

ESTUDIOS

Incluye



Papel

Digital

ORDALÍAS ALGORÍTMICAS: JUSTICIA, FE Y DATOS EN LA ERA DE LA INTELIGENCIA ARTIFICIAL

MARGARITA SIMARRO PEDREIRA

 ARANZADI

© Margarita Simarro Pedreira, 2026

© ARANZADI LA LEY, S.A.U.

ARANZADI LA LEY, S.A.U.

C/ Collado Mediano, 9

28231 Las Rozas (Madrid)

www.aranzadilaley.es

Atención al cliente: <https://areacliente.aranzadilaley.es/publicaciones>.

Primera edición: enero 2026

Depósito Legal: M-627-2026

ISBN versión impresa con complemento electrónico: 978-84-1085-598-4

ISBN versión electrónica: 978-84-1085-599-1

Diseño, Preimpresión e Impresión: ARANZADI LA LEY, S.A.U.

Printed in Spain

© ARANZADI LA LEY, S.A.U. Todos los derechos reservados. A los efectos del art. 32 del Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba la Ley de Propiedad Intelectual, ARANZADI LA LEY, S.A.U., se opone expresamente a cualquier utilización del contenido de esta publicación sin su expresa autorización, lo cual incluye especialmente cualquier reproducción, modificación, registro, copia, explotación, distribución, comunicación, transmisión, envío, reutilización, publicación, tratamiento o cualquier otra utilización total o parcial en cualquier modo, medio o formato de esta publicación.

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra solo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la Ley. Diríjase a **Cedro** (Centro Español de Derechos Reprográficos, www.cedro.org) si necesita fotocopiar o escanear algún fragmento de esta obra.

El editor y los autores no asumirán ningún tipo de responsabilidad que pueda derivarse frente a terceros como consecuencia de la utilización total o parcial de cualquier modo y en cualquier medio o formato de esta publicación (reproducción, modificación, registro, copia, explotación, distribución, comunicación pública, transformación, publicación, reutilización, etc.) que no haya sido expresa y previamente autorizada.

El editor y los autores no aceptarán responsabilidades por las posibles consecuencias ocasionadas a las personas naturales o jurídicas que actúen o dejen de actuar como resultado de alguna información contenida en esta publicación.

ARANZADI LA LEY no será responsable de las opiniones vertidas por los autores de los contenidos, así como en foros, chats, o cualesquiera otras herramientas de participación. Igualmente, ARANZADI LA LEY se exime de las posibles vulneraciones de derechos de propiedad intelectual y que sean imputables a dichos autores.

ARANZADI LA LEY queda eximida de cualquier responsabilidad por los daños y perjuicios de toda naturaleza que puedan deberse a la falta de veracidad, exactitud, exhaustividad y/o actualidad de los contenidos transmitidos, difundidos, almacenados, puestos a disposición o recibidos, obtenidos o a los que se haya accedido a través de sus PRODUCTOS. Ni tampoco por los Contenidos prestados u ofertados por terceras personas o entidades.

ARANZADI LA LEY se reserva el derecho de eliminación de aquellos contenidos que resulten inveraces, inexactos y contrarios a la ley, la moral, el orden público y las buenas costumbres.

Nota de la Editorial: El texto de las resoluciones judiciales contenido en las publicaciones y productos de ARANZADI LA LEY, S.A.U., es suministrado por el Centro de Documentación Judicial del Consejo General del Poder Judicial (Cendoj), excepto aquellas que puntualmente nos han sido proporcionadas por parte de los gabinetes de comunicación de los órganos judiciales colegiados. El Cendoj es el único organismo legalmente facultado para la recopilación de dichas resoluciones. El tratamiento de los datos de carácter personal contenidos en dichas resoluciones es realizado directamente por el citado organismo, desde julio de 2003, con sus propios criterios en cumplimiento de la normativa vigente sobre el particular, siendo por tanto de su exclusiva responsabilidad cualquier error o incidencia en esta materia.

Índice General

Página

PARTE I ORDALÍAS Y LA VERDAD REVELADA

INTRODUCCIÓN	17
CAPÍTULO I	
EL JUICIO DE DIOS	21
1. Que era una ordalía: definición, contexto y función social ...	21
1.1. Entre el rito y la regla. Anatomía conceptual de la ordalía	21
1.2. La ordalía como encrucijada interpretativa. Distintos enfoques	23
2. Las ordalías en la historia	25
2.1. Tipologías: fuego, agua, combate y otras formas rituales	26
2.2. La sombra del Juicio divino en Roma	30
2.3. Las ordalías en la cosmovisión germánica	32
2.4. La fe como prueba en el medievo español	33
3. La epistemología de la fe	37
3.1. Verdad divina y ausencia de prueba racional	37
3.2. Verdad algorítmica	48

PARTE II
DEL RITO A LA MÁQUINA: GÉNESIS
DEL ALGORITMO COMO ORDALÍA

CAPÍTULO II

EL ALGORITMO COMO NUEVA ORDALÍA	55
1. El carácter ritual de la decisión algorítmica.....	55
1.1. <i>El silencio del Código: la revolución 4.0 como orden jurídico encubierto.</i>	55
1.2. <i>Qué es un algoritmo, más que una fórmula, menos que una norma.....</i>	60
2. De la comunidad a la computación	65
2.1. <i>Juicio público vs. decisión automatizada.....</i>	65
2.2. <i>La tensión entre las garantías procesales y la automatización judicial</i>	72
2.3. <i>La tutela de los derechos digitales en la Unión Europea y España .</i>	78

PARTE III
EL ALGORITMO
EN EL PROCESO PROBATORIO

CAPÍTULO III

EL ALGORITMO COMO TESTIGO: DESAFÍOS PARA EL DERECHO PROBATORIO	99
1. El <i>ius probandi</i> frente a la opacidad algorítmica.....	99
1.1. <i>El derecho a la prueba como garantía esencial de la tutela judicial efectiva</i>	99
1.2. <i>La irrupción de la prueba algorítmica en el proceso judicial...</i>	107
2. El control de los oráculos	116
2.1. <i>¿Es posible auditar al algoritmo?</i>	116
2.2. <i>El problema de la opacidad</i>	123

PARTE IV
INTELIGENCIA ARTIFICIAL Y JUSTICIA PREDICTIVA

CAPÍTULO IV

INTELIGENCIA ARTIFICIAL Y JUSTICIA PREDICTIVA	131
1. ¿Qué es la inteligencia artificial en el contexto jurídico?.....	131
1.1. <i>Concepto de IA y distinción con el algoritmo</i>	131
1.2. <i>Origen y clasificación</i>	136
2. Sistemas algorítmicos en decisiones judiciales	142
2.1. <i>¿Qué es justicia predictiva?</i>	142
2.2. <i>El algoritmo COMPAS CORE: el oráculo matemático de la justicia predictiva</i>	149
2.3. <i>Justicia predictiva en España. Promesas digitales y espejismos normativos</i>	156

