



***MÁSTER UNIVERSITARIO EN ACCESO A LA
PROFESIÓN DE ABOGADO POR LA UNIVERSIDAD DE
CANTABRIA
(EN COLABORACIÓN CON EL ILUSTRE COLEGIO DE
ABOGADOS DE CANTABRIA)***

TRABAJO FIN DE MÁSTER

CURSO ACADÉMICO 2022-2023

La Justicia Predictiva: Límites y Usos Procesales

Predictive justice: Limits and Procedural Uses

**AUTOR:
Íñigo Gutiérrez Gómez**

**DIRECTOR:
José Ignacio Solar Cayón**

“En aquel Imperio, el Arte de la Cartografía logró tal Perfección que el mapa de una sola Provincia ocupaba toda una Ciudad, y el mapa del Imperio, toda una Provincia. Con el tiempo, estos Mapas Desmesurados no satisficieron y los Colegios de Cartógrafos levantaron un Mapa del Imperio, que tenía el tamaño del Imperio y coincidía puntualmente con él”

Jorge Luis Borges (1974), *Del Rigor En la Ciencia*

ÍNDICE

| | |
|---|--------|
| INTRODUCCIÓN..... | - 4 - |
| CAPÍTULO I: INTELIGENCIA ARTIFICIAL, HISTORIA Y MARCO LEGAL | - 6 - |
| 1.1 ¿UNA DEFINICIÓN DE INTELIGENCIA ARTIFICIAL? | - 7 - |
| 1.2. APROXIMACIÓN HISTÓRICA A LA INTELIGENCIA ARTIFICIAL | - 9 - |
| 1.2.1 La primavera de la IA: “grandes esperanzas” | - 9 - |
| 1.2.2 El (primer) invierno de la Inteligencia Artificial..... | - 11 - |
| 1.2.3 El segundo invierno y el otoño | - 12 - |
| 1.2.4 La actualidad: <i>machine learning</i> , <i>deep learning</i> y <i>big data</i> | - 13 - |
| 1.3 APROXIMACIÓN LEGAL A LA INTELIGENCIA ARTIFICIAL | - 15 - |
| 1.3.1 Regulación de la IA en España | - 16 - |
| 1.3.2 Regulación de la IA en Europa..... | - 18 - |
| CAPÍTULO II: INTRODUCCIÓN A LA JUSTICIA PREDICTIVA..... | - 23 - |
| 2.1 DEFINICIÓN DE JUSTICIA PREDICTIVA: EL SUEÑO DEL HOMBRE MALO | - 24 - |
| 2.2 ALGUNAS EXPERIENCIAS DE JUSTICIA PREDICTIVA | - 27 - |
| 2.1.1 La predicción de sentencias de la Corte Suprema de EE.UU. y del TEDH | - 27 - |
| 2.1.2 Usos comerciales de la justicia predictiva: | - 30 - |
| 2.1.3 Usos de justicia predictiva en la Administración de Justicia | - 31 - |
| CAPÍTULO III: PREDICCIÓN, INTELIGENCIA ARTIFICIAL Y PROCESO | - 35 - |
| 3.1 UNA APROXIMACIÓN REALISTA: LOS LÍMITES DE LA IA | - 37 - |
| 3.1.1 Límites técnicos | - 38 - |
| 3.1.2 Límites jurídicos | - 40 - |
| 3.1.3 Límites técnico-jurídicos: la transparencia | - 43 - |
| 3.1.4 Límites de legitimidad | - 45 - |
| CAPÍTULO IV: POSIBLES USOS DE LA INTELIGENCIA ARTIFICIAL EN EL PROCESO..... | - 47 - |
| 4.1 ÁMBITO DE LA TRAMITACIÓN | - 47 - |
| 4.1.1 Usos presentes en la Administración de Justicia..... | - 48 - |
| 4.1.2 Usos futuros en la Administración de Justicia | - 50 - |
| 4.2 ÁMBITO PROCESAL Y/O PROBATORIO: | - 52 - |
| 4.2.1. Los sistemas de codificación predictiva..... | - 52 - |
| 4.2.1. La IA y los medios de prueba y de evaluación de la prueba..... | - 54 - |
| 4.2.2 Los sistemas de evaluación de riesgos de reincidencia criminal | - 56 - |
| 4.2.3 Sistemas de búsqueda de información jurídica | - 61 - |
| 4.3 ÁMBITO DE LAS DECISIONES JUDICIALES | - 61 - |
| CONCLUSIONES..... | - 65 - |
| BIBLIOGRAFÍA..... | - 68 - |

INTRODUCCIÓN

La inteligencia artificial es el tema del futuro, y por eso inunda el presente. Su implantación en los Tribunales, que parece inevitable, puede ser acogida con exceso de entusiasmo por aquellos que anhelan un cambio o de escepticismo por quienes ven la máquina como una amenaza, es decir, visiones tecnofóbicas o tecnófilas que son causa de los reduccionismos que rápidamente afloran al surgir la discusión de las tecnologías disruptivas. En medio de la pugna entre estas visiones, la inteligencia artificial en el Derecho avanza, sembrando el descontento tanto entre quienes quieren que se ralentice o se revierta como entre quienes defienden su acelerada adopción.

El debate sobre la aplicación que han de tener otras disciplinas extrañas en el Derecho no es novedoso. Desde la geometría jurídica de Gottfried Wilhelm Leibniz, con su *Codex Leopoldus*, se ha intentado promover la autonomía del Derecho, haciéndolo impermeable al exterior mediante la lógica, es decir, entendiéndolo como un fin en sí mismo. En contra de esta idea se han posicionado juristas como Oliver W. Holmes, conceptualizando el Derecho como medio para alcanzar fines sociales. La llegada de internet y el aumento de la capacidad de tratamiento de datos parece querer aunar estas posturas, instaurando un sistema mixto: ni deducción, ni profecía, sino probabilidad. Es la denominada “justicia predictiva”.

En principio, gran parte de la literatura actual se refiere a estos avances mediante la remisión al “Juez-robot” o al “abogado-robot”. Sin embargo, tras estos grandilocuentes titulares se encuentran más bien herramientas que, sin sustituir al juez, al auxiliar o abogado, le ayudan en su actividad, aumentando su productividad. Los beneficios son innegables, contribuyendo a la Justicia formal mediante la eficiencia de Tribunales y abogados. No obstante, visto el uso de los algoritmos en perspectiva comparada, hay que tratar también los riesgos y límites de tales herramientas. La actividad de los Tribunales y los abogados sirve para tratar conflictos humanos mediante reglas que, para ser justas, requieren sintaxis, pero también semántica ¿puede considerarse Justicia únicamente un resultado?

Si bien aplicada la inteligencia artificial puede aumentar la confianza y el acceso del conjunto de la ciudadanía a servicios jurídicos, la banalización de la Justicia puede llevar

a la vulneración de derechos fundamentales, como la tutela judicial efectiva, el juez objetivo predeterminado por ley, la presunción de inocencia o la privacidad. Estos daños ocurren a menudo, no de forma dolosa, sino por el desconocimiento de los límites de lo que la máquina puede hacer, así como por los errores humanos en interpretar su *output*, confusión quizá alentada por el -justificado- entusiasmo de sus desarrolladores. En este trabajo se verán varias herramientas que sirven para preparar las sentencias y asistir a los Tribunales a la hora de tramitar y enjuiciar, siempre intentando no caer en el lamento “solo es inteligencia artificial si no funciona”. Acercarse a la realidad actual, sin remitirse al futuro, puede ser un itinerario suficientemente arduo.

CAPÍTULO I: INTELIGENCIA ARTIFICIAL, HISTORIA Y MARCO LEGAL

Actividades tan variadas como invertir en renta variable, conducir o escribir, pueden ser asistidas o directamente realizadas por una máquina que emplea inteligencia artificial (IA en adelante). No es inusual en la literatura especializada que, como *captatio*, se expongan los ejemplos de los logros de la IA -que no son pocos-, pese a que no nos encontramos ante casos aislados, sino ante lo que es un “tsunami” con implicaciones estructurales que van más allá, afectando en el plano social. Actualmente estamos inmersos en la Cuarta Revolución Industrial¹, que está cambiando el entendimiento del tiempo y el espacio. Además, mediante el aumento de su capacidad, las tecnologías de la información y la comunicación (TIC), han logrado el tratamiento de datos disponibles en internet, logrando que la IA alcance a realizar tareas complejas, antes exclusivas de la humanidad.

Sin embargo, para que la IA sea implementada en el ámbito del Derecho tiene algún dique que rebasar. Cualquier espectador de la labor de los tribunales puede vislumbrar la falta de financiación que existe para implementar las más básicas tecnologías. También se ha alegado el -supuesto- celo por las tradiciones de los operadores jurídicos. De hecho, si se pregunta al ciudadano profano sobre la opinión que tiene de la Administración de Justicia, le suele parecer un “abstruso arcano de no fácil descodificación” (CGPJ, 2021). Sin embargo, en una perspectiva interna, algo parece estar cambiando entre los miembros más jóvenes de la carrera judicial. Es común en ellos la queja de que la tecnología esta infradotada o no recibe un empleo correcto en el funcionamiento de los Tribunales².

Desde la lógica sociológica de la teoría de sistemas, Luhmann considera el Derecho como un (sub) sistema, abierto ante la complejidad exterior, pero cerrado en cuanto a las fórmulas que usa y los agentes que lo componen (Ruiz Manotas & Bermeo Álvarez, 2018). Esta característica, que le aísla del mundo gracias a la complejidad, puede explicar su resistencia a adoptar fórmulas o incluir agentes de otras disciplinas, pues solo opera en el ámbito binario (jurídico o antijurídico) (Montoro Ballesteros, 2007). Sin embargo, el

¹ Como veremos en adelante, este progreso es desigual pues aproximadamente el 17 % de la población mundial, 1.300 millones de personas, no ha experimentado la segunda revolución industrial. Y más de la mitad de la población mundial, 4.000 millones de personas, todavía no está conectada a internet (Schwab, 2016)

² Es común entre los Jueces y Magistrados de menor edad encontrar deficientes los medios de que dispone la Administración de Justicia, por ejemplo, un 59% entre la franja de edad de 30-35 años y un 58% entre las de 36-40 años, consideran inadecuada la formación que da el Consejo General del Poder Judicial (CGPJ) en materia tecnológica (CGPJ, 2015)

sistema está formado por personas, que ineludiblemente se ven afectadas por el cambio social. Por ello, la implantación de la IA como elemento externo al Derecho conlleva el reto de intentar definir qué es, desde el papel del jurista, para conocer qué deparará su inclusión en el sistema legal.

1.1 ¿UNA DEFINICIÓN DE INTELIGENCIA ARTIFICIAL?

Si bien adquiere cada vez mayor importancia, alcanzar una definición de IA es todo un reto. En primer lugar, existe una definición incluso en el diccionario, que considera la IA como aquella “disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico” (RAE, 2022), pero la realidad es más complicada. Se trata de un campo de estudio en ebullición, que surge de la confusión con otras disciplinas, no existiendo una definición de trabajo, es decir, una definición entre los expertos inmersos en proyectos de investigación en IA que sea común para todos ellos (Wang, 2019). La IA se puede dividir a su vez en subcampos o áreas de muy diversa índole, entre ellas la representación del conocimiento y razonamiento, el aprendizaje automático, el procesamiento del lenguaje natural (PLN), el reconocimiento de imágenes, la robótica y el reconocimiento automático del habla (Wollowski et al., 2016).

Antes de abordar el concepto en toda su amplitud, se ha de conocer cómo funciona, mediante un “conjunto definido de reglas o procesos para la solución de un problema en un número finito de pasos” (Salazar García & Benjamins, 2020, p. 270). Esto es, los llamados algoritmos.

Debido a la multitud de definiciones y la dificultad de encontrar una completa, más allá de la incluida en el diccionario, los autores Stuart J. Russell y Peter Norvig (2004) han creado una matriz que recoge elementos de las distintas sensibilidades, que coinciden en definir la IA como máquinas que piensan y/o actúan como humanos o racionalmente. Esta distinción se centra en dos grupos de autores: aquellos que buscan conseguir una inteligencia de la máquina similar a la humana, que siguen el denominado modelo cognitivo, y los que entienden que la falta de emociones de la máquina difiere de la inteligencia humana, siguiendo un tipo de inteligencia llamado “racionalidad” (Russell & Norvig, 2004 p. 4).

Como se ve, el óbice para definir la IA es la “inteligencia”, inherente a la concepción de inteligencia humana, que ha evolucionado desde finales del siglo XIX de ser una

mensurable cuantitativamente a múltiples inteligencias, algunas difíciles de cuantificar³. Sin afán de abarcar todas las teorías sobre la inteligencia, por ejemplo, para Robert J. Sternberg (1985), defensor de una visión triárquica de inteligencia, esta puede dividirse en creativa, práctica y analítica. La inteligencia creativa se centra en dar a luz ideas nuevas e innovadoras a partir de la conexión subjetiva con el conocimiento; la práctica supone la adaptabilidad en el entorno para convertir la teoría en práctica; y la analítica se refiere a la capacidad de captar y retener información para resolver problemas, estando esta última íntimamente relacionada con la IA. Siguiendo la noción difusa de inteligencia, y la matriz dada por Russel y Norvig, hemos de distinguir, al menos, dos enfoques tradicionales sobre la IA: como inteligencia “fuerte” o inteligencia “débil” (Searle, 1980). Distinción a la que el filósofo Nick Bostrom ha añadido una última etapa futurista, la “superinteligencia” (Bostrom 2014). En general podemos distinguir:

- Inteligencia Artificial especializada o IA débil. Esta IA realiza tareas y actividades específicas que requieren “inteligencia”, como jugar al ajedrez, responder a preguntas en un programa de televisión, conducir un coche, etc.
- Inteligencia Artificial general, también conocida como IA fuerte. Aunque “general” y “fuerte” no son estrictamente lo mismo, toda IA fuerte ha de ser general (Meseguer González & López de Mántaras, 2017). Sería una IA capaz de emular la inteligencia humana, pudiendo aprender y actuar de una manera indistinguible a esta.
- Inteligencia Artificial superinteligente, que excedería por mucho las mejores mentes humanas, dividiéndose en “superinteligencia” de velocidad, colectiva o de calidad (Bostrom, 2016, pp. 52-27).

Hoy en día, los mayores avances vienen de la IA “débil” o especializada. Débil, porque la inteligencia de la IA es diferente respecto a la forma de pensar humana, no restándole por eso utilidad. Según Richard Feynman, no es necesario comprender la forma en que las aves baten sus alas y cómo están diseñadas sus plumas para crear una máquina voladora” (Feynman, 2008, p. 110). Y especializada, porque depende del ser humano para

³ Entre finales del siglo XIX y principios del siglo XX psicólogos como Francis Galton, James McKeen Cattell y Wilhelm Stern, que buscaban medir la inteligencia, acuñaron términos como "cociente mental" o "cociente intelectual". Con el transcurso del tiempo fue quedando atrás el concepto único de inteligencia y se comenzó a distinguir otros tipos de inteligencia, como la teoría de las inteligencias múltiples propuesta por Howard Gardner (1983) quien sostiene que existen diferentes tipos de inteligencia, como la lingüística, la lógico-matemática, la espacial, la musical y la interpersonal, entre otras.

ser orientada en su labor o tarea específica, es decir, la IA es *task-oriented* (Solar Cayón, 2021 p. 7).

Para una comprensión mejor de lo que es la IA se ha de trazar brevemente su historia hasta la actualidad (una aproximación histórica), así como analizar su tratamiento en las normas jurídicas (una aproximación normativa), hallándonos hoy, como veremos, inmersos en plena labor legislativa sobre la materia.

1.2. APROXIMACIÓN HISTÓRICA A LA INTELIGENCIA ARTIFICIAL

Para abordar la historia de la IA es frecuente el uso recursivo de las estaciones del año, aunque el periodo abarcado por cada una de ellas es objeto de controversias entre los historiadores (Haenlein & Kaplan, 2019). Pese a ello, todos coinciden en diferenciar dos tendencias alternas que se suceden en el tiempo a lo largo de la existencia de la IA: una optimista, durante la cual los avances producidos llevan a una mayor inversión -la denominada “primavera de la IA”-, y otra pesimista -el denominado “invierno”-, en el que la inversión se estanca o desciende por la falta de los resultados esperados.

1.2.1 La primavera de la IA: “grandes esperanzas”

Para encontrar el inicio de la IA ha de tomarse como referencia la década de 1940, con la creación de los primeros ordenadores modernos, destinados entonces al esfuerzo bélico, como el ENIAC (*Electronic Numerical Integrator and Computer*), para el cálculo de tablas balísticas, o el célebre “Bombe” de Alan Turing y el posterior “Colossus” de Tommy Flowers, diseñados con objeto de descifrar las comunicaciones enemigas. Turing, una de las figuras más importantes en el campo, diseñó además una máquina imaginaria, conocida como la Máquina de Turing, con finalidad de explicar que todos los ordenadores son equivalentes en su lógica (Hawkins & Blakeslee, 2005).

Coetánea a estos inventos es la obra de Warren McCulloch y Walter Pitts (1943), que constituye la base de las redes neuronales artificiales (RNA) o sistemas conexionistas, y es considerada una de las primeras incursiones en el campo común de la cibernética y de la IA. Posterior es el trabajo de Norbert Wiener, que con su obra *Cybernetics* (1949) delimita el campo de estudio de la primera en la teoría del control y la comunicación, tanto de sistemas biológicos como mecánicos. Gran parte de los autores consideran la cibernética como el origen de la IA, separándose posteriormente ambas disciplinas en lo que se podría llamar una ruptura “dulce” (McCorduck, 2004, p. 113; R. Kline, 2011, p.

6). El horizonte de esta disciplina común se amplía también a finales de la década, cuando se crean los primeros robots cibernéticos, como el homeostato de Ross Ashby, que perseguía mantener un equilibrio electromagnético soportando un amplio abanico de alteraciones (homeostasis), o las tortugas mecánicas de Grey Walter, apodadas “*machina expeculatrix*”, que se recargaban de manera automática cuando su batería se descargaba.

Por entonces se produce el que es considerado el primer hito fundacional de la IA, con la publicación del artículo de Turing titulado *Computing Machinery and Intelligence* (1950), en el que se dieron a conocer por primera vez “el aprendizaje automático, los algoritmos genéticos y el aprendizaje por refuerzo” (Russell & Norvig, 2004, p. 20). Turing conceptualiza, mediante su prueba, lo que pasaría a ser la IA “fuerte”. La prueba consiste en un juego de imitación en el que un interrogador hace una serie de preguntas durante cinco minutos para descubrir si su interlocutor es realmente un humano o una máquina, pasando la prueba la máquina si el interrogador cree que está hablando con una persona al menos el 70% del tiempo. El conocido matemático puso como horizonte para que tal cosa sucediera el año 2000. Sin embargo, hoy en día, aun cuando reina el entusiasmo por los agentes conversacionales o *chatbots*⁴, la profecía no se ha cumplido (Russell & Norvig, 2004, p. 1077).

Tomando en cuenta los precedentes en el campo de la computación, la cibernética y la robótica, en 1955, John McCarthy logra convencer en un principio a diez participantes, escribiendo a la Fundación Rockefeller en busca de financiación, para celebrar un seminario de verano sobre la IA en Dartmouth, Nuevo Hampshire (R. Kline, 2011, p. 8). El congreso se llamaría “*Dartmouth Summer Research Project on Artificial Intelligence*” y su objeto sería discutir “la conjetura de que cada aspecto del aprendizaje o cualquier otra característica de la inteligencia puede, en principio, describirse con tanta precisión que se puede hacer una máquina para simularlo” (McCorduck, 2004, p. 111). El congreso se celebró finalmente en verano de 1956 y, aunque, según McCarthy, esa no era la primera vez que se hablaba de inteligencia artificial, la Conferencia de Dartmouth hizo que ese concepto se impusiera sobre todos los demás (McCorduck, 2004, p. 115).

Los críticos ven en la noción de IA un triunfo de una marca diseñada para vender, con muy buen resultado, una idea distorsionada de las capacidades de un ordenador (Acosta,

⁴ Un chatbot, es un “software de aplicación que proporciona un entorno de comunicación para la conversación entre máquina y humanos utilizando un lenguaje natural” (Srinivasa, Mounika, & Fareed, 2020).

2019). Y es que dicho concepto ni siquiera obtuvo consenso cuando se propuso en Dartmouth. Allen Newel y Herbert Simon se decantaban por hablar de “procesamiento complejo de información” (McCorduck, 2004, p. 115), una definición más ajustada a la realidad, pero sin duda no tan llamativa. La IA cosechó grandes éxitos en sus inicios, como el programa ELIZA, capaz, mediante la formulación de preguntas, de hacerse pasar, con relativo éxito, por un psicoterapeuta. En esta primera etapa los investigadores se dedicaron mayoritariamente a rebatir los argumentos de quienes negaban la posibilidad de una IA bajo la premisa de: “¡ninguna máquina podrá jamás hacer X!” (Bostrom, 2016, p. 5).

1.2.2 El (primer) invierno de la Inteligencia Artificial

Tras la primera serie de éxitos, la IA experimenta un declive donde predominan corrientes pesimistas. No existe un consenso sobre cuándo se inicia este invierno, si a mediados de la década de 1960 o a principios de la década siguiente (Bostrom, 2016 p. 7). Entre las razones de esta pérdida de interés se encontraba la sobrestimación de las capacidades de la IA, que viene a denunciarse en un conocido informe de James Lighthill (1973) en el que se reconocía la incapacidad de la IA del momento para afrontar la “explosión combinatoria”⁵. Entre las figuras críticas, cabría detenerse brevemente en algunos autores como Ulric Neisser, Hubert Dreyfus o John Searle.

El conocido como padre de la psicología cognitiva, Ulric Neisser, fue uno de los primeros críticos de la IA fuerte en su artículo *The Imitation of Man by Machine* (1963). Para el autor, el hecho de defender que las máquinas podían pensar como humanos se debía a una interpretación errónea de la naturaleza del razonamiento humano. Neisser distingue las máquinas de los humanos en base a tres características: i) el pensamiento humano siempre sucede de forma introspectiva, estando involucrado un proceso acumulativo de madurez y desarrollo; ii) el pensamiento humano tiene sus orígenes en íntima asociación con las emociones y sentimientos, que nunca se dejan de presentar y iii) toda actividad humana, incluido el pensamiento, sirve no a uno, sino a una multiplicidad de motivos al mismo tiempo (Neisser, 1963, p.195).

⁵ La “explosión combinatoria”, como su nombre indica, supone un límite cuantitativo a la hora de afrontar la resolución de problemas. Ante un número casi infinito de combinaciones, la máquina tardaría un largo periodo de tiempo en llegar a una respuesta. Sin embargo, hoy existen estrategias avanzadas para enfrentarse a ella, como aproximaciones, heurísticas y el uso de poda (Meseguer González & López de Mántaras, 2017, p. 57)

Un poco después, en la publicación *Alchemy and Artificial Intelligence* (1965), el filósofo Hubert Dreyfus comparó los objetivos de la IA con los intentos de transmutar el plomo por parte de los antiguos alquimistas. En esta obra Dreyfus rebate las hipótesis asociacionistas que defienden que el pensamiento humano puede ser descompuesto en operaciones lógicas simples⁶. Aunque Dreyfus no se opone a que se haya de seguir estudiando, niega cualquier esperanza en la creación de máquinas inteligentes. Posteriormente, Dreyfus se abre a la posibilidad de una máquina corpórea como necesidad esencial para alcanzar la “inteligencia” (Dreyfus, 1992, pp. 235 y ss.).

En 1980, John Searle publicó un artículo titulado *Minds, Brains and Programs*, que contenía una alegoría objeto de gran discusión entre partidarios y detractores de la IA. Para Searle, “comprender un lenguaje, o ciertamente, tener estados mentales, incluye algo más que tener un puñado de símbolos formales” (Searle, 1985, p. 39). La alegoría que usa para explicarlo es la denominada “habitación china”. Una persona que no habla chino se encuentra encerrada en una habitación en la que hay unos caracteres escritos en chino. Para ella, estos caracteres suponen una serie de garabatos sin sentido, pero tiene consigo un pliego de instrucciones en su lengua materna (un símil del algoritmo) que le indica como ordenar estos caracteres. Para el hablante de chino que se encuentre fuera de la habitación, la respuesta dada desde dentro puede ser inteligible e incluso inteligente, pero, al igual que la persona encerrada, la máquina no comprende chino (Searle, 1980, p. 418).

