

ALFREDO ABADÍAS SELMA
PERE SIMÓN CASTELLANO

DIRECTORES

MARIO CATERINI
(COORD. ITALIA)

Desafíos jurídicos de la inteligencia artificial

Reflexiones sobre la
toma de decisiones judiciales

PRÓLOGO
MIGUEL ÁNGEL PRESNO LINERA



JMB
BOSCH EDITOR

COLECCIÓN
«Penalcrim» J.M. Bosch Editor

Partiendo de una reconstrucción histórica y filosófica del papel y del poder del juez, esta obra propone llevar a cabo una reflexión sobre las implicaciones penales del uso de la inteligencia artificial y de los algoritmos, en particular los «sentenciosos», identificando criterios constitucionalmente orientados y capaces de «guiar» el proceso de toma de decisiones, que no puede dejarse únicamente en manos de las competencias informáticas.

Por una parte, intenta demostrar la incompatibilidad de un criterio puramente estadístico-actuarial y, por otra, de evaluar y ofrecer un parámetro de decisión diferente que pueda ser utilizado por la inteligencia artificial, basado en el principio de *in dubio interpretatio pro reo*. Se trata de desarrollar un criterio decisorio maquínico que, ante varias opciones hermenéuticas, no se limite a «calcular» y sugerir al juez la estadísticamente «más probable», sino la que, entre las «probables», resulte más favorable al acusado.

Desafíos jurídicos de la inteligencia artificial

Reflexiones sobre la toma
de decisiones judiciales

ALFREDO ABADÍAS SELMA
PERE SIMÓN CASTELLANO

DIRECTORES

MARIO CATERINI (COORD. ITALIA)

Desafíos jurídicos de la inteligencia artificial

Reflexiones sobre la
toma de decisiones judiciales

PRÓLOGO

MIGUEL ÁNGEL PRESNO LINERA

Catedrático de Derecho Constitucional
Universidad de Oviedo

2024



BOSCH EDITOR

Esta obra ha sido examinada por los siguientes miembros del Comité Científico editorial:

Dr. Alfredo Abadías Selma. Profesor Contratado Doctor de Derecho Penal, Universidad Internacional de La Rioja

Dr. Miguel Bustos Rubio. Profesor Titular de Derecho Penal, Universidad Internacional de La Rioja

Dr. Bernardo Feijoo Sánchez. Catedrático de Derecho Penal. Universidad Autónoma de Madrid

© JULIO 2024 ALFREDO ABADÍAS SELMA
PERE SIMÓN CASTELLANO
(Directores)

© JULIO 2024

**JTB BOSCH
EDITOR**

Librería Bosch, S.L.
<http://www.jmboscheditor.com>
<http://www.libreriabosch.com>
E-mail: editorial@jmboscheditor.com

Coordinación de la obra en Italia: MARIO CATERINI

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra solo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Dirijase a CEDRO (Centro Español de Derechos Reprográficos) si necesita fotocopiar o escanear algún fragmento de esta obra (www.conlicencia.com; 91 702 19 70 / 93 272 04 45).

SBN papel: 978-84-10044-78-4
ISBN digital: 978-84-10044-79-1
D.L.: B 10036-2024

Diseño portada y maquetación: CRISTINA PAYÁ  +34 672 661 611

Printed in Spain – Impreso en España

Colección «Penalcrim» J.M. Bosch Editor

Coordinadores del Comité Científico

Dr. Alfredo Abadías Selma

Profesor Contratado Doctor de Derecho Penal y Criminología
Universidad Internacional de La Rioja

Dr. Miguel Bustos Rubio

Profesor Titular de Derecho Penal
Universidad Internacional de La Rioja

Miembros del Comité Científico

Dra. María del Carmen Armendáriz León

Profesora Contratada Doctora de Derecho Penal
Universidad Complutense de Madrid

Dr. Ignacio Berdugo Gómez De La Torre

Catedrático de Derecho Penal
Universidad de Salamanca

Dr. Sergio Cámara Arroyo

Profesor Titular de Derecho penal
Universidad Nacional de Educación
a Distancia UNED