PARTE V
EPITAFIO DE LA PERSONALIDAD ROBÓTICA:
EL RETORNO ENCUBIERTO DE LAS ORDALÍAS

1. Una breve reflexión sobre la personalidad jurídica de la IA	177
2. Obituario de una categoría sin sujeto	182

POST SCRIPTUM.....	185
BIBLIOGRAFÍA	187
WEBGRAFÍA	197

Índice Imágenes

	<i>Página</i>
Imagen 1: En el mapa mostrado se pueden ver en rojo las CCAA donde es utilizado el sistema Minerva: Comunidades del Territorio del Ministerio (Baleares, Castilla-la Mancha, Castilla y León, Ceuta, Extremadura, Melilla y Región de Murcia) y aquellas Comunidades que tienen competencias transferidas en materia de justicia (Asturias, La Rioja y Galicia), así como la AN y el TS y la Fiscalía Europea	69
Imagen 2: Gráfico de comparativa de las medidas adoptadas por individuos en España frente a la media de la UE (2023) ..	84
Imagen 3: Gráfico donde se muestra la comparación de incidentes y medidas de seguridad adoptadas por las empresas en España frente a la media de la Unión Europea (2024)	86
Imagen 4: Competencias digitales en España.	89
Imagen 5: Gráfico de Comparación de las competencias digitales básicas por género, edad y situación laboral: España frente a la media de la UE (2023).....	90
Imagen 6: Gráfico de la Actividad en Derechos y Principios Digitales en España (2024)	94
Imagen 7: Gráfico del impacto percibido de las iniciativas sobre Derechos Digitales en España.....	95

Capítulo IV

Inteligencia Artificial y justicia predictiva

1. ¿QUÉ ES LA INTELIGENCIA ARTIFICIAL EN EL CONTEXTO JURÍDICO?

1.1. CONCEPTO DE IA Y DISTINCIÓN CON EL ALGORITMO

Se ha examinado el concepto de algoritmo desde una pluralidad de enfoques (económico, técnico y procesal), lo que permite una aproximación interdisciplinar que revela tanto su dimensión instrumental como sus implicaciones normativas. Asimismo, se han contrastado las dos principales corrientes doctrinales en la materia: la teoría funcionalista, orientada a la eficacia operativa y a la optimización de resultados, y la teoría garantista, centrada en la preservación de derechos fundamentales y en la necesidad de controles jurídicos que limiten eventuales efectos adversos¹.

Sobre la base de ese análisis, cabría preguntarse: ¿hasta qué punto es posible trasladar al ámbito de la inteligencia artificial la estructura conceptual aplicada al algoritmo, y en qué medida la diferenciación entre ambos resulta esencial para delimitar su tratamiento jurídico, identificar los riesgos específicos que plantea la IA y justificar la existencia de un régimen regulatorio autónomo? La respuesta radica en que, en el debate contemporáneo sobre tecnología y Derecho, se ha generalizado el uso indistinto de los términos *algoritmo* e *inteligencia artificial*.

Esta confusión semántica no es inocua: afecta a la delimitación del objeto regulatorio, a la interpretación de las normas vigentes y a la forma en que los tribunales deben abordar problemas relativos a transparencia, responsabilidad y tutela judicial efectiva.

Así, el algoritmo puede definirse como una secuencia finita, ordenada y precisa de instrucciones destinada a la resolución de un problema o a la

1. Vid. Cap. II apdo. 1.2 de este libro.

ejecución de una tarea concreta. Su lógica es esencialmente predecible y determinista: dadas unas entradas, produce invariablemente las mismas salidas. Constituye, por tanto, la base de la programación y se aplica en ámbitos diversos, desde las matemáticas hasta la informática. Un ejemplo clásico es el algoritmo de búsqueda binaria, empleado para localizar un elemento en un conjunto ordenado de datos. Otro ejemplo más sencillo sería que el algoritmo funciona como una receta de cocina (primer paso: la entrada, que serían los ingredientes de la receta; segundo paso: las instrucciones, que constituirían los pasos a seguir para la elaboración; y el tercer paso: la salida, la comida lista para servir).

Por su parte, la inteligencia artificial (IA) no se reduce a un único algoritmo, sino que comprende un conjunto de técnicas, modelos y arquitecturas, que integran algoritmos de distinta naturaleza, diseñados para reproducir, en mayor o menor medida, funciones asociadas a la cognición humana, tales como el aprendizaje, el razonamiento, la percepción o la toma de decisiones en entornos de incertidumbre.

A diferencia de los algoritmos tradicionales, los sistemas de IA pueden generar resultados distintos ante los mismos datos de entrada, en la medida en que incorporan procedimientos de carácter estadístico, probabilístico y adaptativo.

La diferencia fundamental entre ambos conceptos radica en la capacidad de aprendizaje y adaptación: mientras que un algoritmo ejecuta un conjunto de instrucciones previamente definido y no varía su comportamiento, la inteligencia artificial puede ajustar su rendimiento en función de la experiencia acumulada y de la incorporación de nuevos datos. En consecuencia, los algoritmos constituyen la infraestructura básica sobre la cual se erigen los sistemas de inteligencia artificial, proporcionando las reglas elementales de funcionamiento, mientras que la IA introduce la posibilidad de mejorar, generalizar y evolucionar a partir de la información procesada.

Sin embargo, no existe una definición universal de lo que es la IA debido a que “su raudo desarrollo se encarga de imposibilitar cualquier intento de consolidación conceptual: toda formulación definidora nace anudada a una pronta obsolescencia”². Se pueden extraer conceptos de la IA a partir de los entornos en los que es utilizada, pero sin que sea de manera uniforme ni

2. JULIÁ PIJOAN, M.: “La computarización del derecho: a partir del proceso y de los procedimientos judiciales”. Ed. Dikynson, 2024. Pág. 25.

igualitaria (no es lo mismo la IA en un contexto informático, donde primarán las cuestiones técnicas explicativas de las tecnologías base de la IA como *machine learning* y *deep learning*, que en el escenario jurídico).

En el ámbito normativo, el Reglamento de IA de la UE ofrece una definición jurídica precisa de la inteligencia artificial. Conforme al artículo 3.1, se entiende por *sistema de inteligencia artificial* “un *software* desarrollado con una o varias de las técnicas y enfoques enumerados en el anexo I, que, para un conjunto dado de objetivos definidos por el ser humano, puede generar resultados tales como contenido, predicciones, recomendaciones o decisiones que influyen en los entornos con los que interactúa”.

En el Anexo I del Reglamento se concretan los métodos incluidos bajo esta definición, destacando:

- a) técnicas de aprendizaje automático (incluidos el aprendizaje supervisado, no supervisado y por refuerzo)
- b) enfoques basados en lógica y conocimiento (lógica inductiva, deductiva, sistemas expertos, representaciones de conocimiento)
- c) métodos estadísticos, de estimación bayesiana y de optimización.

De esta manera, la Unión Europea diferencia explícitamente entre el algoritmo aislado (como ya se ha dicho, regido por un conjunto de instrucciones predefinidas y de comportamiento invariable) y el sistema de IA, entendido como un *software* con capacidad de aprendizaje, adaptación y generación de resultados no estrictamente deterministas, con efectos potenciales sobre derechos e intereses jurídicamente protegidos.

La diferencia esencial radica, pues, en que el algoritmo constituye un elemento técnico básico, mientras que la inteligencia artificial se configura como un sistema complejo, autónomo y evolutivo, sujeto a un marco regulatorio específico orientado a la gestión del riesgo, la transparencia, la trazabilidad, la seguridad y la rendición de cuentas (arts. 5 a 29 AI Act).

Una vez establecida la diferenciación entre algoritmo e IA, hay que profundizar en el propio concepto de IA, estableciendo sus efectos a distintos niveles, como políticos, sociales y económicos.