1.2.3 El segundo invierno y el otoño

El deshielo del primer invierno se produjo durante la década de 1980, cuando Japón comenzó a trabajar en el Proyecto de Sistemas Computacionales de Quinta Generación, arrastrando a otros países a retomar la inversión en proyectos de IA (Bostrom, 2016, p. 7). Este es el periodo de auge de los denominados sistemas expertos de segunda generación, que ya habían dado lugar a sistemas como Heuristic Dendral, en la geología, o Mycin, en la medicina. Un sistema experto podría definirse como el “sistema informático que simula el proceso de aprendizaje, de memorización, de razonamiento, de comunicación y de acción de un experto humano en una determinada rama de la ciencia” (Amador Hidalgo, 1996, p. 138). Para simular al experto, la máquina usa reglas lógicas

⁶ Para una definición de asociacionismo vid. Manuel Carabantes López (2014): “el asociacionismo establece que los estados o contenidos mentales se enlazan unos con otros de tal manera que la aparición de uno determina la presencia del siguiente en virtud de su semejanza, contraste o contigüidad” (pp. 85-86).

del tipo “SI-ENTONCES”, la conocida como *knowledge based approach* o GOFAI (*Good Old-Fashioned Artificial Intelligence*).

Pese a ello, el entusiasmo por los sistemas expertos decae tras el fracaso del proyecto japonés y las limitaciones de estos sistemas para codificar la realidad⁷. Tras este periodo viene un segundo invierno precedido de la obra *The emperor's new mind: concerning computers, minds, and the laws of physics* (1989) del físico Roger Penrose. Penrose aborda en esta obra algunas críticas anteriores a la IA “fuerte”, comentando sus debilidades y defendiendo su propia aproximación crítica: que nunca se podrá llegar a la IA “fuerte” en base al teorema de incompletitud de Gödel. Este teorema, de gran complejidad, que en verdad son dos subordinados, se puede resumir en la idea de que “hay sentencias aritméticas que son verdaderas y que no se pueden demostrar” (Russell & Norvig, 2004, p. 143). Para superar a Gödel, Penrose propone una “nueva Física”, aunque no existe consenso actual sobre cómo llegar a ese fin (Meseguer González & López de Mántaras, 2017, p. 55).

El optimismo se retomará de nuevo gracias a técnicas que en apariencia suponían una alternativa al paradigma “logicista tradicional” de la GOFAI (Bostrom, 2016, p. 7). Desde principios de 1990 hasta 2015, la IA ha ido progresando en lo que se ha conocido como el “otoño” (Haenlein & Kaplan, 2019). Se ha ido recogiendo la cosecha del desarrollo tecnológico que ha experimentado la tecnología de semiconductores, expresado en la conocida “ley de Moore” formulada por Gordon Moore en 1965. Esta ley afirma la duplicación de la capacidad de los chips en un ciclo -con mayor o menor desviación- de 18 meses⁸. Es en este otoño cuando se retoma la investigación en redes neuronales artificiales (RNA) -que había sido abandonada con el primer invierno, tras la obra *Perceptrons* (1969) de Marvin Minsky y Seymour A. Papert- mediante el “nuevo conexionismo” (Schwartz, 1988), constituyendo un hito en esta evolución la victoria en 2015 del programa de IA *AlphaGo* frente al campeón mundial del juego de mesa Go, en lo que se ha denominado el “Sputnik” de nuestro tiempo (Schiavenza, 2018).

1.2.4 La actualidad: *machine learning*, *deep learning* y *big data*

⁷ La principal limitación es el denominado “cuello de botella de adquisición del conocimiento” (*knowledge acquisition bottleneck*), término acuñado por Ed Feigenbaum en 1973 para referirse a los problemas de obtención y uso del conocimiento por parte de las máquinas (Casanovas, 2015).

⁸ Cuando Moore formuló su teoría esta consistía en que el número de transistores por centímetro cuadrado en un circuito integrado se duplicaba cada año, aunque posteriormente, en 1975, se rebajó el ritmo de evolución a la duplicación cada 18 meses (Cheang Wong, 2005, pp. 4-5).

Con la evolución de la IA pierde protagonismo la primera idea de codificar en lenguaje formal el razonamiento humano mediante reglas del tipo SÍ-ENTONCES (Janiesch, Zschech, & Heinrich, 2021, p. 685). En la actualidad, aunque siguen existiendo y son de utilidad los sistemas GOFAI, el protagonismo lo está tomando otro tipo de IA basada en el procesamiento y análisis de grandes volúmenes heterogéneos de datos, denominados *Big data* o macrodatos⁹, para encontrar correlaciones, permitiendo “predecir” hechos presentes, pasados y futuros (Agrawal, Gans, & Goldfarb, 2019, p. 28). El papel de los macrodatos en las predicciones de la IA sería análogo al del combustible para un vehículo.

El aumento de la capacidad de procesamiento de datos y análisis que sirve para generar predicciones, reglas, respuestas o recomendaciones no se podría entender sin el desarrollo del *machine learning* o aprendizaje automático. El aprendizaje automático, que se encuentra en la intersección entre la informática y la estadística, se basa en algoritmos que mejoran su rendimiento a medida que van adquiriendo experiencia en el desarrollo de una tarea, pudiendo encontrar relaciones ocultas entre datos sin estar explícitamente programados para ello (Mitchell & Jordan, 2015, p. 255). La principal diferencia entre las técnicas de aprendizaje automático y los métodos estadísticos clásicos, como la regresión, es que el “aprendizaje automático implica probar qué funciona mejor en la práctica”. Por el contrario, “inventar un nuevo método de regresión requiere primero probar qué funciona en la teoría” (Agrawal et al., 2019, p. 41).

⁹ El *Big data* o macrodatos suele definirse mediante las denominadas tres V: volumen, variedad y velocidad. En su obra *Principles of Big Data: Preparing, Sharing, and Analyzing Complex Information* (2013) Jules Berman distingue diez diferencias entre macrodatos y datos a pequeña escala en cuanto a su: i) Meta: los datos a pequeña escala a menudo se recopilan con un propósito específico, mientras que los macrodatos pero pueden ser usados para múltiples propósitos; ii) ubicación: los datos a pequeña escala suelen almacenarse en una única ubicación, mientras que los macrodatos pueden distribuirse en varios archivos y servidores en diferente ubicación; iii) estructura: los datos a pequeña escala están altamente estructurados, mientras que los macrodatos pueden ser no estructurados y tener muchos formatos de archivo diferentes y enlaces a otros recursos; iv) preparación de los datos: los datos a pequeña escala a menudo se preparan por el usuario final para su propósito, mientras que los macrodatos son preparados por un grupo, analizados por un segundo grupo y utilizados por un tercer grupo; v) longevidad: los datos a pequeña escala suelen mantenerse durante un período de tiempo específico mientras que los macrodatos tienen una vida más larga y más incierta debido a su alto costo. Pueden actualizarse continuamente con nuevos datos, factores contextuales del pasado, variables adicionales o enlaces a otros conjuntos de datos; vi) medición: los datos a pequeña escala suelen medirse utilizando un único protocolo con un conjunto de unidades todas a la vez, mientras que los macrodatos pueden medirse utilizando diferentes protocolos por personas en diferentes ubicaciones vii) reproducibilidad: los conjuntos de datos a pequeña escala pueden reproducirse completamente si algo sale mal, cosa que no sucede con los macrodatos; viii) compartición de datos: los datos a pequeña escala a menudo se comparten entre un pequeño grupo de personas, mientras que los datos grandes con un grupo más grande; ix) privacidad: los datos a pequeña escala pueden tener menos preocupaciones de privacidad, mientras que los datos grandes plantean preocupaciones de privacidad más significativas; y x) análisis: los datos a pequeña escala suelen poder analizarse utilizando herramientas y métodos tradicionales, mientras que los datos grandes requieren técnicas avanzadas (Berman, 2013; Poulson, 2019).

Se pueden distinguir tres tipos de aprendizaje automático: el aprendizaje supervisado, el aprendizaje no supervisado y el aprendizaje por refuerzo. El aprendizaje supervisado tiene por objeto “inducir modelos capaces de predecir el valor de ciertas variables dependientes a partir de variables independientes” (González, 2015, p. 77) mediante un conjunto de ejemplos (*training set*) proporcionados por un supervisor humano (Sutton & Barto, 2018, p. 2). El aprendizaje no supervisado, en cambio, no parte de variables dadas, encontrando patrones subyacentes en los datos sin necesidad de ser entrenado, como por ejemplo, las técnicas de análisis de conglomerados, que tratan de encontrar grupos de datos de iguales características (*clustering*). Por último, existe el denominado aprendizaje por refuerzo, consistente en aprender a decidir, ante una situación determinada, qué acción es la más adecuada para lograr un objetivo. Para ello la IA posee una política o plan (*policy*), una recompensa (*reward signal*) y una función de asignación de valor (*value function*) (Sutton & Barto, 2018, p. 7). En el aprendizaje por refuerzo la IA ha de explorar la manera de conseguir una mayor recompensa, explotando las formas descubiertas para determinar cuál es la más eficiente (Sutton & Barto, 2018, p. 3).

De forma paralela se ha avanzado en una familia de métodos conocidos como aprendizaje profundo (*Deep learning*). El aprendizaje profundo contiene técnicas de aprendizaje de representación entre las que destacan las redes neuronales profundas (RNP). Las RNP son redes neuronales artificiales (RNA) de múltiples capas que “aplican en cascada un conjunto de transformaciones no lineales, de forma que cada nivel recibe en entrada la salida del nivel anterior y lleva a cabo una abstracción, aprendiendo así representaciones cada vez más complejas” (Meseguer González & López de Mántaras, 2017, p. 82), y se hallan basadas en el algoritmo “backpropagation”, creado en 1986 por Rumelhart, Hinton y Williams (Lee & Lin, 2020). Un ejemplo usual de cómo funcionan las RNP es el reconocimiento de imágenes. La máquina parte de píxeles: la primera capa de representación detecta la presencia o ausencia de contornos; la segunda capa reconoce diseños; la tercera capa ensambla los diseños en combinaciones más grandes que corresponden a partes de objetos; y las capas posteriores detectan estos objetos como las combinaciones de estas partes (Meseguer González & López de Mántaras, 2017, p. 88).

1.3 APROXIMACIÓN LEGAL A LA INTELIGENCIA ARTIFICIAL

El impacto de la IA no distingue los límites fronterizos de los estados, impulsando al legislador, nacional y supranacional, a abordar los posibles efectos que pueda tener en los derechos de los ciudadanos. No obstante, el principal problema al que debe hacer frente

la regulación es la naturaleza volátil de la innovación tecnológica, que provoca la denominada “disrupción regulatoria” (Grimm, 2022) o la capacidad de eludir la sistematización legal gracias al avance tecnológico¹⁰.

Por eso, las definiciones contenidas en las normas han de ser realistas, a la vez que adaptativas, recurriendo a menudo a la remisión a otros textos que usualmente no poseen carácter vinculante (la *soft law* de la IA). Este apartado pretende dar a conocer el estado de la regulación en la materia en dos planos: el nacional, español, y el supranacional o regional europeo.

1.3.1 Regulación de la IA en España

La primera mención explícita a la IA en el Boletín Oficial del Estado (BOE) la encontramos en el verano de 2020, con motivo de la creación del Consejo Asesor de Inteligencia Artificial. La IA es, según esta primera aproximación,

la rama de la informática y la Ingeniería que se ocupa del estudio, diseño y construcción de sistemas físicos (hardware) y lógicos (software) que manifiestan un comportamiento que podría considerarse «inteligente», dada su capacidad para analizar su entorno y ejecutar acciones con cierto grado de autonomía con el fin de alcanzar unos objetivos específicos (Preámbulo de la Orden ETD/670/2020, de 8 de julio, por la que se crea y regula el Consejo Asesor de Inteligencia Artificial)

En noviembre de 2020, en las conclusiones de la Estrategia Nacional de Inteligencia Artificial (ENIA), parte de la agenda digital del Gobierno “España Digital 2025”, se define de nuevo la IA como

sistemas de software (y posiblemente también de hardware) diseñados por humanos que, ante un objetivo complejo, actúan en la dimensión física o digital: percibiendo su entorno, a través de la adquisición e interpretación de datos estructurados o no estructurados, razonando sobre el conocimiento, procesando la

¹⁰ Para adelantarse a esta posibilidad se han creado grupos de expertos a nivel español y europeo. En Europa, el *Libro Blanco sobre la inteligencia artificial* expone que, “debido a la rapidez con la que evoluciona la inteligencia artificial, el marco regulador debe dejar margen para abordar su desarrollo en el futuro. Toda modificación debe limitarse a aquellos problemas detectados con claridad para los que existan soluciones factibles”. En España la *Carta de Derechos Digitales* manifiesta que “la propia naturaleza rápidamente cambiante del entorno digital hace necesario asegurar la existencia de un proceso abierto de reflexión que permita mejorar la adecuación del marco jurídico a las nuevas realidades.” Y en la Unión Europea, el proyecto de Reglamento también se preocupa en su exposición de motivos de encontrar “una definición única de la IA que puede resistir el paso del tiempo” (CT 1.1) COM(2021) 206 final 2021/0106(COD).

información derivada de estos datos y decidiendo las mejores acciones para lograr el objetivo dado. Los sistemas de IA pueden usar reglas simbólicas o aprender un modelo numérico, y también pueden adaptar su comportamiento al analizar cómo el medio ambiente se ve afectado por sus acciones previas (ENIA, 2020)

Hay que reseñar el reciente interés nacional por la IA. Antes de 2010 el total de menciones en el BOE se circunscribía al personal investigador. Es en el periodo entre 2010 y 2020 cuando empiezan a surgir primero másteres, e incluso más recientemente, grados en IA¹¹. Y en un corto espacio de tiempo, la IA ha pasado de los laboratorios y las cátedras universitarias a disponer de una Secretaría de Estado, e incluso de una Fundación dedicada a su promoción en el campo del Derecho¹².

En julio de 2021 se aprueba una Carta de Derechos Digitales que carece de carácter normativo, al estilo de la Declaración realizada en el parlamento italiano *Dichiarazione dei diritti in Internet*. La Carta, más amplia que su homóloga, extiende su campo de acción a los conocidos como neuroderechos y los derechos ante la IA. Para la Carta, “la inteligencia artificial deberá asegurar un enfoque centrado en la persona y su inalienable dignidad, perseguirá el bien común y asegurará cumplir con el principio de no maleficencia” (XXV, 1). La IA deberá en su ciclo de vida: “i) garantizar el derecho a la no discriminación; ii) establecer condiciones de transparencia, auditabilidad, explicabilidad, trazabilidad, supervisión humana y gobernanza; y iii) garantizar su accesibilidad, usabilidad y fiabilidad” (XXV, 2).

La reciente *Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación* (LITND) es la primera norma vigente, con rango de ley, que trata la IA de forma explícita (arts. 3.1 y 23 LITND) aunque sin dar una definición de esta. La ley remite a la ENIA, la Carta de Derechos Digitales y las iniciativas europeas en relación a la definición, clasificando la IA en tres niveles: i) un primer nivel es la IA al servicio de la

¹¹ En cuanto a los másteres vid. resoluciones de 25 de octubre de 2010, Universidad de Málaga; de 20 de julio de 2011, Universidad de Sevilla; de 23 de noviembre de 2012, Universidad politécnica de Valencia; de 1 de marzo de 2013, UNED; de 16 de junio de 2014, Universidad Politécnica de Valencia; de 21 de junio de 2017, UIMP y de 6 de febrero de 2018 de la Universidad internacional de La Rioja. Y en cuanto a los grados, vid. resoluciones de 30 de mayo de 2022, de la Universidad Pontificia Comillas; de 20 de julio de 2022 de la Universidad Alfonso X el Sabio; de 26 de julio de 2022, de la Universidad Rey Juan Carlos de Madrid y de 8 de septiembre de 2022, de la Universidad Complutense de Madrid.

¹² El 16 de diciembre de 2021 se inscribe en el Registro de Fundaciones la *Fundación de la Inteligencia Artificial Legal*, cuyo objeto es “la realización de actividades, iniciativas y proyectos para fomentar el uso de las nuevas tecnologías, y en especial de la inteligencia artificial, en el ámbito legal” (*Resolución de 16 de diciembre de 2021, de la Dirección General de Seguridad Jurídica y Fe Pública, por la que se inscribe en el Registro de Fundaciones la Fundación de la Inteligencia Artificial Legal*)

Administración Pública involucrada en la toma de decisiones; ii) en el segundo nivel se encuentra la IA involucrada en la toma de decisiones fuera de la Administración Pública; y iii) un tercer nivel, para la IA que no interviene en la toma de decisiones (art. 23 LITND). Sólo en el primer nivel se han de tomar en cuenta “criterios de minimización de sesgos, transparencia y rendición de cuentas, siempre que sea factible técnicamente” (art. 23.1 LITND). Como apreciación, ese “siempre que sea factible técnicamente” supone una cláusula abierta, que habrá de completarse en relación con los riesgos que señale la futura normativa europea sobre la IA¹³.

Actualmente, se encuentra en proceso de elaboración también un Anteproyecto de Ley de Medidas de Eficiencia Digital del Servicio Público de Justicia (APLED), en el que más tarde nos detendremos por su interés en el ámbito jurídico, viniendo a completar la ya existente *Ley 18/2011, de 5 de julio, reguladora del uso de las tecnologías de la información y la comunicación en la Administración de Justicia* (LUTICAJ). La APLED busca la “justicia orientada al dato”, con el objetivo de conocer las tendencias de delitos de violencia de género en un período determinado o la situación de vulnerabilidad de las familias en los lanzamientos por impago de rentas, permitiendo diseñar políticas públicas que aborden estos problemas sociales (Llop Cuenca, 2021).

1.3.2 Regulación de la IA en Europa

1.3.1.1 La Unión Europea

El ordenamiento de la Unión Europea (UE) tiene como Derecho primario los tratados constitutivos: el *Tratado de Funcionamiento de la Unión Europea* (TFUE), el *Tratado de la Unión Europea* (TUE), el *Tratado constitutivo de la Comunidad Europea de la Energía Atómica* (EURATOM) y la *Carta de Derechos Fundamentales de la Unión Europea* (CDF). Tienen importancia en la materia aquellos apartados que versan sobre el mercado común (114 TFUE), el tratamiento de datos (39 TUE, 16.2 TFUE y 8 CDF) y la libre circulación de datos en el área comunitaria (16.2 TFUE). No obstante, las primeras menciones a la IA en el Diario Oficial (DO) de la UE se producen en 2017, a principios de mayo, con la revisión intermedia de la aplicación de la Estrategia para el Mercado

¹³ En todo caso, está aún vigente en España la polémica causada con el programa BOSCO, ante la negativa del Consejo de Transparencia a hacer público el código del programa que dejaba fuera de ayudas estatales a personas con derecho a percibir las mismas. (Belloso Martín, 2022).

Único Digital, adoptada en 2015 por la Comisión Europea (CE), y con la Declaración de Tallin sobre administración electrónica de octubre del mismo año.

Uno de los primeros órganos europeos en pronunciarse fue el Comité Económico y Social Europeo (CESE), órgano consultivo compuesto por representantes de empresarios, trabajadores y grupos de interés, que aprobó un Dictamen el 31 de mayo de 2017 tratando las consecuencias de la IA para el mercado único (digital). El Dictamen indicaba once áreas donde se plantearían desafíos: “ética; seguridad; privacidad; transparencia y rendición de cuentas; trabajo; educación y desarrollo de capacidades; (des)igualdad e inclusión; legislación y reglamentación; gobernanza y democracia; guerra; y superinteligencia” (párr. 1.1 2017/C 288/01). Tras este primer informe, el CESE se ha venido refiriendo en sucesivas ocasiones al impacto de la IA en áreas como el empleo o la administración pública.

También el Parlamento Europeo (PE), mediante iniciativas legislativas (INL) y no legislativas (INI), ha tratado los problemas relacionados con la IA¹⁴. Se pueden entrever preocupaciones muy diversas, que van desde la atribución de responsabilidades por los posibles daños ocasionados por su uso a que este derive en prácticas discriminatorias y de control indiscriminado, es decir, en una “mirada tecno-panóptica” (Schussler, 2020). En este sentido es interesante resaltar la llamada de atención del PE sobre el peligroso papel que la IA está jugando en el actual cambio de paradigma del Derecho Penal hacia un modelo actuarial¹⁵.

¹⁴ Aunque la iniciativa legislativa le correspondía originalmente a la Comisión Europea, con la evolución del Derecho de la UE el PE puede solicitar a la Comisión que se le presente una propuesta (arts. 140, 154, 155 TFUE). Las iniciativas a las que se refiere este apartado son las *Resoluciones del Parlamento Europeo, de 3 de mayo de 2022, sobre la inteligencia artificial en la era digital* (2020/2266(INI)); *de 6 de octubre de 2021, sobre la inteligencia artificial en el Derecho penal y su utilización por las autoridades policiales y judiciales en asuntos penales* (2020/2016(INI)); *de 20 de mayo de 2021, sobre la configuración del futuro digital de Europa: eliminación de los obstáculos al funcionamiento del mercado único digital y mejora del uso de la inteligencia artificial para los consumidores europeos* (2020/2216(INI)); *de 19 de mayo de 2021, sobre la inteligencia artificial en los sectores educativo, cultural y audiovisual* (2020/2017(INI)); (2020/2013(INI)); *de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas* (2020/2012(INL)); *de 20 de octubre de 2020, sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial* (2020/2015(INI)); *de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial* (2020/2014(INL)); *de 12 de febrero de 2019, sobre una política industrial global europea en materia de inteligencia artificial y robótica* (2018/2088(INI)); *de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial* (2020/2014(INL)).

¹⁵ Como afirma la *Resolución sobre la inteligencia artificial en el Derecho penal y su utilización por las autoridades policiales y judiciales en asuntos penales*: “el Derecho penal moderno se basa en la idea de que las autoridades estatales reaccionan ante un delito después de que se haya cometido, sin presuponer

Por su parte, la Comisión Europea se ha ido pronunciado en favor de incrementar las inversiones en IA, impulsado la Estrategia Europea de Datos de abril de 2018 y un Plan Coordinado con más de 70 acciones, de diciembre de ese mismo año, así como el *Libro Blanco sobre la inteligencia artificial*, que toma en cuenta las múltiples ventajas que tiene la IA, siempre y cuando sea “antropocéntrica, ética y sostenible y respete los derechos y valores fundamentales” (COM(2020) 65 final).

En esta dirección, ha realizado una propuesta de *Reglamento por el que se establecen normas armonizadas en materia de Inteligencia Artificial*, la futura “Ley de Inteligencia Artificial” (COM(2021) 206 final), prohibiendo ciertos tipos de sistemas basados en IA cuyo uso sea de una peligrosidad inaceptable para los derechos fundamentales o sea dañino física o psíquicamente. Este futuro Reglamento pretende adelantarse a las posibles regulaciones sobre la IA de los legisladores de los Estados miembros para evitar que puedan fragmentar el mercado. Esta propuesta de Reglamento define la IA como

el software que se desarrolla empleando una o varias de las técnicas y estrategias que figuran en el anexo I¹⁶ y que puede, para un conjunto determinado de objetivos definidos por seres humanos, generar información de salida como contenidos, predicciones, recomendaciones o decisiones que influyan en los entornos con los que interactúa.

En cuanto a la clasificación por peligrosidad, la propuesta de Reglamento distingue entre sistemas de IA con i) un riesgo inaceptable, generalmente prohibidos salvo el reconocimiento biométrico (que puede ser usado de forma excepcional bajo autorización judicial o administrativa) ii) un riesgo alto, que han de seguir unos requisitos iterativos de gestión y mitigación del riesgo, siendo examinados *ex ante* (Capítulo II, COM(2021) 206 final) y iii) un riesgo bajo o mínimo, sujetos a unas obligaciones muy limitadas, pero que se pueden adherir voluntariamente a códigos éticos de conducta.

que todas las personas son peligrosas y deben ser vigiladas constantemente para evitar posibles infracciones; que las técnicas de vigilancia basadas en la IA cuestionan profundamente este enfoque, por lo que es de urgente necesidad que los legisladores de todo el mundo evalúen exhaustivamente las consecuencias de permitir el despliegue de tecnologías que reducen el papel de los seres humanos en la garantía del cumplimiento de la ley y las decisiones judiciales” ((Q) 2020/2016(INI))

¹⁶ Según el Anexo I estas son: “a) Estrategias de aprendizaje automático, incluidos el aprendizaje supervisado, el no supervisado y el realizado por refuerzo, que emplean una amplia variedad de métodos, entre ellos el aprendizaje profundo. Estrategias basadas en la lógica y el conocimiento, especialmente la representación del conocimiento, la programación (lógica) inductiva, las bases de conocimiento, los motores de inferencia y deducción, los sistemas expertos y de razonamiento (simbólico)”.

Los sistemas de IA empleados en la Administración de Justicia, y que afectan a funciones relevantes, es decir, los que sean utilizados para “investigar e interpretar los hechos y el Derecho y aplicar la ley a unos hechos concretos” son considerados por esta propuesta de Reglamento como de alto riesgo, por lo que estarían, de aprobarse en su redacción actual, sometidos a estricto control.