Dra. Beatriz Cruz Márquez

Profesora Titular de Derecho Penal y Criminología
Universidad de Cádiz

Dra. Beatriz Escudero García-Calderón

Profesora de Derecho Penal, CUNEF

Dr. Bernardo Feijoo Sánchez

Catedrático de Derecho Penal
Universidad Autónoma de Madrid

Dr. Juan Carlos Ferré Olivé

Catedrático de Derecho Penal
Universidad de Huelva

Dr. Octavio García Pérez

Catedrático de Derecho Penal
Universidad de Málaga

Dra. Alicia Gil Gil

Catedrática de Derecho Penal
Universidad Nacional de Educación a Distancia

Dra. María Concepción Gorjón Barranco

Profesora Titular de Derecho Penal
Universidad de Salamanca

Dr. José León Alapont

Profesor Titular de Derecho Penal
Universidad de Valencia

Dra. Paz Lloria García

Catedrática de Derecho Penal
Universidad de Valencia

Dr. Víctor Manuel Macías Caro

Profesor Ayudante Doctor de Derecho Penal
Universidad Pablo de Olavide de Sevilla

Dr. Esteban Mestre Delgado

Catedrático de Derecho Penal
Universidad de Alcalá de Henares

Dr. Fernando Miró Linares

Catedrático de Derecho Penal,
Universidad Miguel Hernández de Alicante

Dr. Fernando Navarro Cardoso

Catedrático de Derecho Penal
Universidad de Las Palmas de Gran Canaria

Dr. Félix María Pedreira González

Profesor Titular de Derecho Penal
Universidad Complutense de Madrid

Dra. Ana María Peligero Molina

Profesora Ayudante Doctora de Criminología
Universidad Internacional de La Rioja

Dra. Ana Isabel Pérez Cepeda

Catedrática de Derecho Penal
Universidad de Salamanca

Dr. Francisco Rodríguez Almirón

Profesor Derecho Penal
Universidad de Granada

Dr. Enrique Sanz Delgado

Profesor Titular de Derecho Penal
Universidad de Alcalá de Henares

Dr. Pere Simón Castellano

Profesor Titular de Derecho Constitucional
Universidad Internacional de La Rioja

*El problema de la humanidad
es que los estúpidos están seguros de todo
y los inteligentes están llenos de dudas.*

Bertrand Russell

ÍNDICE

PRÓLOGO. Miguel Ángel Presno Linera	17
INTRODUCCIÓN. Mario Caterini	23
 CAPÍTULO I	
DERECHO, ALGORITMOS Y JUSTICIA PENAL: LA DUDA RAZONABLE EN LA ERA DE LA INTELIGENCIA ARTIFICIAL Mario Caterini	27
1. INTRODUCCIÓN	27
2. LA CERTEZA DEL <i>IUS DICERE</i> “ARTIFICIAL”: ¿HA LLEGADO EL FUTURO?	32
3. LOS ALGORITMOS PREDICTIVOS EN EL ÁMBITO JURÍDICO: MODELOS Y DEFECTOS DE FUNCIONAMIENTO.....	34
4. LA “FACTUALIZACIÓN” CIBERNÉTICA DEL DERECHO.....	41
5. LA IMPLEMENTACIÓN ALGORÍTMICA DE LA REGLA DE LA LENIDAD	43
6. CONCLUSIONES.....	48
7. REFERENCIAS BIBLIOGRÁFICAS.....	53
 CAPÍTULO II	
EL ESPÍRITU DE LA LEY Y EL PARADIGMA DEL PODER LIMITADO: EL REGRESO AL JUEZ AUTÓMATA María Princi	63
1. EL JUEZ INANIMADO EN LA ERA DE LAS LUCES.....	63

2.	LA INTERPRETACIÓN SILOGÍSTICA EN LA VISIÓN DE BECCARIA.....	68
3.	EL ESPÍRITU DE LA LEY Y EL PARADIGMA DEL PODER LIMITADO: LOS ORÍGENES DEL GARANTISMO PENAL.....	72
4.	ANTÍTESIS ENTRE ARBITRARIEDAD Y SEGURIDAD JURÍDICA.....	76
5.	EL ESPECTRO DEL JUEZ AUTÓMATA.....	79
6.	¿ES CONFIABLE LA INTELIGENCIA ARTIFICIAL?	80