Una definición sencilla, aunque no por ello conceptualmente más relevante, concibe la inteligencia artificial como una disciplina académica vincu-

lada a la teoría de la computación, orientada a la emulación de determinadas facultades intelectuales humanas en sistemas artificiales³. No obstante, esta caracterización presenta el riesgo de reducir el fenómeno al plano estrictamente técnico, invisibilizando sus dimensiones sociales, jurídicas y políticas, que resultan esenciales para comprender la amplitud de sus implicaciones y el alcance de su impacto.

En un contexto caracterizado por una creciente interdisciplinariedad, resulta más acorde con la realidad reconocer que bajo el rótulo de «inteligencia» se agrupan tecnologías heterogéneas y de naturaleza diversa. Esta amalgama terminológica no es neutra: contribuye a magnificar expectativas, alimentar narrativas de inevitabilidad tecnológica y legitimar intervenciones normativas o económicas cuya fundamentación resulta, en muchos casos, cuestionable. Así, la ambigüedad conceptual no solo obstaculiza un análisis sistemático y riguroso de los impactos sociales, políticos y económicos de la inteligencia artificial, sino que opera como un recurso retórico de poder, orientado a encubrir intereses concretos y a desplazar el debate desde sus dimensiones materiales hacia un plano eminentemente simbólico.

Desde ese enfoque interdisciplinar y con una crítica al mito de la inteligencia artificial como algo autónomo, inmaterial y casi humano: “la IA no es artificial ni inteligente. Más bien existe de forma corpórea, como algo material, hecho de recursos naturales, combustible, mano de obra, infraestructuras, logística, historia y clasificaciones. Los sistemas de IA no son autónomos, racionales ni capaces de discernir algo sin un entrenamiento extenso y computacionalmente intensivo, con enormes conjuntos de datos o reglas y recompensas predefinidas. De hecho, la IA como la conocemos depende por completo de un conjunto mucho más vasto de estructuras políticas y sociales. Y, debido al capital que se necesita para construir IA a gran escala y a las maneras de ver que optimiza, los sistemas de IA son, al fin y al cabo, diseñados para servir a intereses dominantes ya existentes. En este sentido la IA es un certificado de poder”⁴.

En suma, este enfoque desmonta la narrativa tecnocientífica de la IA como una entidad autónoma e inevitable y la reubica como infraestructura de poder material, social y político.

3. BENÍTEZ IGLESIAS, R.: “Inteligencia artificial avanzada”, Ed. UOC, 2014, pág. 12.

4. CRAWFORD, K.: “Atlas de IA; poder, política y costes planetarios de la inteligencia artificial”, Ned Ediciones, 2023, pág. 29.

En el mismo sentido, Shoshana Zuboff ha descrito como *capitalismo de la vigilancia*⁵, en el que los datos y su explotación constituyen una nueva forma de acumulación de poder económico. En paralelo, Mireille Hildebrandt⁶ ha advertido que la IA no es una mera técnica de cálculo, sino una tecnología reguladora que condiciona la arquitectura misma del derecho y de la vida social.

En suma, ambos enfoques tratan de que la IA no solo genera valor económico en el marco de un nuevo capitalismo basado en datos (Zuboff), sino que también reconfigura las formas de regulación social y jurídica (Hildebrandt). Se trata de dos críticas que muestran cómo la IA está íntimamente vinculada al poder: por un lado, a su acumulación económica; por otro, a su ejercicio normativo.

La IA, más que un medio tecnológico, es un cúmulo de factores complejos en el que intervienen infraestructuras materiales, recursos naturales, capital financiero, trabajo humano, marcos normativos, intereses políticos y dinámicas culturales que condicionan su desarrollo y aplicación. Esta complejidad desmiente la concepción de la IA como entidad autónoma o neutra, revelándola, en cambio, como un entramado profundamente situado y orientado por relaciones de poder preexistentes.

En este sentido, las llamadas “ordalías algorítmicas” ilustran con claridad cómo esta complejidad queda oculta tras la apariencia de objetividad técnica: al igual que en las ordalías medievales, donde el fuego o el agua eran presentados como pruebas incuestionables de la voluntad divina, los sistemas algorítmicos se legitiman como procedimientos neutrales de decisión, cuando en realidad encarnan y reproducen intereses políticos y económicos ya establecidos. Reconocer esta dimensión implica entender que la regulación de la IA no puede limitarse a parámetros técnicos de transparencia o eficiencia, sino que debe enfrentar el carácter profundamente político de unas tecnologías que funcionan como rituales modernos de legitimación del poder.

-
5. La era del capitalismo de la vigilancia (Spanish Edition) 2019 https://www.perio.unlp.edu.ar/catedras/ecal/wp-content/uploads/sites/210/2023/06/ZUBOFF_1.pdf. Última consulta el 10/10/2025.
 6. HILDEBRANDT, M.: “Smart Technologies and the End(s) of Law novel entanglements of law and technology. en Smart Technologies and the End (s) of Law. Edward Elgar Publishing. 2016.
- HILDEBRANDT, M.: “Esclavos de los macrodatos. ¿O no?” en IDP. Revista de Internet, Derecho y Política. Núm. 17, 2013. págs. 7-26. UOC.

1.2. ORIGEN Y CLASIFICACIÓN

La comprensión del fenómeno de la inteligencia artificial exige una mirada retrospectiva a sus orígenes y a su evolución conceptual y tecnológica.

En 1950, Alan Turing publicó en la revista *Mind* el artículo “Computing Machinery and Intelligence”, en el que formuló la célebre pregunta: *¿pueden pensar las máquinas?* Para responderla, propuso un método experimental que pasaría a conocerse como Test de Turing, concebido como herramienta para evaluar la capacidad de una máquina de exhibir un comportamiento inteligente indistinguible del humano.

La lógica del *Test* de Turing se fundamenta en un procedimiento mediante el cual no se puede distinguir quien es el que está respondiendo. El evaluador humano interactúa simultáneamente con una persona y con una máquina, sin conocer la identidad de cada interlocutor. Si, tras el intercambio, el interrogador no logra discernir con certeza cuál de las dos entidades corresponde a la máquina, se concluye que esta ha superado la prueba y, por tanto, ha mostrado un comportamiento inteligible equiparable al humano. En cambio, si el evaluador identifica sin dificultad a la máquina, se entiende que no ha alcanzado el umbral de inteligencia propuesto por Turing.

Entre sus fortalezas destacan la simplicidad metodológica y su aplicabilidad a contextos de lenguaje natural. Sin embargo, también se han señalado debilidades: privilegia la apariencia de inteligencia sobre su comprensión real, depende en exceso de la ingenuidad del interrogador y no aborda cuestiones más amplias relativas a la conciencia o al aprendizaje autónomo.

A pesar de estas críticas, el *Test* de Turing ha tenido un impacto decisivo tanto en la filosofía de la mente como en el desarrollo práctico de la IA. Ha inspirado competiciones como el Premio Loebner y ha impulsado la evolución de chatbots y sistemas de interacción conversacional, hoy presentes en ámbitos tan diversos como la medicina, la educación o la atención al cliente.

Por otro lado, el desarrollo del computador digital programable por Konrad Zuse, la Z3 en 1941, considerada la primera computadora programable y plenamente operativa de la historia.

Además, Zuse concibió en 1945 el Plankalkül, un lenguaje de programación de alto nivel de carácter fundamentalmente teórico, ya que no llegó a implementarse durante su vida ni influyó de manera directa en los primeros lenguajes posteriores. En 1946 fundó la que puede considerarse la primera empresa dedicada a la fabricación de ordenadores y, pocos años después,

construyó la Z4, que en 1950 se convirtió en el primer computador en ser comercializado.