1.3.1.2 El Consejo de Europa

Por último, la Comisión Europea para la Eficiencia de la Justicia (CEPEJ), creada por el Consejo de Europa el 18 de septiembre de 2002¹⁷, ha abordado el tema de la utilización de la IA en el Derecho en la *Carta Ética Europea sobre el uso de la Inteligencia Artificial en los sistemas judiciales y su entorno* (en adelante CEEIA o Carta ética), de 2018, estableciendo un marco de cinco principios que ha de seguir la IA en la Justicia: i) principio del respeto de los derechos fundamentales ii) principio de no discriminación; iii) principio de calidad y seguridad; iv) principio de transparencia, imparcialidad y equidad y v) principio “bajo el control del usuario”. Siguiendo estos principios, y el estado de la ciencia, la Carta distingue:

- Usos que deben ser fomentados: i) la mejora en la búsqueda de jurisprudencia más allá de palabras clave gracias al procesamiento de lenguaje natural o PLN¹⁸; ii) el acceso a la legislación y documentos legales gracias al PLN iii) la creación de herramientas estratégicas mediante métodos cualitativos y cuantitativos para mejorar el acceso a la justicia.
- Posibles usos, con precisiones metodológicas: i) ayudar a la elaboración de baremos sobre compensaciones civiles; ii) apoyo a la resolución alternativa de conflictos civiles; iii) la resolución de disputas en línea y iv) el uso de algoritmos en la investigación penal para identificar dónde se están cometiendo los delitos.
- Usos que pueden considerarse tras una mayor investigación: i) perfilado de jueces y ii) anticipación de las decisiones judiciales.

¹⁷ Vid. la *Resolución Res(2002)12 del Comité de Ministros del Consejo de Europa*.

¹⁸ Antes referido como un subcampo de la IA, el procesamiento del lenguaje natural consiste “en la utilización de un lenguaje natural [que es aquel usado para fines de la comunicación humana] para comunicarnos con la computadora, debiendo ésta entender las oraciones que le sean proporcionadas, o bien, desarrollar modelos que ayuden a comprender los mecanismos humanos relacionados con el lenguaje” (Vásquez, Quispe, & Huayna, 2009, p. 48).

- Usos para considerar con extrema reserva: i) utilización de algoritmos para la elaboración de perfiles en materia penal ii) aplicación de normas basadas en criterios cuantitativos (Quantity-based norms), es decir, transmutar la fuente del Derecho de la ley a los “precedentes en masa”.

La Carta ética viene a completar la futura Ley de Inteligencia Artificial de la UE, dando un espaldarazo a las funciones accesorias que poco tienen que ver con el enjuiciamiento o la aplicación del Derecho y sí con la mejora del acceso a la justicia, pero poniendo en duda aquellas que puedan suponer riesgos para los derechos fundamentales.

CAPÍTULO II: INTRODUCCIÓN A LA JUSTICIA PREDICTIVA

Una vez abordada la necesaria introducción sobre la IA y su marco normativo, queda sumergirnos en el encaje de la IA en el Derecho (IA & D). Este trabajo trata en concreto de la IA focalizada en su posible aplicación predictiva en la Administración de Justicia. Sin embargo, no se puede ocultar que esta separación es artificiosa pues las herramientas aquí tratadas, así como la (posible) automatización del proceso judicial afectan a múltiples operadores jurídicos, entre ellos los abogados, de forma directa o indirecta, en, por ejemplo, el pronóstico para la estrategia procesal (Pulido, 2022, p. 305; Susskind, 2017, p. 79).

La historia de la IA & D puede remontarse a los postulados de Lee Loevinger (1949), que contraponía la jurisprudencia, como especulación, con el análisis científico cuantitativo de la justicia o “Jurimetría” (p. 406). Esta relación ha evolucionado hasta la actualidad, donde están presentes técnicas de IA que, como señala Kevin D. Ashley (2017), intentan replicar el razonamiento jurídico humano, los denominados modelos computacionales de razonamiento legal (MCRL) y, dentro de ellos, los modelos que replican el proceso de argumentación legal, los modelos computacionales de argumentación legal (MCAL) (Ashley 2017 pág. 3). En un principio, estos objetivos se pretendían conseguir mediante los ya vistos sistemas expertos legales basados en la GOFAL, como TurboTax (Surden, 2017), donde el conocimiento se daba de “de arriba abajo” (*top-down*). Hoy en día la aproximación se produce de “abajo a arriba” (*bottom-up*) (Ashley, 2019, p. 95), usando algoritmos de aprendizaje automático para procesar los datos de los cada vez más ingentes y completos bancos de datos legales, con la posibilidad de predecir fallos judiciales, e incluso tratar de argumentar decisiones (mediante las técnicas de PLN llamadas *argument mining*).

Aunque aún persiste la incertidumbre ante el potencial que puede alcanzar, existe la discusión acerca de si, en un futuro cercano, las profesiones jurídicas podrían encontrarse en peligro de forma inminente y generalizada (Remus & Levy, 2017), representando una de las controversias mayores en este campo el problema del replazo del profesional, o del abaratamiento de los costes de determinados servicios legales, a causa de la competencia con la máquina (Agrawal et al., 2019, p. 175). Por ejemplo, en el campo de la medicina, Geoffrey Hinton afirmó en 2016 que se debería dejar de formar radiólogos, pues en cinco, quizás diez años, la IA sería mejor que ellos en el diagnóstico a través de

imágenes, afirmación que obtuvo tanta repercusión como rechazo entre las revistas especializadas en radiología por, entre otras cosas, reducir la disciplina a la mera visualización de imágenes (Ramón Alvarado 2018).

El mismo debate se está dando en el Derecho, planteando Ettore Battelli (2020) dos posibles ventajas de la IA en la toma de decisiones judiciales: “(1) proporcionar rendimientos más eficientes que aquellos producidos por el ser humano, y (2) garantizar seguridad jurídica” (Battelli, 2020, p. 46). Es posible que la utilización de la IA en el proceso judicial solviente problemas de calado que enfrentan los Tribunales, como la sobrecarga de trabajo por la acumulación de casos idénticos, los conocidos como “pleitos masa” (Perea González, 2021), o la facilidad de encontrar la norma aplicable, debido a la maraña normativa contemporánea, el denominado “desbordamiento de las fuentes de Derecho” (Pérez Luño, 2011). Sin embargo, algunos defensores de una nueva “justicia robótica” creen que se debería ir más allá, siguiendo la aspiración ilustrada de convertir al Juez en la “boca muda de la ley” y sustituir al juez humano por el denominado “juez-robot” o “juez-IA”.

2.1 DEFINICIÓN DE JUSTICIA PREDICTIVA: EL SUEÑO DEL HOMBRE MALO

En el Derecho estadounidense y escandinavo de principios del siglo XX cobró relieve una corriente de pensamiento jurídico de “revuelta contra el formalismo”, que constituiría el germen del realismo jurídico. Para definirlo de forma sucinta, el realismo se trata de una posición reduccionista del Derecho que condiciona la validez jurídica a la eficacia, es decir, “las normas válidas, para serlo deben ser eficaces” (Peces-Barba Martínez, Asís Roig, & Fernández García, 1999, p. 31). Uno de sus pioneros fue el norteamericano Oliver Wendell Holmes, que llegaría a ser *Associate Justice* de la Corte Suprema de los EEUU. Entre sus frases más famosas está su concepción del Derecho como “las profecías acerca de lo que los tribunales harán en concreto, nada más y nada menos” (Holmes, 1975, p. 21).

Holmes introdujo el prisma del “hombre malo”, tomando como punto de observación del Derecho la perspectiva de aquel individuo que, ajeno a los problemas lógicos, únicamente se interesa por el resultado del litigio, desligando así el Derecho de la moral, el “ser” del “deber ser” (Solar Cayón, 2012, p. 33). Aunque se decanta por la profecía, como operador jurídico Holmes parece conformarse con un Derecho insondable, mientras aboga por una

interpretación jurídica creativa¹⁹. Sin entrar a negar la tesis holmesiana, se ha de matizar que se ha demostrado que la actividad judicial no es homogénea, existiendo casos cuya solución requiere menos creatividad que otros, y donde los jueces pueden seguir una serie de directrices subjetivas que involucran memoria, estadística o emociones, los denominados heurísticos (Nieva Fenoll, 2018; Tversky & Kahneman, 1974). Por ello, si fuera posible realizar modelos sobre las decisiones de los tribunales mediante la IA, se presentarían como un recurso fundamental para que el “hombre malo” contemporáneo logre conocer el (probable) resultado de los Tribunales.

El interés por la previsión de las decisiones en el Derecho se ha revitalizado en la actualidad al amparo de la denominada “justicia predictiva”. Esta terminología no está exenta de crítica, pues para algunos autores “predecir” (*pre-* (antes) y *dicere* (decir)) proviene de las ciencias “duras”, siguiendo la mala traducción del inglés “*predictive justice*” (Pérez Ragone, 2021)²⁰. La palabra “predictiva” es entendida por estos como “predecible” o “probable”, debido a que, en el estadio actual, las herramientas de las que se sirve la justicia predictiva no dan en general argumentos sobre sus decisiones, sino meras probabilidades (Contini, 2019).

Teniendo en cuenta la discusión anterior, en la práctica la justicia predictiva versa sobre herramientas basadas en IA que toman como variable independiente los datos relativos al caso (*input*) y como variable dependiente la probable decisión judicial (*output*), auxiliándose de algoritmos para realizar la predicción²¹. La justicia predictiva hace uso

¹⁹ Holmes nunca llegó a aclarar cómo se podía llegar a la “profecía” del Derecho. De hecho, afirmó que la teoría era su aportación, no la práctica (Ruger, Kim, Martin, & Quinn, 2004, p. 1156). Como señala Francisco J. Laporta, existe una paradoja, pues el “hombre malo” nunca elegiría un juez experimentalista como Holmes (Laporta San Miguel, 2002).

²⁰ De esta distinción se hace eco la CEEIA: “Debe hacerse una distinción desde el principio entre lo que es una “predicción” y lo que es un “pronóstico”. La predicción es el acto de anunciar lo que sucederá (prae, antes de - dictar, decir) antes de futuros eventos (por sobrenatural inspiración, por clarividencia o premonición). El pronóstico, por otro lado, es el resultado de observar (puntería, viendo) un conjunto de datos para prever una situación futura” (par. 60). Sin embargo, para otros autores carece de sentido en la práctica diferenciar entre pronóstico y predicción (Corvalán, 2019a, p. 9).

²¹ Recogemos a continuación diversas definiciones de justicia predictiva como: “la posibilidad de prever el resultado de un juicio a través de algunos cálculos, en particular predecir la probable sentencia relativa a un caso específico, con el auxilio de algoritmos (Battelli, 2021, p. 60); “[la justicia predictiva] analiza masivamente decisiones judiciales anteriores para identificar tendencias en un caso específico” (Pérez Ragone, 2021p. 217); “*Per “giustizia predittiva” deve intendersi la possibilità di prevedere l’esito di un giudizio tramite alcuni calcoli; non si tratta di predire tramite formule magiche, ma di prevedere la probabile sentenza, relativa ad uno specifico caso, attraverso l’ausilio di algoritmi*” (Viola, 2018); “*Some private companies even aim to anticipate judges’ decisions with so-called “predictive justice” tools, powered by the open data of judicial decisions. i.e. making case law databases available for free download. The objective is to predict the outcome of a dispute on the basis of criteria previously provided by the user or predict the risk of infringement....*” (Barbaro & Meneceur, 2018, p. 2); o “*any mechanism and associated*

de una serie de algoritmos de aprendizaje automático (Máquinas de Vectores de Soporte o SVM de Vapnik, bosque aleatorio, etc.) con finalidad tanto “analítica/inductiva” (analizando las decisiones pasadas) como “prospectiva/predictiva” (prediciendo decisiones futuras) (Viola, 2020, p. 157).

Para predecir el resultado, los modelos alcanzan a conocer relaciones entre datos que van más allá de precedentes, leyes, o contratos, pudiendo incluir variables como el perfilado del abogado y de los despachos que son parte en un litigio, así como del juez encargado de resolver (Simón Castellano, 2021, p. 6). Para los partidarios de no limitar el uso de estas herramientas, aquellos datos no suponen más que un sobredimensionamiento del común “ruido de pasillo” (Pérez Ragone, 2021p. 218) que forma parte del conocimiento informal de los abogados sobre circunstancias extrajurídicas, resultando estas tan esclarecedoras para predecir el resultado como cualquier norma o variable legal. Sin embargo, entre sus consecuencias negativas se puede señalar una intromisión en la privacidad e independencia de los jueces²², dando como resultado prácticas de elección de la jurisdicción más favorable, el foro de conveniencia o *forum shopping*, teniendo como efecto la desnaturalización del derecho al juez ordinario predeterminado por la ley” (Hierro Hernández-Mora, 2020, p. 744).

Sin embargo, diversos tipos de sistemas de análisis predictivo basados en inteligencia artificial son empleados ya en muchos países como apoyo de la Administración de Justicia, del juez en el enjuiciamiento, o de la policía en la persecución del crimen (Miró Llinars 2019). En el terreno policial contribuyen a detectar denuncias falsas²³, así como al diseño de estrategias para anticipar el crimen basadas en la localización geográfica (*place-based*) o en la persona, víctima o victimario (*person-based*) (Richardson, Schultz, & Crawford, 2019). Y en el dominio judicial se usan para tareas tan dispares como la redacción de borradores de sentencias o la evaluación del riesgo de reincidencia criminal desde una perspectiva semejante a la actuarial, auxiliando al juez en la toma de decisiones.

algorithms that utilise predictive analytics or artificial intelligence or machine learning to predict the result of any given dispute, or any associated information” (Maramot, 2019, p. 39).

²² Así, en Francia, la *LOI n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice*, en su art. 33, prohíbe con penas de hasta cinco años de prisión y multas de hasta 300.000€ la utilización de los datos de los jueces para “evaluar, analizar, comparar o predecir sus prácticas profesionales reales o presuntas”, pudiendo responder también de este ilícito las personas jurídicas.

²³ España es ejemplo puntero en usar la IA con esta utilidad. Un ejemplo es el programa VeriPol, que, según un estudio realizado en la ciudad de Murcia por Liberatore, Quijano-Sánchez y Camacho-Collados (2019), ha aumentado la resolución de denuncias falsas con gran aceptación por parte de la policía, aunque su uso no está exento de críticas (Salgado, 2021).

Estas utilidades son, como veremos, causa de intenso debate, al basarse en prospecciones estadísticas cuya fiabilidad reside en la cantidad de datos disponibles y en su calidad. Nos enfrentamos aquí al problema de la llamada “*dirty data*”, expresado en el principio informático “basura entra, basura sale” (*garbage in, garbage out*), pudiendo la máquina reflejar resultados errados al haber partido de datos “contaminados” (Richardson et al., 2019). Como bien ha señalado gran parte de la doctrina, la predicción algorítmica no tiene nada que ver con la ciencia ficción. La IA no es sustituta de los “precognitivos” de la película *Minority Report*.

2.2 ALGUNAS EXPERIENCIAS DE JUSTICIA PREDICTIVA

La justicia predictiva ocupa actualmente una parte importante del mercado de herramientas tecnológicas aplicadas al Derecho, las conocidas como *Legal Tech*, en creciente expansión, que han pasado de 99 patentes en el mundo en 2012 a 1.369 en 2019 (Lock, 2020; Thomson Reuters, 2017). En España ya existen varias bases de datos, entre ellas Jurimetría, Tirant Analytics o Vlex Analytics, que incluyen herramientas de análisis predictivo entre sus utilidades, así como numerosas herramientas que pueden asistir al abogado en diversas tareas.

Se pueden mencionar diversas experiencias en el empleo de herramientas de justicia predictiva, tales como: 1) la predicción de sentencias del Tribunal Europeo de Derechos Humanos (TEDH) o la Corte Suprema de EE.UU.; 2) la predicción con usos comerciales mediante sistemas que asisten a profesionales jurídicos y 3) herramientas predictivas que asisten a la Administración de Justicia. A continuación se verán algunos ejemplos.

2.1.1 La predicción de sentencias de la Corte Suprema de EE.UU. y del TEDH

- **La Corte Suprema de los Estados Unidos**

Siempre ha habido interés, por parte tanto de juristas como de politólogos, por obtener predicciones fiables acerca de cómo resuelve la Corte Suprema de EE.UU., cristalizándose dicho interés en el *Supreme Court Forecasting Project*, que tenía por objeto predecir el resultado de las decisiones pendientes de resolución antes de octubre de 2002 en lo relativo a dos cuestiones: 1) ¿afirmará o revocará el tribunal la apelación?, y 2) ¿Qué votará cada juez individual? Con ese objetivo, el trabajo de Andrew D. Martin, Kevin M. Quinn, Theodore W. Ruger y Pauline T. Kim (2004) se servía, por un lado, de un panel de expertos analistas, y, por otro, de un modelo estadístico basado en los datos

de casos precedentes, para predecir el futuro fallo de la Corte. Este modelo discriminaba, primero, mediante dos árboles de decisión si se reunían o no las características de sentencias unánimes, conservadoras o liberales; si estos no pronosticaban una decisión unánime o daban resultados contradictorios, se pasaba a nueve árboles de decisión para prever el voto de cada juez, tomando en cuenta la interdependencia entre ellos²⁴. El modelo propuesto se basaba en seis variables, revelándose finalmente como la más determinante la ideología -conservadora, liberal o independiente- de los jueces²⁵. La fiabilidad de las predicciones mediante el modelo superó a las del grupo de expertos en cuanto al sentido del fallo del tribunal (75% a 59,1%), aunque no en relación al sentido de los votos particulares de los jueces (66,7% a 67,9%).

Pese a que el modelo anterior demostró una considerable eficacia atendiendo al tiempo en que fue creado, servía únicamente para realizar predicciones en relación a la composición de jueces que en aquel momento se sentaban en la Corte Suprema, siendo necesaria una nueva aproximación que pudiera tratar la generalidad de casos, ya con la llegada de las nuevas tecnologías. Este intento vino de la mano del trabajo de Daniel Martin Katz, Michael J. Bommarito y Josh Blackman (2017), quienes aplicaron técnicas de aprendizaje automático a partir de los datos publicados en la Base de Datos de la Corte Suprema, analizando cerca de 240 variables por sentencia que se podían conocer de forma previa al fallo. Los autores usaron como muestra el conjunto de sentencias que va de 1816 a 2015 (28.009 casos y 243.882 votos particulares), valiéndose de un algoritmo de bosque aleatorio (*random forest*) que obtuvo un grado de fiabilidad en sus predicciones del 70,2% en relación al sentido de las sentencias y un 71,9% en los votos individuales de los jueces.

Merece también mención un trabajo posterior, publicado con una diferencia de unos meses por parte de los mismos autores, sobre el denominado “*crowdsourcing*” o

²⁴ Las variables son parte de una aproximación aptitudinal, cuyos orígenes se remontan a 1948 con la obra de Charles Herman Pritchett: *The Roosevelt Court: a Study in Judicial Politics and Values 1937-1947* y con el precedente cercano en el trabajo del politólogo Harold J. Spaeth. Las variables son: (1) el circuito de origen; (2) el asunto (*issue area*) según un protocolo obra de Spaeth; (3) el demandante o *petitioner* (por ejemplo, si se trata del propio estado, un particular, una empresa...); (4) el demandado o *respondent*; (5) la ideología del fallo de la instancia, codificada siguiendo también el protocolo de Spaeth y (6) si existía una cuestión de constitucionalidad (Ruger et al., 2004, p. 762).

²⁵ El modelo acierta gracias a su éxito relativo a la hora de predecir los votos de los cinco jueces más conservadores de la Corte, gracias en parte a la referencia del Juez conservador O'Connor (Ruger et al., 2004, p. 761). Así, Suzanna Sherry (2004) plantea que los votos de los jueces conservadores son los más predecibles para el modelo porque votan siguiendo sus políticas, siendo los votos centristas los menos predecibles: “*Liberal justices vote the law, the conservative justices vote their politics, and the centrist justices do neither*” (pp. 769-775).

inteligencia colectiva (2017). Este estudio se centra en la predicción humana agregada, estudiada a partir de la liga virtual FantasySCOTUS, un juego de predicción con más de 7.000 jugadores que han realizado más de 600.000 predicciones sobre el resultado de cerca de 400 decisiones de la Corte Suprema, alcanzando una exactitud de 80.8%, la mayor hasta el momento (Katz, Bommarito II, & Blackman, 2017).

- **El Tribunal Europeo de Derechos Humanos**

En el plano europeo destaca el experimento, centrado en las sentencias del TEDH, llevado a cabo por Nikolaos Aletras, Dimitrios Tsarapatsanis, Daniel Preotiuc-Pietro y Vasileios Lampos (2016). Estos autores elaboraron un modelo predictivo que usaba como *data set* la jurisprudencia previa del Tribunal en 584 casos sobre tres derechos recogidos en el *Convenio Europeo para la Protección de los Derechos Humanos* (CEDH): la prohibición de la tortura, del art. 3 (250 casos); el derecho a un proceso equitativo, del art. 6 (80 casos) y el derecho al respeto a la vida privada y familiar, del art. 8 (254 casos).

La propuesta del trabajo era analizar sentencias pasadas para ver si resultaba posible predecir si había existido una vulneración de uno de estos derechos (una única respuesta binaria) a partir de la evidencia textual dada en la sentencia (mediante PLN), descomponiendo esta en sus partes constitutivas: el procedimiento (*procedure*), las circunstancias (*circumstances*), los hechos (*facts*), el marco jurídico relevante (*relevant law*), la legislación (*law*) y el conjunto de todas (*full*). Los autores utilizaron para ello un algoritmo de Máquinas de Vectores de Soporte, que analizaba las relaciones entre unidades lingüísticas (*N-grams*) y conglomerados (*clusters*) en sentencias clasificadas, asignando valores positivos o negativos a las características que señalaban si el Derecho había sido conculcado o no (+1 si existe vulneración, -1 no existe). La hipótesis del trabajo parte de la tesis originaria de Reed C. Lawlor, consistente en que, mediante el análisis de la ley y los hechos, una máquina puede conocer la decisión judicial, logrando finalmente el algoritmo predecir el resultado en un 79% de los casos.

El estudio concluye que los jueces están fuertemente influenciados por factores extrajurídicos, en concreto, por la información contenida en los apartados relativos a las circunstancias (*circumstances*) y los hechos (*facts*), lo que, en opinión de los autores, vendría a validar las premisas del realismo jurídico americano. Su obra ha tenido una gran repercusión mediática y académica, pero estudios posteriores han desafiado sus resultados, criticando la metodología empleada por contener sesgos. El modelo empleado

por Aletras et al. toma en cuenta para predecir el fallo con el apartado de legislación (*law*), una sección de motivación de la decisión que, en ocasiones, puede contener el mismo fallo (Medvedeva, Vols, & Wieling, 2020, p. 262). Además, las circunstancias (*circumstances*) son contemporáneas a la sentencia, lo que para sus críticos sería “como “predecir” si un juez desayunó cereales ayer basándose en un informe nutricional de los alimentos que tenía en el plato al tiempo exacto que los masticaba” (Pasquale & Cashwell, 2018, pp. 67-68)

2.1.2 Usos comerciales de la justicia predictiva:

- **Lex Machina:**

Lex Machina es una herramienta pionera, diseñada en la Universidad de Stanford por el profesor de Derecho Mark Lemley, para resolver, en un principio, litigios de propiedad intelectual (PI). Propiedad de Lexis Nexis desde 2015, Lex Machina usa un algoritmo de aprendizaje automático supervisado centrado en los diversos actores que participan en el proceso judicial. El algoritmo toma en cuenta variables como la firma de abogados que representa a cada parte, el perfil del Juez y el circuito donde se va a realizar el juicio, con el fin de predecir el probable resultado (Ashley, 2019).

Lex machina consiguió predecir el resultado de un 64% de los litigios sobre propiedad intelectual de EE.UU. durante una década, gracias a un método de regresión logística, encontrando como variable más significativa el perfil del Juez (Surdeanu, Nallapati, Gregory, Walker, & Manning, 2011). El principal escollo al que se ha de enfrentar Lex Machina es su propia naturaleza de predicción apriorística (*prior factors*), no tomando en cuenta los derechos subjetivos que se discuten en el caso (*merits*), cuestión en la que los desarrolladores dicen estar trabajando (Ashley, 2017). De hecho, es posible cuestionar si las variables que usa pueden servir a otros supuestos ajenos a la PI, dado que el éxito de la herramienta se centra en un área del Derecho en el que un acuerdo previo puede ser altamente beneficioso si las pretensiones corren riesgo de ser desestimadas, antes incluso de conocer en profundidad los pormenores del caso (Sil, Roy, Bhushan, & Mazumdar, 2019).

- **Ravel Law**

Ravel Law fue creada por estudiantes de la Facultad de Derecho de Stanford con el apoyo de la biblioteca de Derecho de Harvard. Este proyecto se basó en el escaneo de 40

millones de páginas del contenido de la biblioteca, parte de la iniciativa de Harvard “Caselaw Access” (HLS, 2015). La función de la aplicación consistía en el mapeado de los precedentes legales para relacionarlos con un concepto jurídico concreto. La predicción que realiza Ravel Law se situaría en el campo de las micro-predicciones, como las reglas y el lenguaje específico que un juez prefiere y suele citar, o las distinciones que diferencian a un juez de otro, dando al abogado herramientas para conocer la posibilidad de éxito de un argumento o una tendencia que sigue un tribunal, pero no pretende predecir el sentido de la sentencia (Lewis, 2017).