CAPÍTULO III

ALGORITMOS: EVOLUCIÓN, INTENTOS DE DEFINICIONES E IMPLICACIONES PENALES Valentina Aragona		89
1.	EL NACIMIENTO Y EVOLUCIÓN DE LOS ALGORITMOS: DE LA INTELIGENCIA ARTIFICIAL CLÁSICA AL <i>MACHINE LEARNING</i>	89
2.	POSIBLES APLICACIONES ALGORÍTMICAS EN EL MUNDO DEL DERECHO PENAL.....	100
3.	VENTAJAS Y CUESTIONES CRÍTICAS DE LA TECNOLOGÍA ALGORÍTMICA EN EL PROCESO PENAL	104
4.	INTENTOS REGULATORIOS ENTRE <i>CIVIL LAW</i> Y <i>COMMON LAW</i>	108
4.1.	El enfoque de <i>soft law</i> estadounidense	109
4.2.	<i>The Humman in command</i> : la orientación europea.....	113

CAPÍTULO IV

EL ALGORITMO DE JUZGAMIENTO: NUEVO FACTOR DE RIESGO PARA LA LEGALIDAD PENAL Mariana Rocca		129
1.	TOMA DE DECISIONES AUTOMATIZADA Y EL IMPACTO CRÍTICO EN LA LEGALIDAD PENAL. PREMISAS GENERALES	129
2.	UN INTENTO DE REVISIÓN DOGMÁTICA DEL PRINCIPIO DE LEGALIDAD	136
3.	LOS COROLARIOS DE LA LEGALIDAD PENAL EN LA APLICACIÓN PRÁCTICA DEL DERECHO: LA INFLUENCIA DE LA <i>LAW IN ACTION</i>	149
4.	LEGALIDAD PENAL EN LA JURISPRUDENCIA DE LA CORTE EDH	165

5.	EL VALOR DEL PRECEDENTE EN LOS SISTEMAS DE <i>COMMON LAW</i>	176
6.	EL VALOR DEL PRECEDENTE EN EL ORDENAMIENTO JURÍDICO ITALIANO, BASADO EN EL PRINCIPIO DE LEGALIDAD	182
7.	EL <i>FAVOR REI</i> COMO BARRERA A LOS ÁRBITROS INTERPRETATIVOS	185
8.	REFLEXIONES FINALES	187

CAPÍTULO V

LA HERMENÉUTICA PENAL EN LA ERA ROBÓTICA: EL HORIZONTE DEL <i>FAVOR REI</i> PARA LA ‘PREVISIBILIDAD’ JUDICIAL Alberto D’Aiello		189
1.	RELEVANCIA DEL <i>ARS INTERPRETANDI</i> EN LOS ‘PLIEGUES’ DE LA JURISDICCIÓN	190
2.	NECESARIA CONFORMIDAD DE LA INTERPRETACIÓN CON LOS PRINCIPIOS CONSTITUCIONALES Y EUROPEOS	199
3.	LA ACTIVIDAD HERMENÉUTICA A LA LUZ DE LA RELACIÓN ENTRE LENGUAJE NORMATIVO, DERECHO PENAL E INTELIGENCIA ARTIFICIAL	209
4.	<i>QUAESTIO FACTI</i> Y <i>QUAESTIO IURIS</i> : UN VÍNCULO (IN)SEPARABLE EN EL JUICIO ROBÓTICO.....	218
5.	EL PARADIGMA HERMENÉUTICO DEL <i>FAVOR REI</i> ENTRE JUSTICIA PREDICTIVA Y SEGURIDAD JURÍDICA.....	228
6.	LA INCIDENCIA DE LA DIRECCIÓN INTERPRETATIVA UNÍVOCA SOBRE LA DURACIÓN RAZONABLE DEL JUICIO	237
7.	<i>MÁQUINA INTERPRETARI POTEST?</i> REFLEXIONES SOBRE LA VIRTUD DE LA ARGUMENTACIÓN JURÍDICA	244