El contexto de la Segunda Guerra Mundial limitó la difusión internacional de sus avances, que permanecieron relativamente desconocidos fuera de Alemania. No obstante, en 1946 se documenta una primera influencia tangible de su trabajo en el extranjero, cuando IBM adquirió varias de sus patentes.

El término “inteligencia artificial” fue acuñado en 1955 por John McCarthy, y un año después la Conferencia de Dartmouth consolidó el campo como disciplina autónoma, sin olvidar que, en 1960, Donald Michie, investigador británico y considerado uno de los mayores contribuyentes en la creación de la IA, desarrolló Machine Educable Noughts and Crosses Engine (MENACE), uno de los primeros programas que jugaba de un modo perfecto al tres en raya, que ilustra cómo la IA emergió en un contexto histórico marcado por la guerra, la criptografía y la creciente confianza en las máquinas como auxiliares de la decisión. Desde una perspectiva jurídica, resulta significativo que estos primeros sistemas se concibieran no como entes autónomos, sino como prolongaciones de la agencia humana.

En las décadas posteriores, el avance de la IA se articuló en hitos significativos: ELIZA fue el primer programa informático de procesamiento del lenguaje natural, desarrollado entre 1964 y 1966 en el Massachusetts Institute of Technology (MIT) por Joseph Weizenbaum. Concebido como un experimento sobre la interacción entre seres humanos y máquinas, el sistema se basaba en un mecanismo de concordancia y sustitución de patrones que generaba en los usuarios la ilusión de mantener una conversación significativa. Sin embargo, ELIZA carecía de cualquier forma de representación semántica o de comprensión genuina del lenguaje, limitándose a reproducir respuestas predeterminadas sin interpretar el contenido del diálogo.

ELIZA operaba mediante la detección de palabras clave en las frases introducidas por el usuario, a partir de las cuales seleccionaba y devolvía una respuesta previamente registrada en su base de datos.

En determinadas ocasiones, las réplicas generadas resultaban lo suficientemente verosímiles como para inducir a algunos interlocutores, en su primer contacto con el sistema, a creer que estaban interactuando con un agente humano. No obstante, esta ilusión de inteligibilidad se desvanecía cuando el usuario intentaba atribuir significados implícitos o matices interpretativos a las respuestas del programa, pues tales contenidos no existían

en la lógica de funcionamiento del sistema. En ese punto, la interacción tendía a tornarse incoherente.

Asimismo, cuando se introducía una expresión que no coincidía con ningún patrón previsto, ELIZA solía reformularla en forma de pregunta o de reflexión general, estrategia que reforzaba el efecto conversacional pese a la ausencia de comprensión real del lenguaje. Sin embargo, ELIZA puede considerarse tanto uno de los primeros *chatterbots*, antecesores de los actuales *chatbots*, como uno de los primeros programas diseñados para poner a prueba, al menos de manera embrionaria, los supuestos de la célebre prueba de Turing.

En 1996, la empresa IBM desarrolló la supercomputadora Deep Blue con el objetivo de enfrentar a una máquina contra un jugador humano en partidas de ajedrez bajo un ritmo de juego lento. Ese mismo año, Deep Blue se midió con el entonces campeón mundial, Gary Kaspárov, quien resultó vencedor por 4-2.

Un año más tarde, IBM presentó una versión mejorada del sistema, denominada Deeper Blue, preparada ya para competir bajo los estándares temporales habituales de las competiciones oficiales. En este nuevo encuentro, la supercomputadora logró imponerse a Kaspárov, convirtiéndose así en la primera máquina en derrotar a un campeón humano en un juego que exige inteligencia estratégica.

Siguiendo con la política de mejoras de la empresa informática IBM, en 2005 comenzó a diseñar un sistema que considerara un desafío mucho más complejo que ganar partidas de ajedrez. Así surgió Watson, un sistema de inteligencia artificial concebido para responder a preguntas formuladas en lenguaje natural. Su creación se enmarca en el proyecto DeepQA, dirigido por el investigador David Ferrucci, y su denominación rinde homenaje a Thomas J. Watson, fundador y primer presidente de la compañía.

El funcionamiento de Watson se basa en una extensa base de datos almacenada localmente, nutrida a partir de una amplia diversidad de fuentes: enciclopedias, diccionarios, tesauros, artículos periodísticos, obras literarias, así como bases de datos externas, taxonomías y ontologías.

La capacidad del sistema fue puesta a prueba públicamente en febrero de 2011, durante un torneo especial del concurso televisivo estadounidense *Jeopardy*, en el que compitió durante tres jornadas contra dos de los concur-

santes más exitosos del programa: Brad Rutter, máximo ganador económico de su historia, y Ken Jennings, poseedor del récord de victorias consecutivas. Watson se impuso con claridad, obteniendo el primer premio de un millón de dólares. Rutter y Jennings recibieron 200.000 y 300.000 dólares respectivamente, comprometiéndose ambos a donar la mitad de sus ganancias a organizaciones benéficas, mientras que IBM distribuyó el premio de Watson entre World Vision y World Community Grid.

En 2016, Google DeepMind, desarrolló un programa informático de IA llamado AlphaGo para jugar al juego de mesa Go. En esa ocasión, se enfrentó al jugador chino Fan Hui en una serie de cinco partidas oficiales, todas ellas ganadas por el sistema, así como en encuentros informales adicionales que finalizaron 3-2 a favor de la inteligencia artificial.

Posteriormente, en marzo de 2016, AlphaGo disputó un histórico enfrentamiento contra uno de los mejores jugadores del mundo, el surcoreano Lee Sedol. El 9 de marzo de ese año, AlphaGo obtuvo la victoria en la primera partida, marcando un hito en la historia de la inteligencia artificial aplicada a juegos de estrategia.

En la actualidad, la IA se ha expandido a sectores estratégicos como la automoción, el turismo, los seguros y la logística, impulsada por el reconocimiento de imágenes, el análisis masivo de datos (*big data*) y los modelos predictivos. Este despliegue plantea no solo oportunidades económicas, estimadas en centenares de miles de millones de dólares, sino también retos éticos, sociales y jurídicos en torno a la privacidad, la transparencia algorítmica y la posible sustitución del trabajo humano.

Sin embargo, es preciso advertir que toda narración sobre los «orígenes» no es nunca inocente: las genealogías de la IA suelen construirse de manera selectiva, destacando ciertos hitos y silenciando otros, de modo que configuran un relato legitimador de la tecnología y de los intereses que la promueven. Así, la historia que sitúa a Alan Turing, la Conferencia de Dartmouth de 1956 o los sistemas expertos de los años setenta como momentos fundacionales ha tendido a invisibilizar otras tradiciones, como los desarrollos de la cibernetica soviética⁷, las contribuciones de ingenieras y programadoras en la implementación práctica de los sistemas o los debates ético-filosóficos que acompañaron al progreso técnico.

7. La cibernetica soviética, aunque poco recordada, constituye un episodio clave para comprender la historia global de la inteligencia artificial y el sesgo occidentalocéntrico de sus relatos fundacionales. Inicialmente rechazada en los años 40 y principios de los 50 como “pseudociencia burguesa”, fue rehabilitada tras la muerte de Stalin,

Como ha mostrado Kate Crawford en *Atlas of AI*, las historias dominantes de la IA suelen borrar las condiciones materiales y sociales de su desarrollo, reforzando una narrativa de inevitabilidad tecnológica, lo que significa que la IA no es solamente el producto del ingenio humano o de los avances tecnológicos, sino que es el resultado de unas condiciones materiales, sociales y políticas concretas.