- **Jurimetría**

Jurimetría es la herramienta de analítica predictiva de Wolters Kluwer, gracias a un proyecto desarrollado en metodología Lean-Agile. En concreto, la aplicación informa sobre: 1) la duración de media de un proceso en el juzgado que le corresponda o en partido; 2) la duración media de un proceso según la materia; 3) las resoluciones judiciales y sentencias que estiman un asunto 4) las resoluciones judiciales y sentencias que desestiman un asunto 5) y la cantidad de asuntos que se han resuelto y tramitado en un año (Molina García, 2017).

La aplicación dispone de Jurimetría del Caso, Jurimetría del Tribunal, Jurimetría del Magistrado y Jurimetría del Abogado, ofreciendo datos sobre “duración estimada del proceso, argumentación legal del juez o la contraparte, línea jurisprudencial o posibilidad de que un asunto sea o no recurrido” (Oliveros Roselló, 2019). Con esta herramienta, que tiene como fuente la base de datos del Consejo General del Poder Judicial, se pretende, según sus creadores, encontrar tendencias ocultas en la jurisprudencia para que el abogado pueda conocer el riesgo de la acción que promueve.

2.1.3 Usos de justicia predictiva en la Administración de Justicia

- **El Fiscal Watson**

Realmente no es una herramienta, sino un conjunto de herramientas con finalidades policiales y de investigación, fabricado con la colaboración de una facultad canadiense y la empresa IBM, que sirve a la Fiscalía colombiana. Basado en Watson Debater, de IBM, ante un tema, analiza textos y selecciona “los argumentos que parecen más sólidos - probablemente, entre otras razones, por ser los más repetidos-y los expresa en un lenguaje natural” (Nieva Fenoll, 2018, p. 30). El Fiscal Watson permite realizar correlaciones para

luchar contra la delincuencia, en concreto usa la base de datos de denuncias del Sistema Penal Oral Acusatorio colombiano que, de 2005 a 2018, acumulaba trece millones de denuncias (Flórez Rojas & Vargas Leal, 2020). Es usado, según la Vicefiscal General de la Nación en su momento, María Paulina Riveros, para analizar esas denuncias, correlacionarlas, hacer análisis de contexto sobre elementos similares, como *modus operandi*, características físicas, tipos de armas y vehículos, entre otros (El Espectador, 2018).

- **Prometea**

El software Prometea ha sido desarrollado por el Laboratorio de Innovación e Inteligencia Artificial (IALAB) de la Facultad de Derecho de la Universidad de Buenos Aires y financiado con fondos públicos del Ministerio Público Fiscal de Argentina. Se halla implementado desde finales de 2017 en la Fiscalía General Adjunta en la sala de lo Contencioso Administrativo y Tributario del Tribunal Superior de Justicia de la Ciudad de Buenos Aires (Estevez, Linares, & Fillotrani, 2020), que es un tribunal de tercera instancia que conoce de una serie determinada de asuntos -amparo habitacional, empleo público, derecho a la salud, derechos de personas con discapacidad, derechos de niños y adolescentes- “sobre los cuales ya existe una tendencia jurisprudencial muy firme” (Corvalán, 2019b). Prometea tiene dos utilidades principales para la Fiscalía: por un lado, apoyar su tarea mediante la búsqueda de información relevante, y, por otro, realizar borradores de dictámenes y propuestas de sentencia mediante un motor predictivo. Si el Fiscal se encuentra conforme con el borrador generado por Prometea, simplemente puede firmarlo y remitirlo al Tribunal, que es quien finalmente decide el caso (Saavedra & Upegui, 2021). La pertinencia de estos borradores es evaluada mediante un “test de concordancia”, relativo al grado de adopción por parte del Tribunal de los borradores generados automáticamente por Prometea. En 2018, cuando se realizó por primera vez esta evaluación, en 452 causas (92,2%) hubo concordancia entre la decisión del Tribunal y el borrador propuesto, en 33 (6,7%) no hubo concordancia y en 5 (1%) hubo una concordancia parcial (Estevez et al., 2020).

Para el desarrollo de Prometea, un grupo heterogéneo de programadores y fiscales entrenaron el algoritmo con 2.400 sentencias y 1.400 opiniones legales. También se realizaron árboles de decisión, tomando como referencia los casos de amparo habitacional, que eran la mayoría del trabajo monótono del Ministerio Público (en 2016 constituyeron el 66% de las tareas de la Fiscalía) (Estevez et al., 2020). En cuanto al *input*,

Prometea emplea palabras clave, permitiéndose trabajar mediante una interfaz con asistente de voz, al estilo de Siri o Alexa, para dirigir a la máquina a una serie de sentencias relevantes almacenadas en su base de datos. Esta herramienta es, siguiendo las palabras de su creador, Juan Gustavo Corvalán, una “caja blanca” o “*white box*” (Corvalán, 2019b), debido a que se puede llegar a conocer cómo alcanza las decisiones que toma mediante reglas lógicas, trazables y prefijadas, no como los algoritmos de “caja negra” o “*black box*”, cuya problemática veremos más adelante.

Entre los logros de Prometea está el aumento de la eficiencia de la Fiscalía. En su diseño se plantearon qué tareas habrían de ser automatizadas, reorganizando las labores de la fiscalía y llegando a conseguir un éxito de un 67% en aquellas áreas que consideraban automatizables (Estevez et al., 2020). En cuanto a la eficiencia ganada con su uso, antes de su implementación, para generar 1000 dictámenes relativos al derecho a la vivienda la Fiscalía necesitaba de 174 días de trabajo, mientras que con Prometea solo se precisan de 45 días (Corvalán, 2019b). El tiempo ganado por la herramienta redundaba en beneficio de la dedicación a los casos más difíciles, aumentando la calidad de las resoluciones gracias a la mejor argumentación en estos casos (Estevez et al., 2020).

- **El “Juez Sabio”**

Por último, China se halla hoy en la vanguardia del uso de la IA y las nuevas tecnologías en la Administración de Justicia, gracias al vertiginoso proceso de digitalización llevado a cabo en los últimos años. En 2015 ya existían tres grandes plataformas digitales de información judicial: *China Judicial Process Information Online*, que proporciona información a las partes y los abogados sobre el estado del litigio; *China Judgments Online*, que publica todas las sentencias de los tribunales, siempre que permitan publicidad; y *China Judgments Enforcement Information Online*, que publica la identidad de personas que han incumplido sentencias firmes, restringiendo su posibilidad de realizar algunas actividades como viajar en avión o comprar un inmueble (Shi, Sourdin, & Li, 2021).

A partir de la información contenida en *China Judgments Online*²⁶, el Tribunal Supremo de Pekín ha desarrollado un sistema de IA basado en *deep learning*, denominado “Juez

²⁶ Cuando se lanzó “*China Judgments Online*” (中国裁判文书网) en 2014, el número de documentos subidos era de 5,58 millones. En 2015, alcanzó casi 9 millones y en 2016, el sitio contenía cerca de 10 millones de documentos. En agosto de 2019, el volumen de datos había superado los 74,4 millones de documentos y hoy se estima que supera los 100 millones (Shi, Sourdin, & Li, 2021; Solar Cayón, 2022).

Sabio” (*Rui Fa Guan*), que puede ser usado también por cualquier tribunal de la provincia para la toma de decisiones judiciales (Solar Cayón, 2022). El objetivo perseguido por el “Juez Sabio” es que “casos con hechos similares reciban sentencias similares” (Shi, Sourdin, & Li, 2021, p. 9). Para lograrlo, el sistema es capaz, mediante la identificación de las cuestiones legales planteadas en el caso presente, de buscar y seleccionar sentencias sobre casos similares a nivel nacional, de analizar esos precedentes judiciales y de elaborar a partir de los argumentos contenidos en ellos una propuesta de decisión, que no es vinculante para el juez (Solar Cayón, 2022). Varias herramientas semejantes al “Juez Sabio” han sido implantadas también en otras provincias del país, como Hainan, Shanghai y Guangzhou (Shi, Sourdin, & Li, 2021).

CAPÍTULO III: PREDICCIÓN, INTELIGENCIA ARTIFICIAL Y PROCESO

El proceso puede ser entendido como “instrumento exclusivo a través del cual los Jueces y Tribunales ejercen la potestad jurisdiccional, y constituye una secuencia de actos, todos jurídicamente regulados, que se suceden en el tiempo con la finalidad de aplicar el Derecho a un caso concreto” (Doig Díaz, 2019, p. 5). Así, se ha de distinguir proceso de procedimiento, siendo este último el que configura el *iter* procesal, tal como expone Piero Calamandrei: “aun empleándose en el lenguaje común como sinónimos, tienen significados técnicos diversos, en cuanto el procedimiento nos indica más propiamente el aspecto exterior del fenómeno procesal, siendo posible que en el curso del mismo proceso pueda, en diversas fases, cambiar el procedimiento” (Calamandrei, 1943, p. 318).

De esta forma, el proceso se representa como un tipo ideal que concatena reglas y órdenes para lograr un fin, lo que, así enunciado, resulta semejante a un algoritmo (Juan Gustavo Corvalán, 2019). Además, la jerga jurídica entendida como un lenguaje simbólico, al estilo del “*Law-French*” o el latín jurídico (Radin, 1931, pp. 171-172), ha impulsado la idea de poder interpretar las leyes según modelos matemáticos (Viola, 2020). Sin embargo, las normas, que reflejan la voluntad del legislador, usan en su mayoría lenguaje natural, con la ambigüedad que esto conlleva. Esta dificultad puede evidenciarse con el ejemplo clásico de Herbert L. A. Hart: “se prohíbe la circulación de vehículos en el parque”, pero ¿qué es un vehículo? (Hart, 1958, p. 607)

La indeterminación a la hora de concretar el significado de un precepto legal constituye el fundamento para que Hart defienda la existencia de “casos difíciles” y, consecuentemente, la existencia de un margen de discrecionalidad del juez para resolverlos (Carrió, 1981, pp. 42-44). Hart, que concibe el Derecho como un conjunto de normas, aborda esta cuestión, conocida como la “textura abierta” del Derecho, mediante la distinción entre un núcleo duro de significado, inalterable para cualquier interlocutor (en el caso del vehículo, se encontraría dentro de este núcleo, por ejemplo, un automóvil), y una zona de penumbra (una bicicleta). Para tratar los problemas interpretativos surgidos de esta última, Hart plantea la necesidad de recurrir a la analogía (Rodríguez, 1997, p. 33).

Esta postura tiene su reverso en el racionalismo de Ronald Dworkin, cuya visión del Derecho no solo comprende reglas, sino también principios (Pérez Jaraba, 2010, p. 13).

Dworkin se opone a la discrecionalidad desde una perspectiva denominada “interpretación constructiva”, sosteniendo que, allí donde las reglas no proporcionan una solución al caso, el juez ha de hallarla tomando en cuenta, mediante un complejo proceso de ponderación argumentativa, principios jurídicos que frecuentemente se hallan en tensión. La eliminación de la discrecionalidad requiere, por tanto, de las capacidades fabulosas de un “juez Hércules” que, a través de dicho ejercicio de ponderación, siempre dé con la única respuesta jurídicamente correcta, representando así el prototipo ideal del “juez-robot” (Solar Cayón, 2021, p. 15).

Sin detenernos en este trascendental debate en el plano de la Filosofía del Derecho, es decir, en el “diseño de las plumas del pájaro”, los investigadores en la IA han seguido trabajando en modelos híbridos que se aproximen a las soluciones dadas por el razonamiento legal. Así, “no es que el razonamiento pragmático y la deducción clásica sean rechazados, sino que ambos deben conjugarse” (Almonacid Sierra & Coronel Ávila, 2020, p. 129). En definitiva, el proceso no es siempre un algoritmo, sino más bien una receta, en la que los humanos toleramos la información con un margen de imprecisión del que no es capaz la máquina (Desai & Kroll, 2017, p. 23), puesto que “los algoritmos, (al menos por ahora) no son capaces de decidir cuál es el criterio interpretativo más razonable ante un caso concreto” (Rodríguez Puerto, 2021).

Si bien puede resultar comprensible que mucha de la literatura jurídica actual sobre el empleo de la IA en la Administración de Justicia centre su atención en la futurista posibilidad de desarrollar un “juez-robot”, no por ello es menos criticable que, en la mayoría de los casos, estas aproximaciones obvien el inmediato impacto que la IA puede tener precisamente en aquellas áreas donde las tareas procesales son más mecánicas. Por ello, se hace necesario adoptar en esta materia “una visión global del sistema judicial y de las distintas funciones y tareas que implica el desarrollo del proceso judicial”, que tenga en cuenta el rol que la IA y otras tecnologías pueden desempeñar en la realización de cada una de ellas (Solar Cayón, 2021, p. 6). Esto es, la adopción de lo que se ha venido en llamar una “visión holística” (Corvalán, 2019a; Solar Cayón, 2022). Hay que valorar las funcionalidades de estas herramientas tanto en el *front office* como en el *back office* de la Administración de Justicia, facilitando el acceso a los diversos protagonistas del proceso -como abogados, procuradores y partes- a los documentos que integran el expediente digital, y posibilitando un trabajo cada vez más colaborativo y simultáneo entre todos ellos y los miembros de la oficina judicial.

3.1 UNA APROXIMACIÓN REALISTA: LOS LÍMITES DE LA IA

Aunque distan de ser perfectas, las herramientas de IA pueden ser de ayuda en muchas tareas, siempre y cuando exista interoperatividad, es decir, que la IA pueda acceder a las bases de datos relevantes para la decisión, disponiendo de una datosfera con información suficiente y de calidad. Por tanto, hay que rechazar la dicotomía entre, por un lado, la sustitución total del humano por la máquina y por otro, la ausencia total de la IA en el proceso judicial, abogando más bien por la colaboración entre ambos. Esta ya se da en otros campos con sus propias particularidades, como la medicina, en la que los patólogos asistidos por la IA logran mayor eficacia que sin ella cuando trabajan en el diagnóstico de enfermedades como el cáncer (Johnson et al., 2021). Parece innegable que la IA se está abriendo paso y tendrá un impacto en la Justicia, al menos, en funciones de información, apoyo y consejo (Shi, Sourdin, & Li, 2021).

En relación con las funciones que puede desempeñar la IA en el proceso, siguiendo la obra de Harbers, Peeters y Neerinx, la relación entre el humano y la IA puede tomar diversas formas:

- (1) *Man in the loop*, cuando la IA necesita aportes humanos a intervalos de tiempo regulares para poder llevar a cabo sus acciones.
- (2) *Man on the loop*, si la máquina es capaz de actuar por sí misma a partir de una programación previa, pero el humano puede intervenir interrumpiendo o modificando las acciones del robot en cualquier momento.
- (3) *Man out of the loop*, un modelo en el que la máquina actúa de manera independiente durante ciertos períodos de tiempo y, en estos intervalos, el ser humano no tiene influencia sobre las acciones del robot. (Harbers et al. 2017 en Miró Llinares, 2018, p. 93)

Aunque los sistemas *Man in the loop* dependen del humano, ya generan problemas de diversa índole que, como veremos, lastran su aplicación en algunas áreas del Derecho, como puede ser el Derecho penal. Esto no es baladí pues, de implementarse, la IA ha de superar las peculiaridades de cada orden jurisdiccional en relación con las garantías exigibles en el proceso²⁷.

²⁷ Hay que destacar que actualmente en el ámbito administrativo laboral se está usando la IA por parte de Inspección de Trabajo para investigar y, en su caso, poder abrir un expediente sancionador contra empresas

Habiendo expuesto esta cautela, la actividad procesal, común en todos los órdenes jurisdiccionales, se puede dividir, a grandes rasgos en dos partes: la tramitación y el enjuiciamiento (Nieva Fenoll, 2018, p. 31). Antes de ver los distintos usos de la IA en estas, hay que enfrentarse con los límites de qué puede hacer una máquina hoy. Este trabajo, basándose en el estadio actual de la cuestión, revisa los límites técnicos, es decir, las posibilidades que ofrece el análisis de información de una IA “débil”; jurídicos, entendidos como los márgenes que tiene la IA para implementarse de conformidad con el ordenamiento jurídico; técnico-jurídicos, tratando el problema de la transparencia en su vertiente legal y técnica; y, por último, de legitimación, analizando las posibilidades de aceptación de la posible inclusión de sistemas de IA en la Administración de Justicia.

3.1.1 Límites técnicos

Los economistas Ajay Agrawal, Joshua Gans y Avi Goldfarb (2018) exponen los límites de las capacidades predictivas de los sistemas de aprendizaje automático a partir de la distinción entre cuatro tipos de conocimientos: conocimientos conocidos, desconocimientos conocidos, desconocimientos desconocidos y conocimientos desconocidos.

En primer lugar, la predicción es altamente precisa si se dispone de datos suficientes sobre aquello que se quiere modelar y nosotros conocemos que la predicción resultante es buena. Estos son los conocimientos conocidos. Juan Gustavo Corvalán (2019b) habla, por ejemplo, del “control de suficiencia” que realiza el sistema Prometea, observando si los documentos jurídicos cumplen con requisitos de forma mediante un árbol de decisión. En ese medio controlado la IA opera con conocimientos conocidos²⁸.

En segundo lugar, se encuentran los desconocimientos conocidos, pues en ocasiones existen situaciones donde no se pueden obtener datos suficientes para alimentar un modelo, como en la predicción de sucesos extraordinarios pero conocidos, tales como los terremotos (Agrawal et al., 2019, p. 71). Ante la existencia de pocos datos, el humano es capaz de servirse de analogías, usando datos semejantes para orientar sus decisiones, extremo que la máquina aún no puede imitar. Frente a los desconocimientos conocidos,

que incumplan la normativa de horas extras. Esto se consigue mediante el análisis de una serie de variables, como el tamaño de la plantilla o la cifra de negocio o el volumen de este (Rodríguez, 2022).

²⁸ En su automatización, Prometea analiza si, junto con el expediente, se adjuntan copias de la demanda, del recurso de inconstitucionalidad, la queja, etc: “si el expediente está incompleto, Prometea asigna un modelo denominado “no autosuficiencia” en el cual se establece que el expediente no cuenta con los documentos indispensables para ser resuelto” (Corvalán, 2019b, p. 62)

previstos ya a la hora diseñar la IA, la máquina puede alertar de su limitación y solicitar el auxilio humano. Las predicciones en relación a personas cuando no existen macrodatos se basa en la construcción de perfiles, el denominado perfilado o *profiling*. Este se basa en asociar datos comunes y definir grupos a los que esa asociación es aplicable, sometiéndoles a un escrutinio añadido (Dolz Lago, 2022). Cuando esa asociación resulta ser espuria pasa a ser, como veremos, un conocimiento desconocido.

En tercer lugar, están los llamados “cisnes negros” (Taleb, 2011), que serían desconocimientos desconocidos, es decir, hechos no previstos que únicamente pueden ser explicados de forma retrospectiva. En el campo de la Justicia, una reforma legislativa supondría un desconocimiento desconocido para la máquina, dejando las predicciones inservibles, ya que estas se servirían de datos pretéritos a la norma en vigor. Esta “petrificación” de la Justicia es uno de los problemas que se ha de afrontar si se decide adoptar el *stare decisis et quia non movere* algorítmico, donde la ley es usada como un dato más entre muchos, fenómeno denominado la “*actualización*” del Derecho (Caterini, 2022, p. 7).

Por último, una de las mayores debilidades de la IA son los conocimientos desconocidos, que ocurren cuando la máquina da respuestas altamente precisas pero erradas (Agrawal et al., 2019, p. 73). Es el problema del sesgo, que no desaparece con el tamaño de la muestra. El sesgo no tiene una definición jurídica conocida (Belloso Martín, 2022) pero su problemática base es el no discernimiento de la máquina, que tiene en cuenta datos desconociendo estructuras subyacentes, es decir, las decisiones que dan lugar a estos²⁹.

El sesgo se produce porque la IA no es capaz de realizar un análisis contrafáctico³⁰. Esta forma de razonar propiamente humana tiene, siguiendo la literatura, dos funciones: una principal, la función preparativa, íntimamente ligada con la causalidad, en que analizamos los errores del pasado para facilitarnos el éxito futuro; y una segunda, subsidiaria, la

²⁹ El sesgo es una realidad compleja difícilmente abarcable en este trabajo. Normalmente se acusa al propio algoritmo proyectando en él rasgos humanos de ser racista, sexista, etc. En verdad se trata de un error. Los sesgos pueden surgir en la fase de entrenamiento, donde se excluyen datos relativos a minorías; en la fase de uso, con correlaciones espurias; o en la fase de validación y presentación, cuando el programador presenta el algoritmo a los promotores del producto. Todo esto hace necesaria una auditoría o un control de impacto (Belloso Martín, 2022, pp. 53-55)

³⁰ Este podría ser definido cómo: “¿Qué hubiera ocurrido si X no hubiera sucedido?”. Nótese que en el Derecho es común esta lógica. Sobre todo se puede ver reflejada en la causalidad penal, según la famosa fórmula de la “*conditio sine qua non*” (Salvador Coderch & Fernández Crende, 2006).

emocional, en que las personas piensan en realidades alternativas para sentirse mejor, o hacer que los demás se sientan mejor (Segura, Fernández Berrocal, & Byrne, 1998).

Un caso paradigmático que sirve para explicar los conocimientos desconocidos es el uso de la prueba probabilística en *Estados Unidos v. Shonubi* (895 F. Supp. 460 (E.D.N.Y. 1995)). En este caso se condenó en primera instancia al acusado, Charles O. Shonubi, un viajero procedente de Nigeria, por tráfico de drogas. En la sentencia, apelada con éxito, el tribunal atribuyó a Shonubi una cantidad superior a la que realmente le incautaron en el aeropuerto, basándose en sus viajes previos. Para ello, el órgano sentenciador se sirvió de un programa informático que calculaba la media de heroína requisada a otros traficantes, una vez hallada, y lo aplicó a todos los trayectos realizados por Shonubi, concluyendo que, con un 99% de probabilidad, había transportado 2090,2 gramos de estupefacientes³¹.

Interesa que uno de los argumentos dados por Shonubi consistía en que los detenidos normalmente lo eran por traficar con grandes portes, infraestimando el algoritmo las cuantías correspondientes a pequeñas cantidades, pues eludían los controles aduaneros más fácilmente. Si se diera por válido el argumento de la defensa, la pena impuesta sería claramente desproporcionada. Únicamente podríamos encontrar la respuesta usando el pensamiento “contrafactual”, debido a que no disponemos de un “grupo de control” para medir cuántos traficantes con pequeñas cantidades logran, efectivamente, traspasar la frontera.

3.1.2 Límites jurídicos

Para poder incluir la IA en el proceso judicial es preciso saber si su utilización sería ajustada al ordenamiento de cada Estado. Tanto la normativa de protección de datos -el *Reglamento (UE) 2016/679 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos* (Reglamento general de protección de datos o RGPD), la *Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales*

³¹ El tribunal consideró varios análisis estadísticos basados en 117 traficantes que usaban el mismo método de transporte de droga, provenientes de Nigeria y arrestados en el aeropuerto JFK durante el período de los ocho viajes de Shonubi. Así, se concluye que “existe un 99 % de posibilidades de que Shonubi llevara al menos 2090,2 gramos de heroína en los siete viajes combinados; hay un 95% de probabilidad de que haya importado más de 2341,4 gramos; una probabilidad del 75% de que llevara más de 2712,6 gramos; y un 55% de probabilidad de que llevara más de 3039,3 gramos (*Additional Material Available to Sentencing Judge On Remand A. Experts' reports I. Government expert United States v. Shonubi [Shonubi III] United States District Court for the Eastern District of New York, 1995 895 F.Supp 460*)

(LOPDGDD) y la *Ley Orgánica 7/2021, de 26 de mayo, de protección de datos personales tratados para fines de prevención, detección, investigación y enjuiciamiento de infracciones penales y de ejecución de sanciones penales* (LOPDPP)- como la *Constitución Española* (concretamente arts. 24 y 117 CE) y el futuro *Reglamento europeo sobre la IA* (COM(2021) 206 final 2021/0106(COD)) parecen impedir la toma de decisiones automatizadas que afecten a la función jurisdiccional.

En cuanto a la protección de datos, el RGPD dispone que “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” (art. 22.1 RGPD). En lo que abunda la legislación española, debiendo el afectado “ser informado de su derecho a oponerse a la adopción de decisiones individuales automatizadas” (11.2 LOPDGDD). La decisión únicamente puede autorizarse si concurren ciertas circunstancias. Estas pueden ser ordinarias, como el consentimiento explícito del afectado, habiendo de disponer siempre de “información significativa sobre la lógica aplicada” y “la importancia y las consecuencias previstas” (art. 13.2.f del RGPD y 11.1 LOPDGDD), o extraordinarias, como razones de orden público.

