madrid. 1977.
- WILLIAMS, R., "Rethinking Deference for Algorithmic Decision-Making", *Oxford Legal Studies Research Paper No. 7/2019*, 2019.