Aquí entra en juego la clasificación de la IA⁸ ya que su importancia radica en el hecho de constituir una herramienta práctica con consecuencias:

convirtiéndose en un área estratégica de investigación bajo el amparo de la Academia de Ciencias de la URSS.

Durante las décadas de 1950 y 1960, se desarrollaron aportes significativos: modelos matemáticos de neuronas artificiales, intentos de planificación económica centralizada mediante ordenadores (“socialismo cibernetico”), y experimentos en traducción automática y programas de juego, en paralelo a los avances occidentales. Sin embargo, la escasez de hardware moderno y las prioridades militares limitaron su consolidación, provocando su declive en los años 70.

Historiográficamente, la experiencia soviética demuestra que la IA no surgió únicamente en torno a Turing o Dartmouth, sino también de proyectos colectivos orientados a fines políticos y sociales. En la URSS, la IA se vinculó a la planificación estatal y al ideal socialista, lo que contrasta con la orientación industrial y mercantil de Occidente. Ejemplo: WEST, D. K.: “Cybernetics for the command economy: Foregrounding entropy in late Soviet planning”. en *History of the Human Sciences*, 33(1), 2020. Págs..36-51; GEROVITCH, S. *From Newspeak to Cyberspeak: A History of Soviet Cybernetics*. Cambridge, MA: MIT Press. 2002 (Disponible en línea) From Newspeak to Cyberspeak: A History of Soviet Cybernetics; https://monoskop.org/images/7/76/Gerovitch_Slavia_From_Newspeak_to_Cyberspeak_A_History-of-Soviet_Cybernetics_2004.pdf.

GÜNTHER, CLEMENS. “The Cultural and Political Imaginary of Cybernetic Socialism.” *Kritika: Explorations in Russian and Eurasian History*, vol. 24 no. 2, 2023, p. 321-348. Project MUSE, (Disponible en línea) <https://dx.doi.org/10.1353/kri.2023.0018>

8. AREND HINTZE distingue cuatro tipos de inteligencia artificial:

1. Máquinas reactivas: sistemas básicos que solo responden a estímulos presentes, sin memoria ni aprendizaje (ejemplo: *Deep Blue* de IBM).
2. Memoria limitada: pueden usar experiencias pasadas de manera transitoria para mejorar decisiones, pero sin aprendizaje a largo plazo.
3. Teoría de la mente: IA capaz de interpretar pensamientos y emociones de otros, ajustando su conducta en consecuencia.
4. Autoconciencia: etapa más avanzada, donde la máquina tendría conciencia de sí misma y podría formar representaciones internas de su propia existencia.

Por su parte, RUSSELL y NORVIG proponen una clasificación de la IA que combina dos ejes: pensamiento vs. comportamiento y humano vs. racional. De este cruce surgen cuatro enfoques:

1. Pensar como humanos: modelar procesos cognitivos (ej. redes neuronales, sistemas de diagnóstico, creatividad artificial).
2. Actuar como humanos: imitar conductas humanas sin importar el proceso interno (ej. chatbots, robots humanoides).

1. Selección de aplicaciones: distinguir, por ejemplo, entre IA débil (centrada en tareas específicas) e IA fuerte (hipotética, de propósito general), o entre sistemas de aprendizaje supervisado y no supervisado, permite elegir qué tipo de tecnología es viable para un problema concreto. Una empresa que necesite predecir ventas no requiere la misma IA que una que busque conducción autónoma.
2. Entender limitaciones: la clasificación ayuda a reconocer qué puede y qué no puede hacer cada enfoque. Por ejemplo, un sistema basado en reglas no es adecuado para problemas de reconocimiento de patrones masivos, y un modelo de *deep learning* no ofrece explicaciones transparentes de sus decisiones. Esta comprensión evita la “mitología de la IA” y frustra expectativas desmedidas.
3. Ética y regulación: al tipificar sistemas según riesgos o funciones, se pueden trazar pautas de uso responsable. La Unión Europea, por ejemplo, clasifica la IA en categorías de riesgo (inaceptable, alto, limitado, mínimo), lo que permite definir obligaciones diferenciadas. Sin clasificaciones, hablar de “ética de la IA” sería demasiado abstracto.

En este sentido, la clasificación de la IA ya sea por niveles de riesgo, por capacidad de autonomía o por áreas de aplicación, no puede entenderse como un ejercicio meramente técnico o descriptivo. Se trata más bien de un dispositivo de poder que, al ordenar un campo heterogéneo, selecciona qué aspectos son visibles y cuáles quedan relegados. Así, mientras se presentan como tipologías neutrales que facilitan la regulación o la seguridad jurídica, estas clasificaciones contribuyen a reforzar la ilusión de objetividad y a estabilizar un discurso que naturaliza la expansión de la IA, ocultando tanto sus costes materiales (energéticos, ecológicos, laborales) como las asimetrías sociales y políticas que la sostienen.

El mundo del Derecho no resulta ajeno a la problemática de la clasificación de la IA, ya que cumple varias funciones estratégicas:

-
3. Pensar racionalmente: aplicar lógica formal y razonamiento óptimo (ej. sistemas expertos, planificación).
 4. Actuar racionalmente: orientarse a maximizar objetivos en distintos entornos (ej. agentes inteligentes en juegos, robótica autónoma).

En conjunto, esta tipología ofrece una visión multidimensional de la IA, distinguiendo entre su dimensión interna (procesos) y su manifestación externa (acciones). RUSSEL, S.J.; NORVIG, P.: “Inteligencia Artificial. Un enfoque moderno”, Pearson Educación S.A, 2004.

1. Determinación normativa: al ordenar los sistemas de IA en categorías (por riesgo, función o grado de autonomía), el legislador puede asignar obligaciones diferenciadas. Ejemplo: el *AI Act* de la UE establece niveles de riesgo que van desde el “inaceptable” hasta el “mínimo”.
2. Facilitar la imputación de responsabilidad: al distinguir tipos de IA, es más fácil determinar qué estándares de diligencia aplicar y quién responde ante un fallo: ¿el programador, el proveedor del modelo, el usuario final?
3. Orientación judicial: los jueces necesitan criterios para interpretar casos en los que interviene la IA. Una clasificación permite traducir complejidad técnica a categorías jurídicas manejables.
4. Seguridad jurídica y previsibilidad: las empresas y administraciones requieren un marco estable para invertir y desplegar tecnologías. Clasificar reduce la incertidumbre (aunque nunca del todo).
5. Instrumento retórico y político: la clasificación no solo ordena, también legitima. Al trazar fronteras entre lo “aceptable” y lo “prohibido”, el derecho fija narrativas sobre qué IA es deseable y cuál no.

Ahora bien, clasificar en derecho corre el riesgo de fosilizar un fenómeno tecnológico en evolución continua. Lo que hoy se encaja como “bajo riesgo” puede mañana tener impactos sociales graves. Además, la clasificación puede ser manipulada: los actores poderosos tenderán a influir para que sus sistemas queden en categorías benignas. Lo que nos lleva al inicio: la inteligencia artificial es menos un logro puramente técnico que el resultado de contextos materiales, sociales y políticos determinados.

En suma, en derecho la clasificación de la IA sirve como técnica de gobernanza, pero también puede convertirse en un velo de certeza que oculte la contingencia, los conflictos de intereses y la dimensión política de la tecnología.