Con fines de prevención, detección, investigación y enjuiciamiento de infracciones penales o de ejecución de sanciones penales también quedan “prohibidas las decisiones basadas únicamente en un tratamiento automatizado, incluida la elaboración de perfiles, que produzcan efectos jurídicos negativos para el interesado o que le afecten significativamente, salvo que se autorice expresamente por una norma con rango de ley o por el Derecho de la Unión Europea” (art. 14 LOPDPP).

Sin embargo, la LUTICAJ, previa a la aprobación de estas normas, en su art. 42, se refiere a una “actuación judicial automatizada”, definida en su Anexo como la “actuación judicial producida por un sistema de información adecuadamente programado sin necesidad de intervención de una persona física en cada caso singular. Incluye la producción de actos de trámite o resolutorios de procedimientos, así como de meros actos de comunicación”. Como se argumentará a continuación, esta automatización no debe entenderse en sentido amplio, pues nunca debiera afectar a “resoluciones procesales motivadas o que afecten a derechos procesales fundamentales” (Pérez Estrada, 2022, p. 100).

Esto se debe a que automatizar el núcleo de las decisiones judiciales contravendría la Constitución Española, tanto por una posible vulneración del derecho fundamental a la tutela judicial efectiva (art. 24 CE) como por la sustitución de la labor de Juzgados y Tribunales, que son los únicos que pueden realizar la función jurisdiccional (117.3 CE). En cuanto a lo primero, es cierto que no se incluye de forma explícita en la Constitución que el juez ordinario haya de ser un juez humano. Sin embargo, sería errado igualar al Juez o Magistrado con una máquina pues, siguiendo la CEEA y el antevisto principio “bajo control del usuario”, el uso de la IA debe ser facultativo y la autonomía del juez-usuario debe aumentarse, no restringirse. En ello incide la Propuesta de Reglamento de inteligencia artificial, debiendo poder el juez “en cualquier situación particular, no emplear el sistema de inteligencia artificial o, si no, no tomar en consideración, ignorar o revocar su resultado” (Solar Cayón, 2022, p. 413). Sin acudir ya al *softlaw* o a la futura regulación sobre la IA, el segundo aspecto lo podemos relacionar con la redacción del Título VI de la Constitución: un juez robot no podría “jubilarse” (117.2 CE) ni tendría la posibilidad, por ejemplo, de estar vinculado a partidos o sindicatos (127 CE). De una interpretación sistemática de la norma constitucional se puede inferir sin riesgo a equivocarse que los Jueces y Magistrados (humanos) predeterminados por la ley son los únicos que pueden juzgar y hacer ejecutar lo juzgado en España. Para tomar parte en la función jurisdiccional, una IA no podría juzgar sin una reforma constitucional. Cuestión futura a dilucidar es cuál es el núcleo de la función jurisdiccional.

No obstante, se plantea la duda de qué papel puede jugar la IA en los mecanismos alternativos heterocompositivos de resolución de conflictos cuyas resoluciones son vinculantes para las partes, es decir, el Arbitraje (*Ley 60/2003, de 23 de diciembre, de Arbitraje*). Para algunos autores estos pueden ser el “banco de pruebas” antes de cualquier uso ulterior en la Administración de Justicia (Volokh, 2018, p. 1160).

En el Arbitraje, donde las partes disponen del objeto del litigio, no vulnera ni la exclusividad judicial ni la tutela judicial efectiva someter el asunto a un tercero³². ¿Por

³² Sobre la posible vulneración de la tutela judicial efectiva por someter las disputas a un tercero se pronuncia el TC: “Tal planteamiento, sin embargo, no puede ser compartido, ya que supondría privar al arbitraje, cuya licitud constitucional hemos declarado reiteradamente (SSTC 43/1988, 233/1988, 15/1989, 288/1993 y 174/1995), de su función como medio heterogéneo de arreglo de controversias que se fundamenta en la autonomía de la voluntad de los sujetos privados; lo que constitucionalmente le vincula con la libertad como valor superior del ordenamiento (art. 1.1 CE). De manera que no cabe entender que, por el hecho de someter voluntariamente determinada cuestión litigiosa al arbitraje de un tercero, quede menoscabado y padezca el derecho a la tutela judicial efectiva que la Constitución reconoce a todos” (STC 176/1996, de 11 de noviembre, F.J. 4).

qué no a una máquina? Que la IA pueda arbitrar no es un fenómeno extraño en los contratos privados, ya que se asemejaría, con matices, a lo que sucede en los denominados contratos inteligentes (*smart contracts*) mediante tecnología *blockchain* o los suscritos por agentes artificiales (*artificial agents*), tema que merece un estudio más allá de estas líneas³³.

Es también conocida la existencia de mecanismos alternativos para la resolución en línea de disputas (ODR), como el de la UE en materia de consumo, el Centro de Resolución de eBay, Modria, Cybersettle y Weclaim, que por ahora proponen medios autocompositivos para encontrar una solución o derivan eventualmente a un órgano de arbitraje³⁴. Particularmente, destaca el caso de Cybersettle al ser uno de los primeros ODR en aplicar una resolución automática al proceso, con el conocido método del “*double-blind binding*”, en el que cada parte propone, sin conocer la propuesta de la otra, la cifra máxima o mínima de la cantidad que está dispuesta a aceptar para satisfacer su disputa, resultando la posibilidad de acuerdo si las cifras propuestas se solapan (Susskind, 2017, p. 163).

En un futuro, no parece descabellado que la legislación se abra a que en ciertas áreas delimitadas exista un “árbitro-robot”, la denominada “*robotration*” (Snijders, 2021). La revisión de sus decisiones plantea problemas, pues debiera ser flexible en todo caso, no como la impugnación de un laudo arbitral dictado por un árbitro humano, sirviendo así para mejorar la confianza en las decisiones de la máquina. Con todo, la automatización del arbitraje plantea algunas dificultades técnicas, pues la IA precisa de una gran cantidad de datos para aprender a realizar una tarea y en el arbitraje aquellos son confidenciales (Scherer, 2019, p. 15).

3.1.3 Límites técnico-jurídicos: la transparencia

La transparencia es un requisito legal que ha de orientar el despliegue de la IA, recogiendo el RGPD los principios de “licitud, lealtad y transparencia” (art. 5 RGPD). La transparencia es también un problema técnico que, por la complejidad de los algoritmos

³³ Los conocidos como contratos inteligentes (*smart contracts*) tienen su origen en tecnología *blockchain* como Ethereum. Cuando dos usuarios de Ethereum celebran un acuerdo, pueden traducir el cumplimiento de ese acuerdo en un *script* de *software*. Las decisiones que se toman versan sobre hechos fácilmente determinables, como por ejemplo una apuesta deportiva, sin embargo, no pueden valorar cuestiones abstractas como la buena fe contractual (Ortolani, 2019, pp. 430–448; Arruñada, 2018, p. 8)

³⁴ Así, como ejemplos de plataformas que, directa o indirectamente, contienen una lista de árbitros o mediadores a los que derivar: ANDRC, Cibertribunal Peruano, Electronic Courthouse, Eucon, ICC, JAMS, Mediation Now, Mediation Room, Forrest Mosten, ODRWorld, Private Judge, Resolution F., Settle the Case, SmartSettle, TRUSTe, WIPO, Camera Milano, V. Courthouse, Arbitrare (Carneiro, Novais, Andrade, Zeleznikow, & Neves, 2014)

y las RNP, no se logra paliar en todas las ocasiones. Se trata del conocido problema de la caja negra o “*black box*”, que afecta fundamentalmente a los sistemas basados en algoritmos de aprendizaje profundo (*deep learning*): si bien la existencia de múltiples capas de redes neuronales artificiales posibilita que estos algoritmos puedan analizar simultáneamente la interacción recíproca de decenas o cientos de indicadores para arrojar una predicción o un resultado más preciso, esa misma complejidad en los niveles de análisis hace imposible conocer “el peso relativo que ha tenido cada uno de los indicadores analizados en el resultado” (Solar Cayón, 2022, p. 405).

La transparencia se puede entender como la capacidad de la IA de ser fiscalizada o fiscalizable (Martínez Martínez, 2019, p. 75) o “la capacidad de saber qué datos se utilizan, cómo se utilizan, quiénes los utilizan, para qué los utilizan y cómo se llega a partir de los datos a tomar las decisiones” (Sangüesa i Solé, 2018). Puede servir para cuatro fines: i) posibilitar la participación del ciudadano en los procesos que puedan afectarle; ii) dar a conocer al interesado la forma en que se ha adoptado una decisión para poder impugnarla; iii) promover el debate público y iv) prevenir que se lleven a cabo conductas ilícitas (Arnanz, 2021, pp. 101-102).

La transparencia técnica está íntimamente ligada con la rendición de cuentas (*accountability*). Tomando un ejemplo de Catherine H. O'Neil, si en un concierto se dispone que no se pueden ocupar las primeras filas de asientos, esa decisión en principio podría parecer arbitraria. Sin embargo, si se explica que las plazas en esta fila están reservadas para personas con movilidad reducida, entonces a juicio de todos estaría justificada la decisión. Así, la autora concluye: “la transparencia importa, marca la diferencia” (O'Neil, 2018, p. 26).

Pese al deber genérico de transparencia, en algunos campos podría carecer de importancia conocer cómo llega la IA a sus resultados. Esto sucede a menudo cuando los posibles errores no tienen gran impacto y los aciertos tienen una gran utilidad. No obstante, en otras áreas la falta de transparencia, como veremos, puede ser una causa fundamental para descartar el uso de una máquina. Especialmente si las decisiones erróneas individuales resultan en un peligro para la vida, la salud y el patrimonio de las personas, como se da en la conducción autónoma de vehículos, en las finanzas o, particularmente, en la Justicia (Coca Vila, 2017; Nieva Fenoll, 2018; Solar Cayón, 2022).

Las dificultades para lograr una total transparencia en el ámbito de la Administración de Justicia pueden obedecer, como expone José Ignacio Solar Cayón, a dos tipos de razones principales: por un lado, a razones jurídicas, como la protección del secreto de los algoritmos desarrollados por compañías privadas al ser considerados propiedad intelectual -como el algoritmo COMPAS, al que más tarde me referiré-, y por otro, a las ya referidas razones técnicas que impiden la transparencia de los resultados de los sistemas basados en múltiples capas de redes neuronales profundas (Solar Cayón, 2022). El primer tipo de obstáculos pueden ser fácilmente salvables, pues dependen del propio Derecho. Sin embargo, en el caso de los algoritmos de aprendizaje profundo, ni siquiera el acceso público y abierto al código fuente garantiza la transparencia, en cuanto los resultados del sistema pueden resultar incomprensibles no solo para los ciudadanos sino también para los expertos en datos e IA (Sangüesa i Solé, 2018). Para intentar desentrañar este problema técnico ha de acudir a las “explicaciones algorítmicas”, es decir, proporcionar una información sobre la lógica aplicada por el sistema que no consiste en un conocimiento de los pormenores técnicos, sino en explicar la decisión particular que afecta al usuario final (Pérez Bernabeu, 2021). Sin embargo, las técnicas para conseguirlo aún no se han desarrollado en todos los casos (Samek & Müller, 2019, pp. 5-22). Por ello, algunos algoritmos corren (por ahora) el riesgo de no cumplir el criterio de transparencia.

3.1.4 Límites de legitimidad

Por último, nos encontramos con el límite de la legitimidad, debido a que existe disenso en la opinión pública acerca de la función que puede desempeñar la IA en los tribunales (Morales Moreno, 2021). En un Estado Social y Democrático de Derecho, el papel a jugar por los algoritmos en la adopción de decisiones judiciales ha de pasar necesariamente por su aceptación por parte de la ciudadanía, pues, al final, todos podríamos estar sujetos a sus resultados.

Mientras que algunos autores sostienen que el uso de la IA vendrá impulsado por el propio desgaste de las instituciones del Estado, descargando a estas de sus propios errores (Vivar Vera, 2021, p. 258), otros sostienen que la legitimidad de las decisiones automáticas será un problema acomodaticio que se solventará en el futuro, cuando se deje de desconfiar en las decisiones de la máquina (Nieva Fenoll, 2018, p.14). Empero hoy surgen cuestiones de índole sociológica que minan la legitimidad de la IA, como la distribución del denominado “capital digital”, formado tanto por competencias como por recursos tecnológicos y concentrado predominantemente en segmentos de población joven,

urbana, masculina, con rentas altas y estudios, a la vez que muchos ciudadanos continúan sin acceso a tecnologías más básicas, siendo estos los nuevos vulnerables o “vulnerables tecnológicos” (A. Meroi, 2022; Ragnedda, Ruiu, & Addeo, 2020).

La desigualdad en el reparto del “capital digital” nos llevaría al problema de garantizar la independencia de quien programa el algoritmo: ¿dejaríamos el *ius puniendi* en manos de una empresa? ¿qué independencia tendría el programador? Para superar esta futurible “tiranía” se deberían tomar ciertas medidas, ya que, según Bonet Navarro, “el programador comparte el ejercicio de función jurisdiccional, de modo que ha de estar sometido a un régimen equivalente al judicial en cuanto a garantías, principalmente en lo relativo a la independencia e imparcialidad” (Navarro, 2020, p. 97).

Sin embargo, quizás una de las más importantes razones del rechazo a la IA en la Administración de Justicia sea la carencia de empatía de la máquina (Nieva Fenoll, 2018). Esta razón está relacionada en cierta medida con la esperanza de que un humano sintiente aplique las leyes, no en rigor, sino en equidad. Aunque este extremo, que se puede relacionar con la ya vista discrecionalidad judicial, cae en la paradoja de que la discrecionalidad humana es el cobijo de los sesgos. Únicamente una ausencia de empatía acabaría con ellos. Sin embargo, se podría argüir que esta solución poco tiene que ver con el necesario enfoque antropocéntrico ha de tener la implementación de la IA.

Ejemplos como Prometea nos muestran que nos encontramos en fase de preguntar ya no por la posibilidad, sino por la legitimidad de la inclusión de algunos sistemas de IA en los Tribunales (A. Meroi, 2022, p. 155). Habríamos de fijarnos bien en aquellos sistemas de IA que son técnicamente viables, solo si el ordenamiento legal lo permite o si estos poseen la capacidad de explicar su decisión al justiciable u operador jurídico.

CAPÍTULO IV: POSIBLES USOS DE LA INTELIGENCIA ARTIFICIAL EN EL PROCESO

Tomando en cuenta la diferencia entre tramitación y enjuiciamiento, este trabajo se aviene a la distinción entre diversos niveles de riesgo seguida por la Propuesta de Reglamento Europeo sobre la IA, por creer que puede ser el mejor abordaje para un tema que, previsiblemente, se regirá por la futura norma cuando esta se apruebe³⁵. El reglamento distingue, por un lado, aquellos sistemas destinados a ayudar a las autoridades judiciales a investigar e interpretar los hechos y el Derecho, y a aplicar la ley a unos hechos concretos, y, por otro, aquellos que, operando en la Administración de Justicia, son accesorios a estos fines. Para trazar un esquema de estos dispositivos se puede acudir a una triple distinción de ámbitos:

1. **Ámbito de la tramitación (riesgo bajo o inexistente)**, donde se ubicarían sistemas de IA que realizan labores accesorias.
2. **Ámbito de la investigación procesal (riesgo alto)**, donde incluimos aquellos sistemas de IA cuya función es ayudar a las autoridades judiciales a investigar e interpretar los hechos.
3. **Ámbito de la decisión judicial (riesgo alto)**, en el que incluimos los sistemas de IA empleados para interpretar el Derecho y aplicarlo a unos hechos concretos.

A continuación veremos, basándonos en estos tres niveles, los posibles usos de la IA en el proceso.

4.1 ÁMBITO DE LA TRAMITACIÓN

La tramitación, es decir, las labores accesorias y cotidianas que ha de realizar un juzgado a la hora de enjuiciar, constituyen un área libre para el desarrollo de la IA. Así, siguiendo el proyecto de Reglamento sobre IA, no se debe “obstaculizar el desarrollo y el uso de enfoques innovadores en la Administración pública, que se beneficiarían de una mayor utilización de sistemas de IA conformes y seguros, siempre y cuando dichos sistemas no

³⁵ Esta no es la única manera de abordar el tema, aunque es quizás la más útil entre las taxonomías posibles, seguida tanto por el magistrado Antonio del Moral García como por el profesor José Ignacio Solar Cayón, quienes, de conformidad con la propuesta de Reglamento, catalogan los sistemas de IA empleados en la Administración de Justicia según las tareas que realizan, llevando a cabo una aproximación basada en la evaluación de sus riesgos (del Moral García, 2022; Solar Cayón, 2022). Otras taxonomías, como la dada por la CEPEJ, se basan en los objetivos perseguidos: acceso a la justicia, comunicación entre los tribunales y los profesionales, administración del tribunal y asistencia directa al trabajo del juez y el secretario (European Commission for the Efficiency of Justice, 2016).

conlleven un alto riesgo para las personas jurídicas y físicas” ((EM37) COM(2021) 206 final 2021/0106(COD)). Entre los casos de uso de la IA en estas actividades de tramitación podemos distinguir aquellos que ya están presentes en algunos tribunales y los que se prevén en un futuro gracias al desarrollo del marco normativo.

4.1.1 Usos presentes en la Administración de Justicia

Uno de los casos más interesantes de implementación de herramientas en el presente es el que nos ofrece el Tribunal de Justicia de la Unión Europea, que dispone de su propio centro de innovación, *Innovation Lab*. Las herramientas desarrolladas por este centro sirven a variadas funcionalidades, como el reconocimiento óptico de caracteres (*Optical Character Recognition*), el escaneo de documentos; el reconocimiento automático de voz (*Automatic Speech Recognition*) para la transcripción de las sesiones en sala de juicios; la traducción automática de textos a otros idiomas (*eTranslation*), la búsqueda de jurisprudencia y la resolución de dudas legales mediante respuestas automáticas obtenidas a través de un *chatbot* (Baena Pedrosa, 2020; Solar Cayón, 2022).

En el caso de España hay que destacar el esfuerzo en la digitalización de expedientes y en la obtención de datos e información de calidad, paso previo para cualquier uso de técnicas de aprendizaje automático o -tal vez en un futuro más lejano- para la implantación de tribunales en línea (Solar Cayón, 2022). En el informe de la Comisión Europea *Study on the use of innovative technologies in the justice field*, de 2020, se exponen hasta siete proyectos actualmente en proceso de desarrollo en España, tres llevados a cabo por el Ministerio de Justicia y cuatro por el Centro de Documentación Judicial (CENDOJ), que se encontraban en aquel momento en diferentes etapas de diseño.

Por parte del Ministerio de Justicia, se presenta como primer proyecto la transcripción de habla a texto mediante el programa Textualización de Vistas, que veremos más adelante; en segundo lugar, un proyecto de automatización de tareas documentales sirviéndose de técnicas de aprendizaje automático y PLN, entre cuyas funcionalidades se recoge, por ejemplo, la detección de cláusulas que pueden dar lugar a controversias en documentos y contratos; y por último, el diseño de una herramienta basada en el aprendizaje automático y el reconocimiento de imágenes para posibilitar el acceso mediante datos biométricos a servicios públicos y la justicia (European Commission, 2020, pp. 214-215).

Por parte del CENDOJ se estaba desarrollando o planeando, al tiempo del informe, un proyecto de optimización de búsqueda de sentencias, usando aprendizaje profundo y

PLN; otro de creación de datos estructurados que permitan optimizar las funcionalidades de búsqueda; un proyecto de *business intelligence*, que sirva tanto a la optimización de criterios de búsqueda de documentos como a obtener conocimientos sobre su contenido, como oraciones, legislación, publicaciones y documentos del Ministerio Fiscal; y un programa para la anonimización y seudonimización de sentencias (European Commission, 2020, p. 215).

Por otro lado, ya se pueden encontrar activas o en desarrollo tanto en el Territorio Ministerio como en Comunidades con competencias delegadas en la Administración de Justicia algunas herramientas de IA. Merecen mención las tres presentadas en noviembre de 2019 por el Ministerio de Justicia, alguna coincidente con las expuestas en el anterior informe de la Comisión Europea. Estas son:

4.1.1.1 Calculadora 988

La Calculadora 988 es la primera herramienta de IA empleada en la Justicia penal española. Utiliza un algoritmo creado por Subdirección General de Nuevas Tecnologías de Justicia del Ministerio de Justicia que sirve para calcular las penas de los condenados cuando existe una acumulación de sentencias. Para su fórmula usa la alternativa más favorable al reo, basándose en el art. 988 de la *Ley de Enjuiciamiento Criminal*, el art. 76 del *Código Penal* y el *Acuerdo de 27 de junio 2018 del Pleno No Jurisdiccional de la Sala Segunda del Tribunal Supremo*. La herramienta, que cuenta con más de cinco mil usuarios, libra de errores en el cálculo e impone un único criterio para realizar las operaciones de acumulación de penas (SGNTJ, 2020a).

4.1.1.2 Textualización de Vistas

Esta herramienta transcribe las vistas o declaraciones del vídeo a soporte textual que facilite la localización del contenido mediante palabras clave. La Textualización de Vistas se encuentra en continuo desarrollo, usando un algoritmo de aprendizaje automático, en especial de aprendizaje por refuerzo, mejorando a medida que los usuarios lo utilizan. La dificultad consiste en distinguir con claridad los distintos registros lingüísticos adaptándose “a rasgos culturales concretos como expresiones, acentos etc., así como a acepciones o conceptos jurídicos, que el sistema tiene que aprender de forma previa para incrementar sus prestaciones” (SGNTJ, 2020b). La transcripción supone un ahorro de un 60% del tiempo de búsqueda en las grabaciones de los juicios.

4.1.1.3 La firma digital manuscrita

Por último, se ha desarrollado una aplicación que recoge la rúbrica de los ciudadanos que no disponen de firma electrónica cuando estos han de comparecer ante la Administración de Justicia. El procedimiento anterior obligaba en tales casos a imprimir y escanear el documento físico firmado, mientras que, mediante el nuevo sistema, la firma se realiza a través de un dispositivo electrónico portátil que incorpora la firma directamente al expediente digital (Dolz Lago, 2022).

4.1.2 Usos futuros en la Administración de Justicia

El *Anteproyecto de Ley de Medidas de Eficiencia Digital del Servicio Público de Justicia* contiene la incorporación de técnicas de inteligencia artificial como instrumento de apoyo al ejercicio de la actividad judicial mediante “la producción de actuaciones judiciales y procesales automatizadas, asistidas y proactivas” (art. 35 APLED). El art. 56 APLED, que contiene las actuaciones automatizadas y proactivas, supone un desarrollo y ampliación del art. 42 LUTICAJ. No obstante, el art. 57 APLED excede el enfoque de riesgo planteado, es decir, ya no es tramitación, cabiendo enunciar en este apartado únicamente las actuaciones Automatizadas y las Proactivas, por dedicarse a tareas accesorias.

4.1.2.1 Las Actuaciones Automatizadas

Las actuaciones automatizadas tienen un precedente en la ya vista “actuación judicial automatizada” (art. 42. LUTICAJ). No obstante, en el APLED, a diferencia de la cláusula abierta de la LUTICAJ³⁶, se recogen los supuestos susceptibles de automatización. El Anteproyecto define actuación automatizada como “la actuación procesal producida por un sistema de información adecuadamente programado sin necesidad de intervención de una persona física en cada caso singular” (art 56.1 APLED). Estas versarán sobre

³⁶ La LUTICAJ no contiene un listado, pero se refiere de forma explícita a ciertos ejemplos de actuaciones judiciales automatizadas relacionadas con el expediente digital, tales como: el escaneo de documentos (“Las oficinas judiciales podrán obtener imágenes electrónicas de los documentos privados aportados por los ciudadanos, con su misma validez y eficacia, a través de procesos de digitalización que garanticen su autenticidad, integridad y la conservación del documento imagen, de lo que se dejará constancia” (art. 28.3)), la emisión de recibos de los documentos presentados (“Los registros electrónicos emitirán automáticamente un recibo consistente en una copia autenticada del escrito, documento o comunicación de que se trate, incluyendo la fecha y hora de presentación y el número de entrada de registro” (art. 31.1)) o el examen de los requisitos de formato del documento (“Los documentos que se acompañen al correspondiente escrito o comunicación, deberán cumplir los estándares de formato y requisitos de seguridad que se determinen en el marco institucional de cooperación en materia de administración electrónica. Los registros electrónicos generarán recibos acreditativos de la entrega de estos documentos que garanticen la integridad y el no repudio de los documentos aportados, así como la fecha y hora de presentación y el número de registro de entrada en la correspondiente sede judicial electrónica” (art. 31.2)).

actuaciones “de trámite o resolutorias simples, que no requieren interpretación jurídica” (art. 56.2 APLED). Esta última especificación, que puede encontrarse implícitamente en la LUTICAJ, es de relevancia debido al avance de la IA en los últimos tiempos.

Las actuaciones automatizadas se encuentran sometidas a un control común, junto con las proactivas y las asistidas, realizado por parte del Comité Técnico Estatal de la Administración Judicial Electrónica, para “la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, la auditoría del sistema de información y de su código fuente” (art. 58.1 APLED), y a uno específico, debiendo ser sus criterios de decisión “públicos y objetivos, dejando constancia de las decisiones tomadas en cada momento” (art. 58.2 APLED). Este requisito daría la espalda a sistemas de “caja negra”, debiendo tomar siempre decisiones trazables.

En cuanto al listado de tareas automatizables, se compone de monótonas labores del servicio de auxilio judicial. Su automatización ha sido reivindicada por el mismo cuerpo de auxiliares en los últimos años, al ocupar un tiempo que podría ser dedicado a labores no susceptibles de ser realizadas por una máquina (Lorenzo Pérez, 2022). Estas tareas son: “a) el numerado o paginado de los expedientes b) la remisión de asuntos al archivo cuando se den las condiciones procesales para ello c) la generación de copias y certificados d) la generación de libros e) la comprobación de representaciones, y f) la declaración de firmeza, de acuerdo con la ley procesal” (art. 56.2 APLED).

4.1.2.2 Las Actuaciones Proactivas

Están contenidas en el artículo 56.3 APLED como “actuaciones auto-iniciadas por los sistemas de información sin intervención humana, que aprovechan la información incorporada en un expediente o procedimiento de una Administración pública con un fin determinado para generar avisos o efectos directos a otros fines distintos, en el mismo o en otros expedientes, de la misma o de otra Administración pública, en todo caso conformes con la ley”. Se trata de un subtipo de decisiones automatizadas, poco claro en cuanto a su redacción³⁷, cuya aplicación podría garantizar, por ejemplo, que no existan

³⁷ En tal sentido se pronuncia el CGPJ en su informe sobre el anteproyecto de la ley: “Las referencias a las Administraciones públicas y a los expedientes tramitados por ellas revelan una falta de conexión de la regulación de estas actuaciones proactivas con el ámbito propio de la Administración de Justicia y, en particular, la tramitación de los procedimientos judiciales. Por otro lado, el objeto de estas actuaciones proactivas se define de forma bastante evanescente, pues aparte de la generación de avisos, cuyo significado se entiende perfectamente, la posibilidad de generar “efectos directos a otros fines distintos” constituye un arcano difícil de descifrar. Por todo ello, es necesario que por parte del prelegislador se reformule el artículo

desahucios sin alternativa habitacional, dado que podría servir para comunicar a servicios sociales la situación de vulnerabilidad de forma automática al decretarse el lanzamiento (150.4 LEC).

4.2 ÁMBITO PROCESAL Y/O PROBATORIO:

En esta materia la IA es prolífica, aunque se ha de seguir con cautela pues es un área que de alto riesgo. Cabe distinguir varios tipos de sistemas de IA dependiendo de sus tareas, con diferentes niveles de aceptación y uso en el plano internacional. Entre ellos el profesor José Ignacio Solar Cayón distingue: i) los sistemas de codificación predictiva para la selección del material relevante en el proceso; ii) la Inteligencia Artificial y *Blockchain* como medios de prueba y como herramientas para la valoración de los medios de prueba; iii) los sistemas algorítmicos de evaluación de riesgos de reincidencia criminal iv) y los sistemas de búsqueda y análisis de la información jurídica (Solar Cayón, 2020b).

4.2.1. Los sistemas de codificación predictiva

Uno de los mayores impactos procesales de la IA se ha dado en el mundo anglosajón, en concreto en el campo de la selección de la información electrónica relevante en el litigio. El Título V de las Reglas Federales en materia procesal civil de EE.UU. (*Federal Rules of Civil Procedure* o FRCP), así como las normas de otros países pertenecientes al *Common Law*, establecen una etapa procesal denominada *discovery* en la que las partes en el litigio han de aportar toda la información relevante sobre el objeto de la disputa, salvo aquella que esté protegida por un *privilege*³⁸(Hazard Jr, 1997; Oard & Webber, 2013). El *discovery* se extiende a todo tipo de información contenida en cualquier clase de formato. Pero en la actualidad, con la digitalización, la tarea de revisión y selección de la información ha pasado a tener por objeto en su mayor parte documentos y otro tipo de archivos electrónicos, así como correos electrónicos y toda clase de contenidos digitales (*Electronically Stored Information*, o ESI), por lo que es frecuente ya referirse a esta etapa del proceso como la fase de *e-discovery*. Además, la digitalización también ha provocado un incremento exponencial de la información disponible, y que en caso de litigio puede ser objeto de revisión.

56.3 APL con el fin de ofrecer una regulación más precisa y conectada con la Administración de Justicia de las denominadas actuaciones proactivas” (par. 159 CGPJ)

³⁸ El *privilege* es, precisamente, en relación con la exigencia de transparencia, la principal figura jurídica que convierte a los algoritmos en “cajas negras” pues “exceptúa *prima facie* a las compañías propietarias de la obligación de revelar el código fuente de sus sistemas” (Solar Cayón, 2020b, p. 141).

La satisfacción de este deber procesal de *discovery* obliga a la parte requerida (mediante la solicitud conocida como “*request for production*”) a realizar un esfuerzo razonable de búsqueda (“*reasonable search*”) de la información relevante que obre en su poder sobre el objeto del litigio, esfuerzo que en todo caso ha de estar regido por un principio de proporcionalidad (Solar Cayón, 2018)³⁹. Pero en algunos casos complejos esta tarea puede llevar mucho tiempo, teniendo en cuenta que en ocasiones es necesario revisar millones de documentos (Solar Cayón, 2018). Tradicionalmente, esta tarea de revisión de la información ha sido realizada por los paralegales y por abogados junior (Susskind, 2017, pp. 87-88), resultando además una de las fases procesales más costosa.

Hoy, sin embargo, puede ser realizada de manera automática por sistemas de IA diseñados específicamente para este fin. Esta revisión asistida por tecnología u ordenador (*technology-assisted review (TAR)* o *computer-assisted review (CAR)*) se lleva a cabo a través de los denominados sistemas de “codificación predictiva” (*predictive coding*).

La codificación predictiva se basa en un “algoritmo de aprendizaje automático activo que “aprende” los criterios de relevancia jurídica en relación a un caso concreto, permitiendo identificar los documentos relevantes de un universo a partir de un subconjunto de documentos que han sido previamente “codificados” (es decir, clasificados cada uno de ellos como relevante o no relevante) por un abogado experto” (Solar Cayón, 2019, p. 149). Este proceso inicial de entrenamiento del algoritmo a partir de una serie de ejemplos de la tarea a realizar ha de ser llevado a cabo por abogados que conozcan bien el caso. Pero, una vez que se halla bien entrenado, el sistema es capaz de “realizar la revisión automatizada de enormes volúmenes de datos registrados en cualquier tipo de formato digital (documentos de texto, imágenes y videos, audios, correos electrónicos, bases de datos, calendarios, hojas de cálculo, programas informáticos, comunicaciones a través de internet, etc.), procedentes de múltiples y heterogéneas fuentes (servidores, ordenadores, discos duros, memorias USB, tabletas, teléfonos móviles, correos electrónicos, CDs y DVDs, cintas de *backup*...), e identificar cualquier tipo de información relevante que pueda ser presentada como evidencia ante el tribunal” (Solar Cayón, 2022, p. 396).

³⁹ En tal sentido la Regla 26(b)(2)(C)(iii) de las *Federal Rules* establece que el Tribunal puede limitar el *discovery* cuando “la carga o el gasto del descubrimiento propuesto supere su posible beneficio, considerando las necesidades del caso, el monto en controversia, los recursos de las partes, la importancia de los asuntos en juego en la acción, y la importancia del descubrimiento para resolver los problemas” FED. R. CIV. P. 26(b)(2)(C)(iii).

La primera vez que se permitió la utilización de la codificación predictiva en el proceso judicial fue en el caso *Da Silva Moore v. Publicis Groupe & MSL Group* (287 F.R.D. 182 (S.D.N.Y. 2012), en el cual, ante la necesidad de revisar más de tres millones de correos y otras piezas de información electrónica, el Juez Andrew J. Peck autorizó a la parte demandada a usar herramientas de IA bajo ciertas condiciones, abriendo así la puerta a la generalización de su uso -incluso frente a la oposición de la parte requirente- siempre que la parte obligada a proporcionar la información se comprometiera a cooperar con su contrario en el diseño de un protocolo consensuado de entrenamiento del sistema y a garantizar la transparencia de su actuación (Solar Cayón, 2018, p. 95). Hoy en día, el empleo de la codificación predictiva para satisfacer el deber procesal de *discovery* se halla admitido en la mayoría de las jurisdicciones del *Common Law*, pudiendo ser considerado “el principal “caso de éxito” en el desarrollo de una inteligencia artificial jurídica dirigida al proceso judicial” (Solar Cayón, 2022, p. 395).

Tal como expone José Ignacio Solar Cayón (2022), los efectos jurídicos de este éxito se han hecho notar, sobre todo, en la superación de la cultura tradicional de confrontación entre las partes en esta fase del proceso y su sustitución por una cultura de cooperación en el diseño de una metodología consensuada para el empleo de esta herramienta. Actitud que resalta el Juez Peck mediante lo que denomina “estrategia de entrega de información proactiva” (“*strategic proactive disclosure of information*”), que puede evitar litigios, y el consecuente gasto que acarrearán gracias a que, algunos demandados, mediante el uso de la codificación predictiva, pueden conocer si las pretensiones del demandante están fundadas o no, antes incluso de la solicitud formal de *discovery* (Borden & Baron, 2014). La codificación predictiva también ha permitido superar algunos condicionantes que tradicionalmente imponía la exigencia de proporcionalidad en el desarrollo de esta fase procesal, modificando los parámetros de evaluación de la relación coste-eficacia (Solar Cayón, 2022 p. 397).

4.2.1. La IA y los medios de prueba y de evaluación de la prueba

Tal y como sucedía en las instituciones totales de las que hablara Erving Goffman, con la llegada de la IA se tienden a juntar diferentes esferas de vida, de modo que la conducta en un campo de actividad puede ser echada en cara en otros contextos (Goffman, 2001, p. 47). Con este estrechamiento de la intimidad es razonable esperar que el papel de la IA como medio de prueba sea cada vez mayor, pues, por un lado, existe una huella digital dejada de manera voluntaria o involuntaria y, por otro, cada vez se dispone de mayor

cantidad de dispositivos que controlan diferentes aspectos de nuestras vidas (Solar Cayón, 2022, p. 398).

En este contexto, el Derecho procesal ha de hacer frente a nuevos riesgos derivados del empleo de la tecnología para escudriñar toda esta información, tal como expone Cristina San Miguel Caso. Así, el actual Anteproyecto de Ley de Enjuiciamiento Criminal, en su artículo 516, habilita, a instancias del Ministerio Fiscal, al Juez de Garantías para autorizar la utilización de sistemas automatizados o inteligentes de tratamiento de datos para cruzar e interrelacionar la información disponible sobre la persona investigada, con otros datos obrantes en otras bases de titularidad pública o privada, si concurren ciertos requisitos (San Miguel Caso, 2021, p. 287)⁴⁰.

Algunas administraciones ya utilizan la IA en la persecución de ilícitos penales, debido a su capacidad para mejorar las labores detectivescas. Podemos recurrir al ejemplo del agente de policía que recibe, entre las decenas que tramita al mes, una denuncia semejante a otra interpuesta hace un año. En el mundo analógico aquel ha de poseer una gran memoria e intuición para poder encontrar patrones de conducta criminal que dieran con la detención del delincuente. La cuestión se complica más si, en vez de tramitar la denuncia ese mismo agente, su turno finaliza y lo hiciera otro. Sin embargo, un algoritmo, como el Fiscal Watson, es capaz de establecer patrones, sin tener en cuenta estas limitaciones, revisando las bases de datos de denuncias. Mediante este método en Colombia se pudo detener a un agresor sexual, el conocido como “violador de Tolima” (Medina Uribe & Fernanda Gómez, 2020).

Promete además el papel de la IA en el campo de la valoración de la prueba, en un contexto en el que se halla en decadencia el pretérito sistema legal rígido de valoración de prueba para dejar paso a nuevas aproximaciones multidisciplinares dentro de la libre valoración de la misma (Nieva Fenoll, 2018). Por ejemplo, en la prueba documental, la IA puede servir para reconocer la intencionalidad del autor, aunque, como hemos visto en las carencias de la IA “débil”, la valoración estaría limitada a la sintaxis y no a la semántica. También podrían servir para detectar inconsistencias en los testimonios, gracias a criterios como la existencia de contradicciones, la capacidad descriptiva de los

⁴⁰ Que son: a) que existan indicios basados en datos objetivos sobre la participación del investigado en los hechos objeto de investigación; b) que, en base a la naturaleza y características del hecho, resulte necesaria la práctica de la diligencia para esclarecer la responsabilidad del investigado en el mismo; y c) que el hecho investigado sea constitutivo de un delito castigado con una pena igual o superior a los tres años de prisión.

hechos, la existencia de indicios objetivos de prueba externa a la declaración o de comentarios retóricos que buscan afianzar una historia inventada (Nieva Fenoll, 2018).

4.2.2 Los sistemas de evaluación de riesgos de reincidencia criminal

El uso de algoritmos que predicen el riesgo de reincidencia criminal ha pasado del campo actuarial al ámbito penitenciario para, posteriormente, saltar al proceso judicial en algunos países. Es el caso de EE.UU., donde existen diversos sistemas de este tipo que, dependiendo de cada Estado, pueden ser empleados para la determinación de medidas cautelares (por ejemplo, la concesión o no de la libertad provisional en la fase de *pre-trial*) e incluso para la determinación de la extensión de la pena (en la fase de *trial*) en caso de sentencia condenatoria (Solar Cayón, 2020b, pp. 130-131).

Aunque sin llegar a entrar en el ámbito del proceso judicial, en España se usan herramientas de evaluación de riesgos como las aplicadas en las prisiones catalanas, e-Riscanvi, o por la *Direcció General d'Execució Penal a la Comunitat i Justícia Juvenil*, SAVRY (*Structured Assessment of Violence Risk in Youth*), así como el Sistema de Seguimiento Integral de los casos de Violencia de Género (VioGén) (Vilela Komatsu, Wenger, Costa, Rezende Bazon, & Pueyo, 2019). Todas ellas sirven al diseño de estrategias de reinserción de reclusos o a la adopción de protocolos policiales, funcionando de apoyo para instituciones como la Comisión para el Estudio de la Reincidencia en los Delitos catalana⁴¹.

El uso de métodos cuantitativos que miden la reincidencia, el enfoque conocido como “valoración o evaluación del riesgo”, supone el abandono del tradicional enfoque de la “peligrosidad”, apoyado fundamentalmente en informes psicológicos subjetivos, criticados por su falibilidad (Martínez Garay & Montes Suay, 2018, p. 5) y en los que se podía entrar a valorar cuestiones sospechosas de estar cargadas de prejuicios, tales como si los detenidos o enjuiciados son personas toxicómanas, marginales o de etnia gitana (Asociación Pro Derechos Humanos de España, 2015, p. 42). La teoría de la peligrosidad vendría además unida a formas de concebir el riesgo, por parte de los jueces, ajenas a la lógica, como visiones conservadoras, tendentes a dar mayor peso a la materialización de los riesgos, u optimistas, que dan mayor confianza a que el riesgo nunca llegue a

⁴¹ Creada mediante la Resolución JUS / 2362/2007, DOGC nº 4937, 24 de julio.

materializarse (Nieva Fenoll, 2018p. 62)⁴². A causa de las limitaciones del análisis subjetivo de peligrosidad, varios estudios, partidarios de la visión actuarial, consideran que el método estadístico tiene una objetividad superior al juicio humano (Grove et al., 2000).

En base a este nuevo enfoque priman las herramientas que usan aprendizaje automático y grandes volúmenes de datos para predecir la probabilidad de reincidencia criminal, valiéndose de variables como el historial delincriminal, las circunstancias personales, las circunstancias sociodemográficas, llegando a clasificar los resultados en escalas de riesgo (bajo, medio o alto) (Solar Cayón, 2020b, p. 129). Una de las herramientas de evaluación del riesgo que más protagonismo ha adquirido en la bibliografía anglosajona, tratada también por autores españoles como José Ignacio Solar Cayón, Carlos María Romeo Casabona o Lucía Martínez Garay, es COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), desarrollada por la compañía Northpointe⁴³. La constitucionalidad del empleo de esta herramienta fue cuestionada en el conocido caso *State v. Loomis*. Eric Loomis había sido condenado por un tribunal de Wisconsin que le impuso una sentencia más rigurosa porque COMPAS lo catalogó como un individuo de alto riesgo de reincidencia (Martínez Garay, 2018, p. 490). Y esta decisión fue recurrida ante el Tribunal Supremo de Wisconsin en el caso mencionado.

Eric Loomis alegaba la vulneración de su derecho a un debido proceso por tres razones: que no había podido conocer la fórmula por la que COMPAS le había clasificado de “alto riesgo”, pues el algoritmo del sistema fue considerado por el tribunal de instancia un secreto comercial; que el algoritmo discriminaba por razón de género; y que no había sido objeto de una sentencia individualizada⁴⁴. El tribunal desestimó la demanda de Loomis. En relación a la falta de transparencia del algoritmo, argumentó que, aunque Loomis no pudo acceder al código fuente del sistema, esto no vulneró el derecho a ser condenado sobre la base de información fiable y exacta, pues Loomis tuvo acceso tanto al informe

⁴² Se ha demostrado, por ejemplo, que el estado anímico de los jueces al tomar decisiones condiciona de manera significativa su decisión. Al hacer una pausa para comer los jueces deciden con mayor frecuencia en favor del reo, decayendo el número de decisiones favorables cuanto más tiempo pasen de trabajo continuado (Danziger, Levav, & Avnaim-Pesso, 2011).

⁴³ COMPAS clasifica la peligrosidad en una escala decimal, siendo el 10 la mayor peligrosidad, valiéndose de un cuestionario con 137 preguntas. Entre las dirigidas al reo o investigado hay algunas controvertidas, como si se está de acuerdo o no con: que una persona hambrienta robe (pregunta 127); que si una persona delinque es por no tener oportunidad de tener un buen trabajo (pregunta 128) o con que con si mucha de la gente que se salta la ley o consume drogas es causa de que la sociedad no les ha educado bien (pregunta 136).

⁴⁴ Vid. *State v. Loomis* ¶ 109-110.

de resultados de COMPAS como a los datos que tuvo en cuenta el algoritmo para llegar a su resultado⁴⁵. En relación con la discriminación, el tribunal también rechazó que la toma en consideración del género por parte del algoritmo fuera un factor discriminatorio, al constituir un dato criminógeno relevante, ya que las mujeres tienen tasas de reincidencia criminal notablemente inferiores a los varones⁴⁶. Y, por último, en cuanto al derecho a una sentencia individualizada, el tribunal entendió que solo habría sido vulnerado si el resultado de COMPAS hubiera sido determinante para la determinación de la sentencia, entendiendo que no lo había sido.

No obstante, para su futuro uso, el Tribunal Supremo de Wisconsin estableció una serie de cautelas: fundamentalmente, que el resultado del sistema no puede ser vinculante para el juez y que la decisión de este nunca puede estar basada únicamente en ese resultado. Adicionalmente, obligó a realizar cinco advertencias cuando se utilicen este sistema, al objeto de valorar adecuadamente sus resultados: 1) que la naturaleza secreta del algoritmo COMPAS impide el conocimiento de la fórmula de valoración del riesgo que usa; 2) que COMPAS no puede identificar individuos específicos de alto riesgo porque sus puntajes se basan en datos grupales; 3) que no se han realizado estudios que avalen la eficacia de COMPAS en Wisconsin; 4) que algunos trabajos han encontrado que COMPAS discrimina a minorías y 5) que COMPAS está originalmente diseñado para su uso en el ámbito penitenciario⁴⁷.

El sistema COMPAS ha generado una enorme polémica, al haber sido objeto de un proceso de ingeniería inversa cuyos resultados han puesto en cuestión su fiabilidad, encontrando también sesgos discriminatorios respecto de minorías (Dressel & Farid, 2018). Ya en 2016 activistas de la ONG periodística ProPublica concluyeron que los afroamericanos son más propensos a ser clasificados falsamente por COMPAS como individuos de "alto riesgo" de reincidencia, mientras es más probable que clasifique erróneamente a los blancos como de "bajo riesgo"⁴⁸. Según ProPublica, que toma como muestra del desempeño del algoritmo a más de 7000 sujetos de un condado de Florida,

⁴⁵ Vid. *State v. Loomis* ¶ 47 y ss.

⁴⁶ Vid. *State v. Loomis* ¶ 83.

⁴⁷ Vid. *State v. Loomis*, ¶ 98-101.

⁴⁸ Entre mayo de 2016 y diciembre de 2017 se publicaron cerca de 578 artículos académicos sobre el sesgo del algoritmo (Washington, 2018). La empresa propietaria de COMPAS respondió a la acusación de ProPublica usando sus mismos datos con otra noción de equidad: la paridad predictiva. En sentido técnico, la reincidencia era independiente de las características raciales en los clasificados como de "alto riesgo" (en torno a un 60% en estos casos) (Dieterich, Mendoza, & Brennan, 2016).

las tasas de falsos positivos, es decir, de personas clasificadas como de alto riesgo y que a la postre no reinciden, es de 45% en la población afroamericana frente a 23% en la blanca, mientras que las personas clasificadas con un perfil de bajo riesgo y que terminaban reincidiendo -falsos negativos- fue de 48% en la población blanca y 28% en la negra (Angwin, Larson, Mattu, & Kirchner, 2016).

El uso de los algoritmos de predicción de reincidencia en la imposición de sentencias, tal como permite la Corte Suprema de Wisconsin, ha sido ampliamente criticado, porque podría llevar a una desproporción en las penas impuestas, e incluso a una distorsión en el espacio entre el enjuiciamiento (*choosing a verdict*) y la culpabilidad (*given a verdict verdict*) (Pruss, 2021). Es decir, al conocer el juez el resultado de “riesgo alto” de reincidencia del acusado antes de decidir si debe entrar provisionalmente en prisión, puede que se guíe por este riesgo abstracto y no por la relación con los hechos que se le imputan. En caso de decretar la prisión provisional, esa decisión supone “prácticamente una certeza de culpabilidad, lo que, ciertamente, equivale al juicio que se emitirá con la sentencia” (Nieva Fenoll, 2018, p. 72).

Si bien el órgano sentenciador no está vinculado por lo que le propone el algoritmo, cabría preguntarnos si se puede producir el llamado “sesgo de automatización”, estudiado con profusión en el campo de la seguridad en los transportes. Este sesgo, comparable con otros heurísticos que rigen el juicio humano, consiste en guiarse de forma complaciente por las indicaciones que le proporciona la máquina, el piloto automático o cualquier otro sistema guía, aunque de no existir este se hubiera obrado de forma diligente, comprobando si los datos son ciertos, resultando así en acciones u omisiones erradas (Parasuraman & Manzey, 2010).

En cuanto a la interpretación de los resultados, esta tarea no es fácil, pudiendo resultar un problema que el juez tenga que distinguir entre riesgo absoluto (el riesgo de reincidencia de un grupo concreto en un tiempo determinado) y relativo (el riesgo de reincidencia del individuo respecto de su grupo delincencial) (Martínez Garay & Montes Suay, 2018, p. 5). Además, se ha sostenido que el uso de algoritmos opacos en el enjuiciamiento vulneraría el derecho fundamental a la tutela judicial efectiva (Simón Castellano, 2021). Los partidarios de la inclusión de sus resultados como un elemento más a tomar en cuenta en las decisiones judiciales rechazan este argumento, alegando que descomponer el algoritmo para explicar la decisión, cuando este sirve junto con otros datos para auxiliar la decisión judicial, sería como exigir al experto, no los motivos por los que lo es, “sino

todos y cada uno de los elementos que le han conducido a ser un experto” (Castellanos Claramunt & Montero Caro, 2020, p. 80). En base a este argumento, Hans Reichenbach (1938) distingue dos conceptos en la investigación científica que guardan importancia en la argumentación judicial y la opacidad algorítmica. Por un lado, el “contexto del descubrimiento” como el conjunto de factores psicológicos, culturales, o emocionales que llevan a explicar el surgimiento de una idea, por otro, el “contexto de la justificación”, como argumentación y encaje lógico del descubrimiento. El problema del algoritmo es que descubrimiento y justificación forman una tautología de difícil, si no imposible, solución (Fouquet, 2021, pp. 267-269).

Si estas herramientas son finalmente aceptadas a la hora de decretar medidas cautelares o penas, deberían obrar siempre en favor del reo, tal como sostiene Mario Caterini (2022). El auxilio al Juez debería consistir en indicar la argumentación jurídica más favorable al investigado, una especie de “*default option*” que sería, “no la estadísticamente más frecuente, sino la más favorable al imputado” (Caterini, 2022, p. 11). Cabría pensar cómo se puede lograr tal cosa con este tipo de herramientas, diseñadas para predecir el riesgo, sin modificar la ponderación de las variables que realiza el algoritmo, lo que podría traer consigo imprecisión⁴⁹.