2. SISTEMAS ALGORÍTMICOS EN DECISIONES JUDICIALES

2.1. ¿QUÉ ES JUSTICIA PREDICTIVA?

Tras haber examinado la figura histórica de las ordalías como juicios de Dios y su paralelismo con los algoritmos contemporáneos, así como la definición de algoritmo y su distinción respecto de la inteligencia artificial,

junto con el impacto de estas tecnologías en los derechos fundamentales y su empleo en la gestión judicial, corresponde ahora abordar el análisis de la justicia predictiva como manifestación específica del uso de la inteligencia artificial en el ámbito jurídico y como un fenómeno que plantea tensiones profundas entre la promesa de eficiencia y previsibilidad, por un lado, y la preservación de principios básicos del Estado de derecho, imparcialidad, motivación judicial, debido proceso, por otro.

Las predicciones constituyen un elemento constante en la vida cotidiana contemporánea. El recurso a modelos predictivos se manifiesta en ámbitos muy diversos: desde la meteorología, orientada a anticipar fenómenos climáticos, hasta la economía, donde se proyectan índices y tendencias de mercado; también en el ámbito lúdico, como en los juegos de azar, donde se busca acertar combinaciones numéricas futuras, o en el terreno político, a través de encuestas y análisis destinados a anticipar el resultado electoral. Del mismo modo, en el deporte se realizan pronósticos sobre competiciones y rendimientos. En suma, la práctica de anticipar escenarios futuros constituye una dimensión transversal que atraviesa prácticamente todas las esferas de la experiencia social.

Ahora bien, cuando esta lógica se traslada al ámbito jurídico surgen tensiones cualitativamente distintas. A diferencia de la predicción meteorológica, económica o deportiva, que se legitima por su utilidad práctica y por su valor probabilístico, en el derecho la anticipación de resultados no puede reducirse a un mero cálculo de probabilidades.

En definitiva, la predicción ha sido tradicionalmente entendida como una medida de validación del conocimiento científico. En efecto, la fuerza de una teoría no reside únicamente en su capacidad para explicar lo ocurrido, sino también en su aptitud para anticipar lo que ocurrirá. Esta lógica, presente en las ciencias naturales, se ha extendido paulatinamente a las ciencias sociales y, más recientemente, al ámbito jurídico, donde el estudio del comportamiento judicial no debería limitarse a una reconstrucción retrospectiva de decisiones ya dictadas, sino someterse también a la prueba de su capacidad prospectiva. Es por ello que “la mejor prueba de una teoría explicativa es su capacidad para predecir eventos futuros. En la medida en que los académicos de ambas disciplinas (ciencias sociales y derecho) buscan explicar el comportamiento judicial, deberían probar sus teorías no solo contra casos ya decididos, sino también contra resultados futuros”⁹.

9. MARTIN, A.D.; QUINN, K.M.; RUGER, T.W.; KIM, P.T.: “Competing approaches to predicting supreme court decision making” en *Perspectives on Politics*, n.º 2(4),

La noción de *justicia predictiva* no es unívoca, sino que abarca diversas aproximaciones. Una primera corriente la concibe como un instrumento destinado a afinar las expectativas de las partes, en la medida en que permite anticipar razonablemente el sentido de una decisión jurisdiccional mediante la correcta identificación e interacción de las variables interpretativas de los hechos y las normas. En este enfoque, las herramientas algorítmicas no ofrecen certezas absolutas, sino estimaciones probabilísticas que orientan el comportamiento procesal¹⁰.

Otra perspectiva entiende la justicia predictiva como el desarrollo de instrumentos destinados a facilitar la labor de las partes a través del análisis sistemático de los pronunciamientos judiciales previos, con el objetivo de proyectar, con un grado elevado de precisión, las decisiones futuras. Este planteamiento se apoya en modelos matemáticos y estadísticos, en tanto que el razonamiento jurídico se asimila, al menos en parte, al funcionamiento de una función matemática susceptible de ser replicada y explotada tecnológicamente¹¹.

Al igual que no existe un concepto uniforme de justicia predictiva tampoco su aplicación es única, sino que adopta formas y finalidades diferenciadas según el ámbito en el que se despliegue. Sus funciones, riesgos y beneficios varían según el terreno en el que se aplique. El análisis comparado de los cuatro grandes órdenes jurisdiccionales —civil, penal, contencioso-administrativo y laboral— permite advertir tanto coincidencias estructurales (eficiencia, uniformidad, reducción de costes) como diferencias sustanciales derivadas de los bienes jurídicos en juego.

Así, en el ámbito civil, la justicia predictiva se orienta principalmente a la estimación de la viabilidad procesal. Los algoritmos analizan grandes volúmenes de sentencias con el propósito de anticipar la probabilidad de

2004, págs..761-767. (Disponible en línea) Competing Approaches to Predicting Supreme Court Decision Making | Perspectives on Politics | Cambridge Core <https://www.cambridge.org/core/journals/perspectives-on-politics/article/abs/competing-approaches-to-predicting-supreme-court-decision-making/12CB7F54E411F9EA-07DFAD>.

10. LAFONT NICUESA, L.: "Justicia predictiva de riesgos y derecho penal" en Ob. Col. Next Generation Justice: Digitalización e Inteligencia Artificial. CALAZA LÓPEZ, S.; ORDEÑANA GEZURAGA, I. (Dirs.), Ed. La Ley, 2024, pág.1134 y 1135. BATELLI, E.: "La justicia robótica: Algoritmos, interpretación y justicia predictiva" en Revista de Derecho Privado, n.º 38, 2020, págs.73-76.
11. HERNÁNDEZ DE LA FUENTE, J.: "Justicia predictiva, extensión de efectos y acciones de representación" en Ob. Col: Digitalización del servicio público de justicia e inteligencia artificial judicial. Op.Cit. Ed. Dykinson, 2024, pág. 114.

éxito de una demanda o de un recurso, así como de proyectar la actitud probable de los tribunales ante determinados supuestos. El efecto práctico más destacado es la racionalización de las estrategias procesales: las partes pueden ajustar sus expectativas, optar por acuerdos tempranos o recurrir a mecanismos alternativos de resolución de conflictos (ADR/ODR), reduciendo costes y tiempos.

En este contexto, el valor añadido de la predicción se vincula a la eficiencia procesal y a la reducción de litigiosidad innecesaria. No obstante, persisten riesgos asociados a la inflexibilidad de la jurisprudencia y a la pérdida de espacio para la innovación interpretativa, especialmente en aquellos casos en los que se requiera un cambio de criterio jurisprudencial o una interpretación creativa de la norma.

En la esfera penal, la justicia predictiva se presenta bajo una lógica distinta: la gestión del riesgo. Herramientas como *COMPAS* en Estados Unidos o *RisCanvi* en Cataluña calculan probabilidades de reincidencia, incumplimiento de medidas cautelares o quebrantamiento de la libertad condicional. A ello se suman los sistemas de *policía predictiva*, que buscan anticipar la comisión de delitos a partir de patrones espaciales, temporales o conductuales.

A diferencia del ámbito civil, donde la predicción actúa como mecanismo auxiliar de racionalización estratégica, en el ámbito penal la predicción afecta directamente a derechos fundamentales: la libertad personal, la presunción de inocencia, la igualdad y el principio de proporcionalidad.