Como propuesta, creo que en estos casos solo superaría el principio *favor rei* un algoritmo de evaluación del riesgo que no se utilizase como guía o criterio para una mayor pena, sino como alarma al juez de un posible exceso punitivo. Es decir, una vez tomada su decisión, mediante los medios subjetivos tradicionales, el juez ha de pasarla a una escala, consultando el riesgo con la máquina al estilo de un oráculo, sin conocer de antemano el criterio de esta, como el ya visto “*double-blind binding*”. Así, cuando el riesgo valorado valiéndose de otros medios de prueba fuera mayor que el obtenido por el algoritmo, este avisaría de su disparidad, y no en el caso contrario o de ser coincidentes, lo que debiera de conducir a una reflexión sobre por qué el juez o los peritos encuentran discordancias con la máquina. Su uso debe estar dirigido únicamente para garantizar un castigo proporcional, nunca para agravarlo.

⁴⁹ Sobre esta cuestión, existen diferentes estrategias para abordar la imparcialidad de los resultados, pudiendo clasificarse en: anti-clasificación, que omite las categorías protegidas que pueden ser discriminatorias; de paridad clasificatoria, la usada por COMPAS para replicar a ProPublica, que se basa en un rendimiento igual entre categorías sospechosas de discriminación y por último, la calibración, que busca una media entre los distintos grupos con categorías sospechosas de discriminación, a costa, en muchos casos, de la precisión en cada uno (Solar Cayón, 2020b, pp. 160-161). Esta reflexión se refiere a la última.

4.2.3 Sistemas de búsqueda de información jurídica

Los avances en el ámbito de la búsqueda de información también pueden servir para interpretar los hechos y el Derecho, gracias al surgimiento de nuevas herramientas, como la plataforma cognitiva ROSS. Creada por estudiantes de la Universidad de Toronto, ha sido considerada la “más intrigante de las nuevas aplicaciones legales” (Ashley, 2017, p. 351). Se trata de una aplicación de búsqueda de respuestas (*Question Answering*) que se encuadra dentro de la categoría de la búsqueda de información legal automática (*automated legal research*). Basada en la plataforma Watson, de IBM, responde a preguntas legales formuladas en lenguaje natural, como, por ejemplo, “¿Puede una empresa en quiebra seguir haciendo negocios?”. En respuesta a las preguntas formuladas, ROSS elabora un informe estructurado de aproximadamente dos páginas en el que “razona” jurídicamente sobre el asunto planteado, seleccionando normas, jurisprudencia y otros textos jurídicos, ponderando la adecuación de las diversas respuestas posibles a partir de los argumentos contenidos en esos materiales y ofreciendo, finalmente, aquella que, en virtud de una serie de criterios, tiene mayores probabilidades de éxito (Ashley, 2017; Remus & Levy, 2017).

4.3 ÁMBITO DE LAS DECISIONES JUDICIALES

En este ámbito, de riesgo elevado, la IA puede tener utilidad, como hemos visto, en campos alternativos a la jurisdicción, como la resolución de disputas extrajudiciales en línea, cuyos objetivos son la ampliación del acceso a la justicia y la solución de las disputas “a través de un cauce fácil, rápido y barato que pueda ser utilizado por las partes en cualquier momento y desde cualquier lugar” (Solar Cayón, 2019, p. 170). Realmente estos sistemas son útiles asistiendo a los individuos para encontrar una solución, descargando de tareas a los órganos jurisdiccionales, pero quedan fuera del ámbito estricto de la decisión judicial. Es en esta área, el de la decisión judicial, ya se pueden encontrar las futuras “actuaciones asistidas” en la Administración de Justicia contenidas en el APLED. La actuación asistida ha de ser entendida como “aquella para la que el sistema de información de la Administración de Justicia genera un borrador total o parcial de documento complejo en base a datos, que puede ser producido por algoritmos, y puede constituir fundamento o apoyo de una resolución judicial o procesal” (57.1 APLED). Los “documentos complejos” de los que habla el APLED suponen una “caja de sastre” donde podrían caber dictámenes del estilo de los realizados por Prometea para la Fiscalía del Tribunal Superior de Buenos Aires como sentencias judiciales finales.

Uno de los problemas de la aplicación en España de sistemas que generen automáticamente borradores de sentencias, tal como expone el Consejo General del Poder Judicial en el Informe al APLED, sería su incidencia en la independencia judicial⁵⁰. La proposición de borradores podría desdibujar la cada vez más fina línea que separa a la jurisprudencia de su consideración de fuente del Derecho, distanciándonos de nuestro modelo de precedente débil, en que se asientan unos límites de la interpretación legal, pero se da sentido a esta (Mercader Uguina, 2001). Sobre esta cuestión se ha pronunciado repetidamente el Tribunal Constitucional, afirmando que existe diferencia entre el principio de igualdad en la aplicación judicial de la ley, de carácter formal, y el principio de igualdad ante la ley, de carácter material (STC 58/1986, de 14 de mayo, FJ2), pues, “lo que el principio de igualdad en la aplicación de la ley exige no es tanto que la ley reciba siempre la misma interpretación a efectos de que los sujetos a los que se aplique resulten siempre idénticamente afectados, sino que no se emitan pronunciamientos arbitrarios por incurrir en desigualdad no justificada en un cambio de criterio que pueda reconocerse como tal” (STC 49/1985, de 28 de marzo, FJ 2).

Este sistema de precedente débil, así como el *soft law* de la IA, parecen haber sido tomados en cuenta por el legislador español, que se cuida de establecer cualquier vinculatoriedad de los documentos propuestos por el sistema, pues “en ningún caso el borrador documental así generado constituirá por sí una resolución judicial o procesal, sin validación de la autoridad competente” (art. 57.2 APLED). Aun tomándose esta precaución, en el informe del Consejo General del Poder Judicial sobre el anteproyecto se critica la posible falta de control sobre el algoritmo y la carencia de supervisión del mismo por el propio órgano de gobierno de los jueces⁵¹.

⁵⁰ Según el CGPJ en su informe sobre el APLED “[l]a atribución de la potestad jurisdiccional al Poder Judicial supone que corresponde a cada Juez o Tribunal verificar la realidad de los hechos que configuran el objeto del proceso, así como subsumir los hechos en las normas, seleccionando e interpretando el Derecho de aplicación al caso, y emitir los oportunos pronunciamientos resolviendo, conforme a la ley, la controversia existente entre las partes, cuidándose, en su caso, de la ejecución del fallo. Todas estas operaciones deben realizarlas los jueces y magistrados integrantes del Poder Judicial, desplegando las correspondientes capacidades intelectivas y volitivas que conforman el núcleo de la función de juzgar (párr. 163, CGPJ, 2022).

⁵¹ Así, en el Informe al Anteproyecto de Ley, el CGPJ dice que: “No se identifica el responsable de la autorización del uso de los sistemas de actuaciones asistidas, ni los requisitos de gestión de riesgos que deben cumplir. Tampoco se define el sistema de supervisión y control y auditoría al que específicamente deben someterse estos sistemas. Y, por último, pero no menos relevante, nada se dice acerca de la posición que corresponde al CGPJ en la evaluación, supervisión y control de estos tipos de sistemas que impactan directamente sobre el núcleo de la función jurisdiccional” (par.167, 2022).

Con todo lo anterior, se ha adoptado en la jurisdicción contencioso-administrativa una línea favorable a la extensión de efectos en sentencias firmes (art. 110 LRJCA), que expanden sus efectos a otros supuestos idénticos; y a los denominados “casos testigo” (art. 27.2 LRJCA), en los que se vinculan varios casos a la resolución de uno. Esta medida, lógica en el Derecho público contencioso, donde existen múltiples actores frente a una única norma o Administración, pretende ser introducida en otras jurisdicciones, como la civil, estando así contemplado en el Plan de choque del Consejo General del Poder Judicial para la reactivación tras el estado de alarma. Hay que tomar en cuenta que en el ámbito privado no existen ni dos conflictos iguales, ni dos contratos iguales, pero parece seguirse la senda de resolver las disputas de forma general y no singular, sobre todo en contratos de adhesión (Perea González, 2020). En este nuevo escenario la IA podría ser de relevancia para agrupar casos con idénticas características e incluso proponer borradores de resoluciones.

Podemos prever que dentro de los sistemas de información y comunicación involucrados en la APLED se encontrarían sistemas semejantes al ya mentado software Prometea, que funciona examinando los documentos de instancias anteriores, y así “elabora la predicción y crea el documento vinculado a esa predicción. Esto, en definitiva, se vincula con asociar un nuevo caso a una respuesta judicial que se ha emitido previamente en casos de estas características” (Corvalán, 2019b, p. 51). Los borradores creados por este tipo de aplicaciones podrían servir de apoyo a la tarea de los jueces, magistrados, fiscales y letrados de la Administración de Justicia, pero para que revistan utilidad y se adecuen a lo que estos buscan, se han de nutrir de una adecuada base de datos. Para lograrlo, se ha optado por “la orientación al dato” de la Administración de Justicia mediante la “digitalización de la iniciación y tramitación del procedimiento, y el registro de las actividades, asegurando la entrada y tratamiento de información en forma de metadatos, conforme a esquemas y datos comunes e interoperables, que permita la gestión de estos sistemas para el cumplimiento de un compendio de finalidades” (arts. 35 y ss. APLED).

El futuro dirá si el juez adopta por complacencia una propuesta de sentencia realizada por una máquina. Este es el problema del llamado “juez perezoso” o “*juge paresseux*” (Licoppe & Dumoulin, 2019), vinculado a la IA gracias a su propia pasividad, operando una especie de “auto-vinculación” o “auto-precedente” (Abellán, 2011). Para separarse del borrador, el ideal de juez perezoso deberá justificar con especial atención por qué la máquina está errada, lo que, sin duda, le generaría más pereza. Aún de poderse dar estos

casos, el auto-precedente versa, según la doctrina constitucional, siempre sobre la interpretación de la norma, y no sobre los hechos, quedando libre el juez en lo correspondiente a la valoración de la prueba (STC 13/1987, de 5 de febrero, FJ. 2; 134/1991, de 17 de junio FJ. 4). Por tanto, la sentencia propuesta dependería en gran medida de aquellos hechos que el juez introduzca en la máquina, no siendo garantía de que dos casos semejantes sean tratados con idéntica respuesta.

CONCLUSIONES

No existe una definición de inteligencia artificial unánimemente aceptada, pues los propios especialistas disienten en aspectos básicos, elementales para una definición pacífica. En cualquier caso, por ahora, únicamente podemos disponer de una IA del tipo “débil”. “Débil” porque está orientada a la realización de tareas específicas y porque no es capaz de un pensamiento o comportamiento humano. Aunque realmente realice labores complejas, carece de consciencia.

Ante la indefinición del término, se requieren aproximaciones. En el trabajo se realiza, en primer lugar, una aproximación histórica, desde los inicios de la IA a mediados del siglo XX hasta la IA actual, basada en el aprendizaje automático y el análisis de macrodatos, que escapa del cajón logicista tradicional para configurarse como un híbrido entre estadística y computación, usando datos para generar predicciones. A continuación se aborda una aproximación legal, en la que se pasa revista a la regulación nacional y supranacional, en el ámbito regional europeo. En relación a la regulación de la IA, en Europa estamos asistiendo actualmente a un proceso de codificación legal que aún no ha terminado, predominando por ahora el *soft law*. Sin embargo, en este apartado se centra también en la Propuesta de Reglamento Europeo Sobre IA, la futura *hard law* que regulara el tema, que adopta un enfoque de riesgo, habiendo servido de eje troncal para el análisis de su aplicación en el Derecho. La Propuesta, que distingue entre riesgo bajo, alto e IA prohibida, considera de riesgo alto los sistemas de IA “cuyo objetivo es ayudar a las autoridades judiciales a investigar e interpretar los hechos y el Derecho y a aplicar la ley a unos hechos concretos”, dejando de lado funciones accesorias, la denominada tramitación.

La IA como “inteligencia predictiva” y su aplicación en el campo del Derecho ha dado lugar a la denominada “justicia predictiva”, que pretende simular el *output* de los tribunales mediante algoritmos, con una variable dada, relativa al caso, y una por hallar, la decisión judicial. Aunque la IA haya alcanzado una certeza nada desdeñable, queda aún mucho para descubrir los factores que influyen al juez en su decisión, pudiendo surgir en el camino problemas en relación con la intimidad personal de los jueces y la independencia judicial, datos en donde quizá se encuentre la respuesta a la pregunta de qué harán los Tribunales.

La IA es usada cada vez en más ocasiones por los Estados y surge el riesgo de que se apliquen en el plano procesal, es decir, en la aplicación del Derecho a un caso concreto, sin tener en cuenta sus limitaciones. En este trabajo se han propuesto cuatro tipos de límites, aunque estos no tratan de constituir un *numerus clausus*, sino una clasificación orientativa. Estos límites, que conforman una visión realista de las capacidades de la máquina son límites técnicos, jurídicos, técnico-jurídicos y de legitimidad. En relación a los primeros, se expone cómo la IA carece de competencias técnicas en ciertos tipos de conocimientos, si no hay datos o si se ha de interpretar la estructura tras estos. Tampoco jurídicamente puede invadir las competencias del juez para juzgar, o estaría conculcando los arts. 24 y 117 de la Constitución Española. En cuanto a los límites técnico-jurídicos se trata fundamentalmente la cuestión de la transparencia, por la posible amenaza que suponen algoritmos en áreas delicadas, cuyas decisiones no son trazables, y por tanto no se pueden analizar. Por último, se propone un límite relativo a la aceptación ciudadana de las decisiones tomadas por un algoritmo, la legitimidad, algo fundamental para la toma de cualquier decisión en relación a la inclusión de la IA en la Administración de Justicia.

Conocidos los límites, existen no obstante aún amplias posibilidades de uso de las herramientas de IA. Para su examen, se han distinguido sus posibles utilidades en función del criterio de su potencial riesgo, analizándose por un lado las tareas que pueden desarrollar en el ámbito de la tramitación (riesgo bajo o inexistente) y, por otro, en los ámbitos de la investigación procesal y de la toma de la decisión judicial (riesgo alto). Las oportunidades de una mejor gestión en el *back office*, mediante un cambio en el modo de trabajar; y en el *front office*, mejorando la relación con el entorno, se encuentran sobre todo en el ámbito de la tramitación, donde aún queda mucho por hacer. En el ámbito de la investigación procesal y de la toma de decisiones existen ya algunas experiencias provenientes de países avezados en el desarrollo de la IA y el Derecho, como es el caso de EE.UU., tales como la codificación predictiva, claramente positiva, o las herramientas predictivas del riesgo, que no lo son tanto. En cuanto a estas últimas, este trabajo toma en consideración que, de aplicarse alguna herramienta de evaluación del riesgo como las usadas en EE.UU. en el campo del Derecho Penal, estas debieran cumplir los estándares de operar a favor del reo, conociendo el juez de su criterio únicamente cuando la predicción indique un riesgo menor de reincidencia que el obtenido de forma subjetiva, y no al revés, evitando en su caso penas más altas si se tiene un riesgo más bajo de reincidir que el percibido, pero no influenciando al juez de antemano.

Parece que nuestro país va a optar por la inclusión también de la IA en el ámbito de la toma de decisiones, con las nuevas “actuaciones asistidas” recogidas en el Anteproyecto de Ley de Medidas de Eficiencia Digital del Servicio Público de Justicia (APLED), actuaciones asistidas que podrían incluir la generación automática de propuestas o borradores de sentencias. Aunque el legislador deja claro que no pretende invadir la esfera jurisdiccional, su inclusión en nuestro sistema legal puede causar problemas, sobre todo en lo relativo al sometimiento del juez a la ley. Se corre el riesgo de que los jueces, debido al sesgo de automatización, acepten acríticamente los borradores propuestos, de manera que los criterios sobre cómo interpretar la ley no provendrían del juez humano sino de la máquina. En cambio, nada le impediría al juez variar los datos fácticos introducidos en casos semejantes, alegando diferencias en los hechos probados, dando la máquina resultados dispares ante situaciones de hecho iguales, por lo que tampoco una máquina acabaría con los resultados desiguales ante mismos supuestos.

En definitiva, la IA sirve de altavoz a la Administración de Justicia, pudiéndose oír en sitios donde antes era imposible llegar a simple voz por su falta de eficiencia, pero a la vez, en ocasiones, la IA puede emitir de manera eficiente los ecos de la injusticia con la misma intensidad, no porque sea intrínsecamente racista, sexista o machista, sino porque se limita únicamente a ampliar y trasladar injusticias ya existentes en nuestra sociedad.

BIBLIOGRAFÍA

- Acosta, M. (2019). Inteligencia artificial: La cibernética del ser vivo y de la máquina. *Naturaleza y Libertad. Revista de Estudios Interdisciplinarios*, 12.
- Agrawal, A., Gans, J., & Goldfarb, A. (2019). *Máquinas predictivas*. Barcelona: Editorial Reverté.
- Aletras, N., Tsarapatsanis, D., Preoțiuc-Pietro, D. & Lampos, V. (2016). "Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective", *PeerJ Computer Science*, vol. 2, pp. e93.
- Almonacid Sierra, J. J. & Coronel Ávila, Y. (2020). Aplicabilidad de la inteligencia artificial y la tecnología blockchain en el derecho contractual privado. *Revista De Derecho Privado*, (38), 119-142.
- Amador Hidalgo, L. (1996). *Inteligencia artificial y sistemas expertos*, Córdoba: Servicio de Publicaciones de la Universidad, D.L. 1996.
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. *Ethics of data and analytics*, Auerbach Publications, pp. 254-264.
- Arnanz, A. S. (2021). Decisiones automatizadas: Problemas y soluciones jurídicas. más allá de la protección de datos. *Revista de Derecho Público: Teoría y Método*, 3, pp. 85-127.
- Arruñada, B. (2018). "Limitaciones de blockchain en contratos y propiedad". *Minnesota Journal of Law, Science & Technology*, vol. 19, pp. 55-105.
- Ashley, K. D. (2019). A brief history of the changing roles of case prediction in AI and law. *Law Context: A Socio-Legal J.*, 36, 93.
- Ashley, K. D. (2017). *Artificial intelligence and legal analytics: New tools for law practice in the digital age*. Cambridge: Cambridge University Press.
- Asís Pulido, M. de (2022). La justicia predictiva: Tres posibles usos en la práctica jurídica. En Llano Alonso, F. H. (Ed.), *Inteligencia artificial y filosofía del derecho*. Murcia: Laborum, Murcia, pp. 285-312.
- Asociación Pro Derechos Humanos de España. (2015). *La práctica de la prisión provisional en España*. Madrid.
- Baena Pedrosa, M. (2020). *Empleo de inteligencia artificial en la resolución de contenciosos internacionales, públicos y privados. experiencias pioneras* (CUADERNOS DE LA ESCUELA DIPLOMÁTICA NÚMERO 68 ed.) Selección de Memorias del curso selectivo de funcionarios de la Carrera Diplomática 2019-2020.

- Barbaro, C., & Meneceur, Y. (2018). Issues in the use of artificial intelligence (AI) algorithms in judicial systems. *Justice of the Future: Predictive Justice and Artificial Intelligence*, 16°, 2.
- Battelli, E. (2021). La decisión robótica: Algoritmos, interpretación y justicia predictiva. *Revista de Derecho Privado*, (40) Retrieved from <https://dialnet.unirioja.es/descarga/articulo/7865450.pdf>
- Belloso Martín, N. (2022). La problemática de los sesgos algorítmicos (con especial referencia a los de género): ¿Hacia un derecho a la protección contra los sesgos? En Llano Alonso, F. H. (Ed.), *Inteligencia artificial y filosofía del derecho*. Murcia: Laborum, Murcia, pp. 45-78.
- Berman, J. J. (2013). *Principles of big data: Preparing, sharing, and analyzing complex information* Newnes.
- Borden, B. B., & Baron, J. R. (2014). Finding the signal in the noise: Information governance, analytics, and the future of legal practice. *Richmond Journal of Law & Technology*, 20(2), 7.
- Borges, J.L. (1974). El Hacedor" in *Obras completas* Emecé, Buenos Aires, pp. 847.
- Bostrom, N. (2016). *Superinteligencia: Peligros, caminos, estrategias* (M. Alonso Trans.). Teel Editorial.
- Calamandrei, P. (1943). *Instituciones del derecho procesal civil*. Buenos Aires: De Palma Editores.
- Carabantes López, M. (2013). *Inteligencia artificial: condiciones de posibilidad técnicas y sociales para la creación de máquinas pensantes*.
- Carneiro, D., Novais, P., Andrade, F., Zeleznikow, J., & Neves, J. (2014). Online dispute resolution: An artificial intelligence perspective. *Artificial Intelligence Review*, 41(2), pp. 211-240.
- Carrió, G. R. (1981). *Dworkin y el positivismo jurídico* (1st ed.) México, D.F.: Universidad Nacional Autónoma de México, 1981.
- Casanovas, P. (2015). Derecho, tecnología, inteligencia artificial y web semántica: Un mundo para todos y para cada uno. *Enciclopedia de filosofía y teoría del derecho*. Universidad Nacional Autónoma de México (UNAM), Instituto de Investigaciones Jurídicas. Retrieved from <https://dialnet.unirioja.es/servlet/articulo?codigo=6346016&orden=0&info=link>
- Castellanos Claramunt, J., & Montero Caro, M. D. (2020). *Perspectiva constitucional de las garantías de aplicación de la inteligencia artificial: La ineludible protección de los derechos fundamentales*.
- Caterini, M. (2022). El sistema penal en la encrucijada ante el reto de la inteligencia artificial. *IDP: Revista de Internet, Derecho y Política = Revista D'Internet, Dret i*

Política, (35) Retrieved from
<https://dialnet.unirioja.es/descarga/articulo/8399208.pdf>

CGPJ. (2021). *Los españoles y la justicia*. Consejo General del Poder Judicial Metroscopia.

CGPJ (2022). *Informe al Anteproyecto de Ley de Eficiencia Digital Del Servicio Público De Justicia, por la que se transpone al ordenamiento jurídico español la directiva (UE) 2019/1151 del parlamento europeo y del consejo, de 20 de junio de 2019 por la que se modifica la directiva (UE) 2017/1132 en lo que respecta a la utilización de herramientas y procesos digitales en el ámbito del derecho de sociedades*.

Cheang Wong, J. C. (2005). Ley de Moore, nanotecnología y nanociencias: Síntesis y modificación de nanopartículas mediante la implantación de iones. *Revista Digital Universitaria*, 6(7).

Coca Vila, I. (2017). Coches autopilotados en situaciones de necesidad. una aproximación desde la teoría de la justificación penal. *Cuadernos de Política Criminal*, (122), pp. 235-276. Retrieved from
<https://dialnet.unirioja.es/servlet/articulo?codigo=6264427&orden=0&info=link>

Contini, F. (2019). *La inteligencia artificial: ¿un nuevo caballo de troya que influye indebidamente en el poder judicial?* Retrieved from
https://www.unodc.org/dohadeclaration/es/news/2019/06/artificial-intelligence_-_a-new-trojan-horse-for-undue-influence-on-judiciaries.html

Cormack, G. V., & Grossman, M. R. (2014). Evaluation of machine-learning protocols for technology-assisted review in electronic discovery. Paper presented at the *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*, Gold Coast, Queensland, Australia. 153–162. doi:10.1145/2600428.2609601 Retrieved from
<https://doi.org/10.1145/2600428.2609601>

Corvalán, J. G. (2019a). Inteligencia artificial y proceso judicial. desafíos concretos de aplicación. *Diario DPI Novedades-09.09*,

Corvalán, J. G. (2019b). *Prometea: Inteligencia artificial para transformar organizaciones públicas*.

Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences of the United States of America JID - 7505876*, 108(17), 6889-6892.

Del Moral García, A. (2022). Robotización e inteligencia artificial en la justicia. Retrieved from <https://www.youtube.com/watch?v=0S8kfKm8GZI>

Desai, D. R., & Kroll, J. A. (2017). Trust but verify: A guide to algorithms and the law. *Harv.JL & Tech.*, 31, 1.