El uso de algoritmos en decisiones sobre prisión preventiva, libertad condicional o medidas de seguridad plantea un dilema central: el riesgo de sustituir la valoración judicial individualizada por cálculos estadísticos que, además, suelen reproducir sesgos estructurales (raciales, socioeconómicos o de género).

En el ámbito contencioso-administrativo, la justicia predictiva permite identificar patrones en las sentencias de los tribunales, ofreciendo a las partes una estimación de éxito en impugnaciones de actos administrativos, lo que supone que cumple la función práctica de disuadir demandas con baja probabilidad de prosperar y fomentar acuerdos administrativos tempranos.

Por otra parte, la diversidad de criterios entre juzgados y tribunales contencioso-administrativos genera inseguridad. Si existieran algoritmos que sistematicen la jurisprudencia pueden servir como guías de coherencia,

esta función es particularmente valiosa en materias de urbanismo, sanciones administrativas o responsabilidad patrimonial del Estado.

Igualmente, la Administración puede utilizar modelos predictivos antes de dictar actos (p. ej. sanciones, resoluciones en contratación pública) para anticipar su probabilidad de ser anulados en sede judicial. Se refuerza así la eficiencia y la economía procesal, aunque con el riesgo de que la Administración se limite a “blindar” decisiones en función de patrones estadísticos más que de criterios de justicia material.

En la esfera laboral, la justicia predictiva permite anticipar los conflictos individuales y colectivos al estimar el resultado de demandas por despido, reclamaciones salariales o conflictos de representación sindical. Su función práctica es orientar a trabajadores y empleadores hacia acuerdos extrajudiciales, reduciendo la litigiosidad en jurisdicciones habitualmente saturadas.

También permite evaluar los riesgos empresariales ya que las herramientas predictivas pueden ser usadas por empresas para calcular la probabilidad de éxito en caso de litigio laboral, influyendo en decisiones de despido, negociación colectiva o externalización de servicios, aunque esta función plantea dilemas éticos como el riesgo de que la empresa utilice la predicción para maximizar su poder de negociación frente al trabajador.

Además, el derecho del trabajo está en constante transformación, sujeto a reformas legales y a una intensa jurisprudencia constitucional y supranacional, la justicia predictiva podría contribuir a sistematizar criterios y mejorar la seguridad jurídica en un ámbito caracterizado por frecuentes cambios legislativos.

En definitiva, en los cuatro órdenes jurisdiccionales la justicia predictiva busca reducir costes y agilizar la resolución de conflictos, así como conseguir una uniformidad jurisprudencial, especialmente relevante en lo civil y contencioso-administrativo, donde el exceso de criterios dispares afecta la seguridad jurídica¹².

12. Un ejemplo paradigmático del uso de técnicas estadísticas para proyectar decisiones judiciales se encuentra en la obra de KATZ, D.; BOMMARITO, M. y BLACKMAN, J. “A General Approach for Predicting the Behavior of the Supreme Court of the United States” donde los autores logran altos niveles de acierto en la predicción de sentencias mediante machine learning. KATZ, D.; BOMMARITO, M. y BLACKMAN, J. “A general approach for predicting the behavior of the Supreme Court of the United States” en A general approach for predicting the behavior of the Supreme Court of the United States | PLOS One <https://doi.org/10.1371/journal.pone.0174698>

Sin embargo, se plantean dificultades en cuanto a los derechos fundamentales, donde el riesgo es más grave en lo penal y lo laboral y donde las predicciones afectan directamente la libertad personal y la igualdad material.

También plantea problemas en cuanto a las relaciones de poder ya que en lo contencioso-administrativo, el peligro es reforzar la posición de la Administración frente al ciudadano; en lo laboral, la de la empresa frente al trabajador; en lo penal, el Estado frente al acusado; mientras que en lo civil la tensión es más equilibrada entre partes privadas.

Se puede decir, por tanto, que la justicia predictiva no es un fenómeno monolítico, sino un conjunto de aplicaciones que adquieren sentido distinto según el terreno en que se apliquen.

En lo civil y lo contencioso-administrativo, se vincula a la eficiencia y a la reducción de litigiosidad.

En lo penal y lo laboral, toca fibras sensibles del Estado de derecho, porque el uso de algoritmos puede reforzar desigualdades estructurales y afectar de modo directo a derechos fundamentales.

La clave, por tanto y como ya advertimos anteriormente, no es solo tecnológica, sino política y normativa: se trata de decidir en qué medida la predictividad puede ser un apoyo para la justicia sin convertirse en una nueva forma de ordalía estadística que sacrifique la deliberación humana en favor de una ilusión matemática de certeza.

La llamada *justicia predictiva* no solo debe aspirar a estimar comportamientos jurisdiccionales, sino que está intrínsecamente vinculada a principios de legitimidad, normatividad y justicia material. En este terreno, la predicción no se limita a describir lo que probablemente sucederá, sino que se enfrenta al desafío de justificar lo que debería suceder conforme a derecho.

De hecho, la búsqueda de previsibilidad ha acompañado históricamente a los sistemas jurídicos. El precedente judicial en la *common law* fue concebido como un mecanismo para dotar de continuidad y coherencia a las decisiones, de manera que los ciudadanos pudieran anticipar razonablemente las consecuencias jurídicas de sus actos. En el ámbito continental, la codificación decimonónica respondió a un impulso similar: ofrecer un marco normativo estable y sistemático que redujera la incertidumbre interpretativa. En ambos casos, el valor de la previsibilidad no radica en una capacidad técnica de predicción, sino en la construcción de confianza en la justicia como institución. Por ello, la incorporación de algoritmos al proceso de decisión judicial

debe leerse como una nueva etapa de esa tensión histórica: ampliar la previsibilidad sin sacrificar la legitimidad del juicio humano.

En este contexto, conviene advertir un peligro latente: la tentación de concebir los algoritmos predictivos como una suerte de “nueva codificación estadística”. Así como los códigos decimonónicos aspiraron a ofrecer un marco estable y exhaustivo del derecho, hoy los modelos algorítmicos pretenden condensar la experiencia jurisprudencial en fórmulas matemáticas capaces de anticipar decisiones futuras. Sin embargo, mientras la codificación clásica mantenía una apertura interpretativa, que permitía la evolución doctrinal y jurisprudencial, la estadística aplicada al derecho corre el riesgo de cristalizar patrones pasados y reforzar inercias interpretativas. Se sustituye así la seguridad jurídica por una mera *apariencia* de seguridad, sustentada en correlaciones numéricas, que puede terminar rigidizando el sistema y reduciendo su capacidad de adaptación a nuevos contextos sociales y normativos.

Este fenómeno recuerda, en cierto modo, la función de las ordalías medievales. También entonces el conflicto se resolvía mediante un procedimiento ritual que ofrecía la ilusión de certeza, desplazando la decisión humana hacia una instancia externa —la voluntad divina— que legitimaba el resultado sin necesidad de justificación racional. Hoy, el algoritmo puede ocupar el lugar de aquel hierro candente o de aquellas aguas hirvientes: un dispositivo al que se atribuye un valor probatorio o decisorio absoluto, pese a que su lógica interna resulta opaca para la mayoría y pese a que su autoridad descansa más en la fe (esta vez en la matemática) que en la deliberación crítica.

El peligro de esta “codificación estadística” reside, por tanto, en repetir el viejo gesto de las ordalías bajo un ropaje tecnológico: sustituir la incertidumbre propia del juicio humano por la confianza ciega en un mecanismo externo, ya no religioso, sino algorítmico. El riesgo es que, bajo la promesa de objetividad y neutralidad, se consoliden sesgos estructurales y se rigidice la interpretación de las normas, limitando la capacidad del derecho para adaptarse a los cambios sociales y tecnológicos.

En consecuencia, la incorporación de algoritmos al ámbito judicial no debería orientarse a producir una nueva ordalía estadística que clausure la deliberación, sino a constituir herramientas auxiliares que mejoren el acceso a la información, incrementen la coherencia de las decisiones y apoyen el trabajo de jueces y abogados, sin desplazar la dimensión crítica e interpretativa que constituye la esencia misma de la función jurisdiccional.

2.2. EL ALGORITMO COMPAS CORE: EL ORÁCULO MATEMÁTICO DE LA JUSTICIA PREDICTIVA

En este punto se impone el análisis de la justicia predictiva en la práctica, donde el derecho se ve cada vez más permeado por herramientas de inteligencia artificial orientadas a anticiparse a los problemas para lograr una solución más eficiente.

Lejos de ser un simple avance tecnológico, su presencia tiene consecuencias directas en la administración de justicia, la actividad policial y, de forma particular, en el proceso penal. Su uso revela una tendencia hacia la delegación de decisiones fundamentales en sistemas cuya lógica interna permanece, en gran medida, inaccesible. El estudio de tales herramientas se configura como una nueva forma de “ordalía algorítmica” en la que la fe en el cálculo sustituye al ritual medieval, y en la que el resultado, aunque revestido de científicidad exige la existencia de un acto de fe en un procedimiento que escapa al escrutinio ciudadano y jurídico y cuya legitimidad descansa más en la aceptación social que en la transparencia o la verificabilidad.

Los algoritmos predictivos de riesgo pueden definirse como modelos matemáticos orientados a la automatización de la predicción del comportamiento individual mediante la asignación de un índice de riesgo que clasifica al sujeto en diferentes niveles. En el ámbito de prevención de delitos lo que hace es “tratar de detectar patrones específicos de crimen y comportamientos criminales ha sido y es hoy una tarea extremadamente desafiante que exige dos cosas: el almacenamiento masivo de datos y un adecuado análisis para la extracción de inteligencia que oriente su correcta utilización. A día de hoy somos testigos del incremento exponencial de fuentes digitalizadas de información relacionada con la actividad delictiva”¹³.

En tales ámbitos se han incorporado algoritmos de evaluación y predicción de riesgo, en particular: Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) como el más destacado tanto por su utilización como por ser un sistema predictivo novedoso, así como por los estudios que sobre él se han realizado¹⁴.

-
13. MIRÓ LINARES, F.: “Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots” en Revista de Derecho Penal y Criminología, n.º 20, 2018, pág. 98.
 14. Existen otros sistemas predictivos de riesgos que son utilizados en el ámbito del proceso penal como: PSA (Public Safety Assessment) que evalúa si un procesado puede esperar en libertad el juicio, en función del riesgo de reincidencia o incomparecencia. LSI-R (Level of Service Inventory-Revised), usado en permisos y libertad condicional, pondera factores como antecedentes, entorno familiar, educación, empleo, consumo de

El sistema COMPAS, desarrollado por la compañía Northpointe Inc. (ahora Equivant) de EEUU¹⁵ en 1998 (revisado a lo largo de los años), se basa en la realización de una serie de cálculos que establece la posibilidad de que un individuo reincida en la comisión de un delito.

El algoritmo tiene en cuenta factores como la edad, el género, el historial delictivo, la gravedad de los delitos cometidos, la no comparecencia ante un tribunal, cargos actuales, cargos pendientes, la estabilidad en cuanto a residencia y los lazos con la comunidad, el consumo de sustancias y la situación laboral. Además, se realiza una encuesta al delincuente en la que se formulan cuestiones relacionadas con su salud, educación, ocio, problemas familiares etc...¹⁶ y una entrevista personal para realizar una estimación más aproximada de las posibilidades de reincidencia.

En otras palabras: "La herramienta COMPAS pretende predecir el riesgo de reincidencia de un individuo sobre la base de un análisis complejo que implica el uso de la información obtenida de una encuesta de 137 preguntas dividida en varias secciones diferentes, así como de la información correspondiente a los antecedentes penales individuales"¹⁷.

Esos datos se introducen en el sistema para que el algoritmo establezca una puntuación donde 1 es la puntuación más baja y 10 la más alta.

drogas y actitudes personales. Actualmente se utilizan en Estados como Washington y California.

CAS (Crime Anticipation System) que anticipa delitos de alto impacto en zonas específicas, usando datos delictivos previos y variables socioeconómicas y demográficas; se basa en teorías de criminología ambiental.

NDAS (National Data Analytics Solution) que identifica riesgos de delitos violentos con armas y posibles víctimas, a partir de registros policiales, ofreciendo servicios sociales o de salud a sospechosos para prevenir delitos.

HART (Harm Assessment Risk Tool), desarrollado en el Reino Unido, su función es predecir el nivel de riesgo de los individuos de cometer delitos en un lapso de dos años, usando criterios como edad, sexo e historial de delitos; este algoritmo ayuda a determinar si el individuo que tiene un riesgo bajo puede ser enviado por los agentes de policía a un programa de rehabilitación denominado Checkpoint.

15. Este sistema se ha utilizado en varios Estados: Nueva York, Wisconsin, California y Condado de Broward en Florida, entre otros.
16. ROA ABELLA, M.D. P.; SANABRIA-MOYANO, J.E.: "Uso del algoritmo COMPAS en el proceso penal y los riesgos a los derechos humanos" en Revista Brasileira de Direito Processual Penal, Vol. 8, n.º 1, 2022. Págs. 275-310.
17. ROMEO CASABONA, C.M.: "Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad" en Revista de Derecho, Empresa y Sociedad, n.º 13. 2018, pág. 49.



ESTUDIOS

Acceso online a Biblioteca Digital Legalteca:
consulte página inicial de esta obra

La obra ofrece a los profesionales del Derecho y del ámbito digital un marco para evaluar la fiabilidad, la legitimidad y los límites probatorios de los sistemas automatizados, y resulta útil porque no se limita a denunciar fallos técnicos, sino que desentraña los supuestos jurídicos y culturales que permiten que un algoritmo opere como si fuera un oráculo. Al mismo tiempo aborda un problema incómodo: la creciente dependencia de sistemas opacos para decidir sobre derechos, riesgos y responsabilidades.

En esta investigación se identifica un patrón histórico, delegar el juicio en artefactos cuya autoridad no se cuestiona, y muestra cómo ese gesto reaparece hoy en la automatización de decisiones públicas y privadas. Examina la noción de *fe tecnológica* como riesgo regulatorio y plantea métodos de auditoría jurídica que no dependen de la transparencia total del código. Es una invitación a recuperar el control normativo allí donde la técnica amenaza con convertirse en dogma.

¿Cómo controlar decisiones automáticas en procedimientos administrativos? ¿Qué valor probatorio podría tener un modelo predictivo? ¿Cómo evitar que la «confianza ciega» en la tecnología sustituya al juicio jurídico? ¿Qué ocurre cuando los sesgos de datos generan exclusiones sistemáticas sin responsables identificables? ¿Qué Derechos se ven afectados? ¿Cuál es la orientación predominante de la regulación y de la jurisprudencia en la actualidad, y hacia qué dirección parecen proyectarse sus desarrollos futuros?

ISBN: 978-84-1085-598-4



9 788410 855984



ER-0280/2005



GA-2005/0100