- Dieterich, W., Mendoza, C., & Brennan, T. (2016). COMPAS risk scales: Demonstrating accuracy equity and predictive parity. *Northpointe Inc*, 7(4)
- Doig Díaz, Y. (2019). Concepto y fuentes del derecho procesal. En Asencio Mellado, J. M. (Ed.), *Introducción al derecho procesal*. Valencia: Tirant lo Blanch, pp. 4-14.
- Dolz Lago, M. J. (2022). Una aproximación jurídica a la inteligencia artificial. *Diario La Ley*, (10096)
- Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances*, 4(1), eaao5580.
- Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason*. MIT Press.
- Dreyfus, H. L. (1965). *Alchemy and Artificial Intelligence*, RAND Corporation, Santa Monica, CA.
- El Espectador. (2018). Así funciona “Watson”, la inteligencia artificial de la fiscalía. Retrieved from <https://www.elespectador.com/judicial/asi-funciona-watson-la-inteligencia-artificial-de-la-fiscalia-article-809463/>
- Estevez, E. C., Linares, S., & Fillotrani, P. (2020). *PROMETEA: Transformando la administración de justicia con herramientas de inteligencia artificial*.
- European Commission. (2020). *Study on the use of innovative technologies in the justice field*.
- European Commission for the Efficiency of Justice. (2016). *Guidelines on how to drive change towards cyberjustice*
- Feynman, R. P. (2008). The computing machines in the future. *Lect.Notes Phys.*, 746, pp. 99-114. doi:10.1007/978-4-431-77056-5_6"
- Flórez Rojas, M. L., & Vargas Leal, J. (2020). *El impacto de herramientas de inteligencia artificial: Un análisis en el sector público en Colombia* (C. Aguerre ed.). Buenos Aires: CETyS Universidad de San Andrés: Inteligencia Artificial en América Latina y el Caribe. Ética, Gobernanza y Políticas.
- Fouquet, D. B. (2021). Sobre inteligencia artificial, decisiones judiciales y vacíos de argumentación. *Teoría & Derecho. Revista De Pensamiento Jurídico*, (29), pp. 248-277.
- Gascón Abellán, M. (2011). Racionalidad y (auto) precedente: Breves consideraciones sobre el fundamento e implicaciones de la regla del autopercedente. *Teoría & Derecho.Revista De Pensamiento Jurídico*, (10), 133-148.
- Goffman, E. (2001). *Internados, ensayos sobre la situación social de los enfermos mentales*. Buenos Aires: Color Efe.

- González, F. A. (2015). Modelos de aprendizaje computacional en reumatología. *Revista Colombiana de Reumatología*, 22(2), 77-78. doi:10.1016/j.rcreu.2015.06.001
- Grimm, D. (2022). Against regulatory disruption. *Jurimetrics J.*, 62, 324.
- Grove, W. M., S,Zald DH FAU - Lebow, B., E,Lebow BS FAU - Snitz, B., C,Snitz BE FAU - Nelson, Nelson, C., & Assess, P. (2000). *Clinical versus mechanical prediction: A meta-analysis*. Psychol Assess.
- Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), pp. 5-14. doi:10.1177/0008125619864925
- Harbers, M., Peeters, M.M. & Neerinx, M.A. (2017). Perceived autonomy of robots: effects of appearance and context. In *A World with Robots*. Springer, pp. 19-33.
- Hart, H. L. A. (1958). Positivism and the separation of law and morals. *Harvard Law Review*, 71(4), 593-629. doi:10.2307/1338225
- Hawkins, J., & Blakeslee, S. (2005). *Sobre la inteligencia*. Espasa Calpe.
- Hazard Jr, G. C. (1997). Discovery and the role of the judge in civil law jurisdictions. *Notre Dame L.Rev.*, 73, 1017.
- Hierro Hernández-Mora, C. (2020). La protección de los derechos fundamentales en el siglo XXI: El abogado y la inteligencia artificial. *Memoria Premios a La Excelencia Cátedra Fundación Mutualidad Abogacía. Un Recorrido Por Los Grandes Retos De La Abogacía Entre 2012 y 2020*, 699.
- HLS. (2015). Harvard Law School launches 'Caselaw access' project. Retrieved from <https://hls.harvard.edu/today/harvard-law-school-launches-caselaw-access-project-ravel-law/>
- Holmes, O. W. (1975). *La senda del derecho*. Buenos Aires: Abeledo Perrot.
- Howard, G. (1983). *Frames of mind: the theory of multiple intelligences*. Basic Books, New York.
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685-695. doi:10.1007/s12525-021-00475-2
- Johnson, K. B., Wei, W., Weeraratne, D., Frisse, M. E., Misulis, K., Rhee, K., Snowdon, J. L. (2021). Precision medicine, AI, and the future of personalized health care. *Clinical and Translational Science*, 14(1), 86-93.
- Katz, D. M., Bommarito II, M. J., & Blackman, J. (2017). Crowdsourcing accurately and robustly predicts supreme court decisions. *arXiv Preprint arXiv:1712.03846*,

- Katz, D.M., Bommarito, M.J. & Blackman, J. (2017) A general approach for predicting the behavior of the Supreme Court of the United States, *PloS one*, vol. 12, no. 4, pp. e0174698.
- Kline, R. (2011). *Cybernetics, automata studies, and the Dartmouth conference on artificial intelligence* doi:10.1109/MAHC.2010.44
- Laporta San Miguel, F. J. (2002). La creación judicial y el concepto de derecho implícito. *Revista Jurídica Universidad Autónoma de Madrid*, (6), pp. 133-151. Retrieved from <https://dialnet.unirioja.es/servlet/articulo?codigo=2867858&orden=192579&info=link>
- Lee, C., & Lin, W. (2020). Induction motor fault classification based on FCBF-PSO feature selection method. *Applied Sciences*, 10(15) doi:10.3390/app10155383
- Lewis, D. (2017). AL interview: Ravel and the AI revolution in legal research. Retrieved from <https://es.coursera.org/lecture/legal-tech-startups/interview-daniel-lewis-ravel-law-ceo-co-founder-xnb5W>
- Liberatore, F., Quijano-Sánchez, L. & Camacho-Collados, M. (2019). Applications of data science in policing. *European Law Enforcement Research Bulletin*, nº 4 SCE, pp. 89-96.
- Licoppe, C., & Dumoulin, L. (2019). Le travail des juges et les algorithmes de traitement de la jurisprudence. premières analyses d'une expérimentation de «justice prédictive» en france. *Droit Et Société*, (3), pp. 535-554.
- Llop Cuenca, M. P. (2021). Ley de eficiencia digital. *Diario La Ley* (9957)
- Lock, S. (2020). Legal tech patent filings jumped to all-time high in 2019. Retrieved from <https://www.law.com/international-edition/2020/10/29/lawtech-patent-filings-jumped-to-all-time-high-in-2019/>
- Loevinger, L. (2004). Jurimetrics the next step forward. *Jurimetrics*, vol. 44, nº 4, pp. 405- 408.
- Lorenzo Pérez, C. (2022). *Inteligencia artificial en la administración de justicia: Regulación española y marco europeo e internacional. proyectos desarrollados por el Ministerio de Justicia de España* (Tramitación orientada al dato de procedimientos judiciales y estadística judicial ed.) Centro de Estudios Jurídicos.
- Maramot, S. (2019), Keeping Up with Legal Technology. *ITA In Review*, vol. 1º, no. 2, pp. 37.
- Martin, A. D., Quinn, K. M., Ruger, T. W. & Kim, P. T. (2004). Competing approaches to predicting supreme court decision making. *Perspectives on Politics*, vol. 2, nº. 4, pp. 761-767.
- Martínez Garay, L. (2018). Peligrosidad, algoritmos y due process: El caso State v Loomis. *Revista de Derecho Penal y Criminología (Uned)*, nº 20, pp. 485-502.

- Martínez Garay, L., & Montes Suay, F. (2018). El uso de valoraciones del riesgo de violencia en derecho penal: Algunas cautelas necesarias. *InDret*.
- Martínez Martínez, R. (2019). Intelligència artificial des del disseny. reptes i estratègies per al compliment normatiu (ES-EN). *Revista Catalana De Dret Públic*, pp. 64-81.
- McCorduck, P. (2004). *Machines who think*. A. K. Peters.
- Medina Uribe, P., & Fernanda Gómez, L. (2020). 'Watson', el investigador inteligente con el que la fiscalía busca cerrarle el paso al crimen. Retrieved from <https://www.elpais.com.co/judicial/watson-el-investigador-inteligente-con-el-que-la-fiscalia-busca-cerrarle-el-paso-al-crimen.html>
- Medvedeva, M., Vols, M., & Wieling, M. (2020). Using machine learning to predict decisions of the european court of human rights. *Artificial Intelligence and Law*, 28(2), pp. 237-266.
- Mercader Uguina, J. R. (2001). El imposible derecho a la igualdad en la aplicación judicial de la ley en la doctrina constitucional. *El proceso laboral: Estudios en homenaje al profesor Luis Enrique de la Villa Gil*, Lex Nova, pp. 587-636
- Meroi, A. (2022). Normas procesales e inteligencia artificial: Condiciones para una intersección posibilista. En D. Guerra Moreno (Ed.), *Constitución e inteligencia artificial en el proceso*. Bogotá, D.C.: Universidad Libre Grupo Editorial Ibáñez.
- Meseguer González, P., & López de Mántaras, R. (2017). *Inteligencia artificial*. La Catarata.
- Miró Llinares, F. (2018). Inteligencia artificial y justicia penal: Más allá de los resultados lesivos causados por robots. *Revista de Derecho Penal y Criminología*, (20), pp. 87-130. Retrieved from <https://dialnet.unirioja.es/servlet/articulo?codigo=7228581&orden=0&info=link>
- Mitchell, T. M., & Jordan, M. I. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), pp. 255-260. doi:10.1126/science.aaa8415
- Molina García, M. J. (2017). Jurimetría: Tecnología punta y su aplicación práctica en el sector legal. *Diario La Ley* (9029)
- Montoro Ballesteros, M. A. (2007). El funcionalismo en el derecho: Notas sobre N. Luhmann y G. Jakobs. *Anuario de Derechos Humanos*, (8), pp. 365-374. Retrieved from <https://dialnet.unirioja.es/servlet/articulo?codigo=2748189&orden=0&info=link>
- Morales Moreno, Á M. (2021). Algoritmos en el estrado, ¿realmente los aceptamos? Percepciones del uso de la inteligencia artificial en la toma de decisiones jurídico-penales. *Ius Et Scientia*, 7(2), pp. 57-87.
- Navarro, J. B. (2020). Algunas consideraciones acerca del poder configurador de la inteligencia artificial sobre el proceso. En Ramírez Carvajal, D. y Vásquez

- Santamaría, J. (Eds.), *Debates contemporáneos del proceso en un mundo que se transforma*. Medellín, Colombia: Universidad Católica Luis Amigó, pp. 95-119.
- Neisser, U. (1963). The imitation of man by machine. *Science*, 139(3551), 193-197. Retrieved from <http://www.jstor.org/stable/1710006>
- Nicholas Taleb, N. (2008). *El cisne negro: el impacto de lo altamente improbable*. Barcelona: Paidós.
- Nieva Fenoll, J. (2018). *Inteligencia artificial y proceso judicial*. Madrid: Marcial Pons.
- Oard, D. W., & Webber, W. (2013). Information retrieval for e-discovery. *Foundations and Trends in Information Retrieval*, 7(2-3), pp. 99-237.
- Oliveros Roselló, M. J. (2019). La jurimetría: ¿Nos sustituirán por Ross? Retrieved from <https://apm3-9.studi-web.com/jurimetria.pdf>
- O'Neil, C. (2018). *Armas de destrucción matemática: Cómo el big data aumenta la desigualdad y amenaza la democracia*. Madrid: Capitan Swing.
- Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, 52(3), pp. 381-410.
- Pasquale, F., & Cashwell, G. (2018). Prediction, persuasion, and the jurisprudence of behaviourism. *University of Toronto Law Journal*, 68(supplement 1), pp. 63-81.
- Peces-Barba Martínez, G., Asís Roig, R. F. d., & Fernández García, E. (1999). *Curso de Teoría del Derecho*. Madrid: Marcial Pons.
- Penrose, R. (1996). *La mente nueva del emperador: En torno a la cibernética, la mente y las leyes de la física*. México: FCE.
- Perea González, Á. (2021). Justicia predictiva: Una solución al 'pleito masa'. Retrieved from https://cincodias.elpais.com/cincodias/2021/03/16/legal/1615930000_454211.html
- Perea González, Á. (2020). Hacer generalidad de la singularidad: Pleito testigo y extensión de efectos ¿Una nueva tutela del conflicto privado? *Diario La Ley* (9669)
- Pérez Bernabeu, B. (2021). El principio de explicabilidad algorítmica en la normativa tributaria española: Hacia un derecho a la explicación individual. *Civitas. Revista Española de Derecho Financiero*, (192), pp. 143-178.
- Pérez Estrada, M. J. (2022). *Fundamentos jurídicos para el uso de la inteligencia artificial en los órganos judiciales*. Valencia: Tirant lo Blanch.
- Pérez Jaraba, M. D. (2010). Principios y reglas: Examen del debate entre R. Dworkin y H.L.A. Hart. *Revista de Estudios Jurídicos*, (10), pp. 247-276. Retrieved from <https://dialnet.unirioja.es/servlet/articulo?codigo=3871466&orden=0&info=link>

- Pérez Luño, A. E. (2011). Las fuentes del derecho y su problemática actual. En *Perspectivas actuales de las fuentes del derecho*. Madrid: Dykinson, pp. 39-68.
- Pérez Ragone, Á. (2021). La justicia civil en la era digital y artificial: ¿Hacia una nueva identidad? *Revista Chilena de Derecho*, 48(2), pp. 203-229. Retrieved from <https://dialnet.unirioja.es/descarga/articulo/8119791.pdf>
- Poulson, B. (2019). Fundamentos de big data: Técnicas y conceptos. Retrieved from <https://es.linkedin.com/learning/fundamentos-de-big-data-tecnicas-y-conceptos>
- Pritchett, C.H. (2014). *The Roosevelt Court: A study in judicial politics and values, 1937-1947*. Quid Pro Books.
- Pruss, D. (2021). Mechanical jurisprudence and domain distortion: How predictive algorithms warp the law. *Philosophy of Science*, 88(5), pp. 1101-1112.
- Radin, M. (1931). Scientific method and the law. *California Law Review*, 19(2), pp. 164-172. doi:10.2307/3475854
- RAE (2022). Diccionario de la lengua española, 23.^a ed., [versión 23.6 en línea] [Homepage of Real Academia Española], [Online]. Available: <https://dle.rae.es> [2022, 10/10].
- Ragnedda, M., Ruiu, M. L., & Addeo, F. (2020). Measuring digital capital: An empirical investigation. *New Media & Society*, 22(5), pp. 793-816.
- Reichenbach, H. (1938). Experience and prediction: An analysis of the foundations and the structure of knowledge.
- Remus, D., & Levy, F. (2017). Can robots be lawyers: Computers, lawyers, and the practice of law. *Geo.J.Legal Ethics*, 30, 501.
- Richardson, R., Schultz, J. M., & Crawford, K. (2019). Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *NYUL Rev.Online*, 94, 15.
- Rodríguez Puerto, M. J. (2021). ¿Puede la inteligencia artificial interpretar normas jurídicas? un problema de razón práctica.
- Rodríguez, C. (1997). *La decisión judicial. El Debate Hart-Dworkin*. Colombia, Siglo Del Hombre Editores, Facultad de Derecho-Universidad de Los Andes, pp. 49-50.
- Rodríguez, T. (2022). Así funciona el algoritmo MAX de la ITSS para controlar las horas extras. Retrieved from <https://noticiastrabajo.huffingtonpost.es/empleo/derechos-trabajador/asi-funciona-algoritmo-inspeccion-de-trabajo-para-controlar-horas-extras/>
- Ruger, T. W., Kim, P. T., Martin, A. D., & Quinn, K. M. (2004). The supreme court forecasting project: Legal and political science approaches to predicting supreme

- court decisionmaking. *Columbia Law Review*, 104(4), pp. 1150-1210. doi:10.2307/4099370
- Ruiz Manotas, P. M., & Bermeo Álvarez, L. F. (2018). La recepción de la teoría de los sistemas de luhmann en la jurisprudencia constitucional colombiana. *Prolegómenos: Derechos y Valores*, 21(42), pp. 161-175. Retrieved from <https://dialnet.unirioja.es/descarga/articulo/6819243.pdf>
- Russell, S. J., & Norvig, P. (2004). *Inteligencia artificial: Un enfoque moderno*. Madrid: Pearson Educación, S.A.
- Saavedra, V., & Upegui, J. C. (2021). *PretorIA y la automatización del procesamiento de causas de derechos humanos* (Inteligencia Artificial e Inclusión ed.) Derechos Digitales.
- Salazar García, I., & Benjamins, R. (2020). *El mito del algoritmo: Cuentos y cuentas de la inteligencia artificial*. Anaya Multimedia.
- Salgado, C. A. (2021). Acerca de la inteligencia artificial en el ámbito penal: Especial referencia a la actividad de las fuerzas y cuerpos de seguridad. *Ius Et Scientia*, 7(1), 25-36.
- Salvador Coderch, P., & Fernández Crende, A. (2006). Causalidad y responsabilidad. *Indret: Revista para el Análisis del Derecho*, (1) Retrieved from <https://dialnet.unirioja.es/servlet/articulo?codigo=1396085&orden=0&info=link>
- Samek, W., & Müller, K. (2019). Towards explainable artificial intelligence. *Explainable {AI}: Interpreting, explaining and visualizing deep learning*. Springer International Publishing, pp. 5-22. doi:10.1007/978-3-030-28954-6_1 Retrieved from https://doi.org/10.1007%2F978-3-030-28954-6_1
- San Miguel Caso, C. (2021). La aplicación de la Inteligencia Artificial en el proceso: ¿un nuevo reto para las garantías procesales? *IUS ET SCIENTIA: Revista electrónica de Derecho y Ciencia*, vol. 7, nº 1, pp. 286-303.
- Sánchez Vilanova, M. (2022). El uso de algoritmos predictivos en el derecho penal. A propósito de la sentencia de la corte de justicia del distrito de La Haya (Países Bajos) sobre syri, de 5 de febrero de 2020. *Teoría y Derecho: Revista de Pensamiento Jurídico*, (33), pp. 252-281.
- Sangüesa i Solé, R. (2018). Intelligència artificial i transparència algorítmica: " It's complicated". *BiD: Textos Universitaris De Biblioteconomia i Documentació*, (41).
- Scherer, M. (2019). Artificial intelligence and legal decision-making: The wide open? *Journal of International Arbitration*, 36(5)
- Schiavenza, M. (2018). China's 'sputnik moment' and the sino-american battle for AI supremacy. Retrieved from <https://asiasociety.org/blog/asia/chinas-sputnik-moment-and-sino-american-battle-ai-supremacy>

- Schussler, A. E. (2020). Artificial intelligence and mind-reading machines - towards a future techno-panoptic singularity. *Postmodern Openings*, 11(4), 334-346. doi:<https://doi.org/10.18662/>
- Schwartz, J. T. (1988). The new connectionism: Developing relationships between neuroscience and artificial intelligence. *Daedalus*, 117(1), pp. 123-141. Retrieved from <http://www.jstor.org/stable/20025141>
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), pp. 417-424. doi:10.1017/S0140525X00005756
- Searle, J. R. (1985). *Mentes, cerebros y ciencia*. Madrid: Cátedra.
- Segura, S., Fernández Berrocal, P., & Byrne, R. M. J. (1998). Razonamiento contrafactual: La posición serial y el número de antecedentes en los pensamientos sobre lo que podría haber sido. *I jornadas de psicología del pensamiento: (Actas) : Santiago de compostela, 22-23 de junio, 1998* (1st ed., pp. 179-187) Universidade de Santiago de Compostela, Servizo de Publicacións. Retrieved from <https://dialnet.unirioja.es/descarga/articulo/2520853.pdf>
- SGNTJ. (2020a). Calculadora 988. Retrieved from <https://aslan.es/wp-content/uploads/2020/02/20200205-Calculadora-988F.pdf>
- SGNTJ. (2020b). Textualización de vistas. Retrieved from <https://aslan.es/wp-content/uploads/2020/02/20200205-Textualizaci00f3n-de-Vistas-F.pdf>
- Shi, C., Sourdin, T. & Li, B. (2021) The Smart Court-A New Pathway to Justice in China?, *IJCAHeinOnline*, pp. 1.
- Sil, R., Roy, A., Bhushan, B., & Mazumdar, A. (2019). Artificial intelligence and machine learning based legal application: The state-of-the-art and future research trends. Paper presented at the *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, pp. 57-62.
- Simón Castellano, P. (2021). Inteligencia artificial y administración de justicia: ¿Quo vadis, justitia? *IDP: Revista de Internet, Derecho y Política = Revista D'Internet, Dret i Política*, (33) Retrieved from <https://dialnet.unirioja.es/descarga/articulo/7962069.pdf>
- Snijders, H. (2021). Arbitration and AI, from arbitration to ‘Robotration’ and from human arbitrator to robot. *Arbitration: The International Journal of Arbitration, Mediation and Dispute Management*, 87(2),
- Solar Cayón, J. I. (2012) Holmes: el inicio de una nueva senda Jurídica pp. 11-53. En *La senda del derecho*. Marcial Pons, pp. 11-53.
- Solar Cayón, J. I. (2018) La codificación predictiva: inteligencia artificial en la averiguación procesal de los hechos relevantes, *Anuario de la Facultad de Derecho*, nº. 11, pp. 75-105

- Solar Cayón, J. I. (2019). *La inteligencia artificial jurídica: El impacto de la innovación tecnológica en la práctica del derecho y el mercado de servicios jurídicos*. Thomson Reuters Aranzadi.
- Solar Cayón, J. I. (2020a). La inteligencia artificial jurídica: Nuevas herramientas y perspectivas metodológicas para el jurista. *Revus. Journal for Constitutional Theory and Philosophy of Law/Revija Za Ustavno Teorijo in Filozofijo Prava*, (41).
- Solar Cayón, J. I. (2020b). Inteligencia artificial en la justicia penal: Los sistemas algorítmicos de evaluación de riesgos. En Solar Cayón, J. I. (Ed.), *Dimensiones éticas y jurídicas de la inteligencia artificial en el marco del estado de derecho*, Universidad de Alcalá - Defensor del Pueblo, pp. 125-172.
- Solar Cayón, J. I. (2021). Reflexiones sobre la aplicación de la inteligencia artificial en la administración de justicia. *Teoría Jurídica Contemporânea*, vol. 6, 2021. Disponible en <https://repositorio.unican.es/xmlui/handle/10902/24149>.
- Solar Cayón, J. I. (2022). Inteligencia artificial y justicia digital. En Llano Alonso, F. H. (Dir.). *Inteligencia artificial y filosofía del derecho*. Murcia: Laborum, Murcia, pp. 381-427.
- Srinivasa, M., Mounika, M., & Fareed, S. (2020). Implementation of service based chatbot using deep learning. *Test Engineering & Management*, (83), pp. 2013- 2019.
- Sternberg, R.J. (1985) *Beyond IQ: A triarchic theory of human intelligence*. Cambridge University Press, New York, NY, US.
- Surdeanu, M., Nallapati, R., Gregory, G., Walker, J., & Manning, C. D. (2011). Risk analysis for intellectual property litigation. Paper presented at the *Proceedings of the 13th International Conference on Artificial Intelligence and Law*, pp. 116-120.
- Susskind, R. (2017). *El abogado del mañana: Una introducción a tu futuro*. Madrid: Wolters Kluwer.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. Cambridge, MA, USA: A Bradford Book.
- Thomson Reuters. (2017). Thomson reuters analysis reveals 484% increase in new legal services patents globally as law firms around the world invest in legal tech. Retrieved from <https://www.thomsonreuters.com/en/press-releases/2017/august/thomson-reuters-analysis-reveals-484-percent-increase-in-new-legal-services-patents-globally.html>
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), pp. 1124-1131. doi:10.1126/science.185.4157.1124
- Vásquez, A. C., Quispe, J. P., & Huayna, A. M. (2009). Procesamiento de lenguaje natural. *Revista de Investigación de Sistemas e Informática*, 6(2), 45-54.

- Vilela Komatsu, A., Wenger, L. S., Costa, R. C., Rezende Bazon, M., & Pueyo, A. A. (2019). Factores protectores en adolescentes infractores: Un estudio tipológico. *International E-Journal of Criminal Sciences*, (14) Retrieved from <https://dialnet.unirioja.es/descarga/articulo/7298460.pdf>
- Viola, L. (2020). *Interpretación de la ley a través de modelos matemáticos*. Milano: Diritto Avanzato.
- Vivar Vera, J. (2021). La sentencia penal, el juez y el algoritmo: ¿las nuevas tecnologías serán nuestros próximos jueces? *Revista Chilena de Derecho y Tecnología*, 10(1), 231-269.
- Volokh, E. (2018). Chief justice robots. *Duke LJ*, 68, 1135.
- Wang, P. (2019). On defining artificial intelligence. *Journal of Artificial General Intelligence*, 10(2), pp. 1-37. doi:10.2478/jagi-2019-0002
- Washington, A. L. (2018). How to argue with an algorithm: Lessons from the COMPAS-ProPublica debate. *Colo.Tech.LJ*, 17, 131.
- Wollowski, M., Selkowitz, R., Brown, L., Goel, A., Luger, G., Marshall, J., Norvig, P. (2016). A survey of current practice and teaching of AI. *Proceedings of the AAAI Conference on Artificial Intelligence*, 30(1) doi:10.1609/aaai.v30i1.9857
- Xu, Z. (2022) Human judges in the era of artificial intelligence: challenges and opportunities, *Applied Artificial Intelligence*, vol. 36, n° 1.
- Yablon, C. & Landsman-Roos, N. (2013) Predictive Coding: Emergin Questions and Concerns, *South Carolina Law Review*, vol. 64, n° 6.