CAPÍTULO VI

OBLIGACIÓN DE MOTIVACIÓN: UN ANÁLISIS COMPARADO POR UNA IA GARANTISTA Gabriele Polito		255
1.	LAS RAÍCES ESCONDIDAS DE LA OBLIGACIÓN DE MOTIVACIÓN: PREMISA SOBRE LA BUENA ARGUMENTACIÓN	255
2.	EL DEBER DE MOTIVAR EN EL SISTEMA PROCESAL PENAL ITALIANO	269
3.	PERFILES COMPARADOS: LA CARENCIA DE MOTIVACIÓN “INSTITUCIONALIZADA” EN LOS SISTEMAS DE <i>COMMON</i>	

	<i>LAW</i> Y LA CONSIGUIENTE NECESIDAD DE LA UNANIMIDAD EN EL VEREDICTO DE CONDENA	281
4.	PERFILES COMPARADOS: EL TRIBUNAL DEL JURADO EN EL SISTEMA ESPAÑOL Y LA MOTIVACIÓN SEMPLIFICADA EN EL VEREDICTO DE ABSOLUCIÓN	291
5.	REFLEXIONES CONCLUSIVAS: ¿CUÁLES SON LOS POSIBLES USOS DE LA INTELIGENCIA ARTIFICIAL A LA HORA DE MOTIVAR LOS ACTOS JURISDICIONALES?.....	302

CAPÍTULO VII

	RAZONAMIENTO PROBABILÍSTICO ARTIFICIAL: HACIA UNA NUEVA CERTEZA DEL SISTEMA PENAL Morena Gallo	309
1.	LOS ORÍGENES DEL PROBABILISMO	309
2.	LA RETÓRICA EN EL <i>ἐνδοξα</i>	313
3.	EL RAZONAMIENTO DIALÉCTICO ARISTOTÉLICO	315
4.	LA LÓGICA DE HERÁCLITO	317
5.	LA LÓGICA DE <i>PORT-ROYAL</i>	317
	5.1. La lógica de Arnould.....	320
6.	EL PROBABILISMO EN TEOLOGÍA MORAL	322
7.	CALCULABILIDAD Y PROBABILIDAD JURÍDICA	325
8.	JUICIO DE HECHO Y JUICIO DE DERECHO	329
9.	LA RAZONABILIDAD DE LA DUDA	333
10.	EL PROBABILISMO COMO TEORÍA ENRAIZADA EN EL PRINCIPIO DEL <i>FAVOR REI</i>	337
11.	CONCLUSIONES SUCINTAS: LA SOLUCIÓN PROBABILÍSTICA EN EL PROCESO ALGORÍTMICO	340

CAPÍTULO VIII

	LA INSUFICIENCIA DEL MARCO NORMATIVO DE LA PROTECCIÓN DE DATOS Y EL ADVENIMIENTO DEL REGLAMENTO EUROPEO SOBRE INTELIGENCIA ARTIFICIAL Pere Simón Castellano	343
--	--	-----

1. A MODO DE INTRODUCCIÓN: RAZONES PARA REGULAR LA INTELIGENCIA ARTIFICIAL MEDIANTE LA TÉCNICA ARMONIZADORA DEL NUEVO ENFOQUE. LA IRRUPCIÓN

	DE LOS SELLOS DE CONFORMIDAD Y LOS SISTEMAS DE GESTIÓN DE RIESGOS	343
2.	SOBRE LA NECESIDAD DE IR MÁS ALLÁ DE LA AUTODETERMINACIÓN INFORMATIVA	345
3.	EL REGLAMENTO EUROPEO DE INTELIGENCIA ARTIFICIAL: CUADRO RESUMEN DEL ENFOQUE, LAS OBLIGACIONES Y EL RÉGIMEN JURÍDICO RESULTANTE	352
4.	CONCLUSIONES	371
 CAPÍTULO IX		
	PROPEDÉUTICA O REALIDAD FÁCTICA ANTE LOS RETOS DE LA INTELIGENCIA ARTIFICIAL Alfredo Abadías Selma	373
1.	A MODO DE INTRODUCCIÓN	374
2.	INTELIGENCIA ARTIFICIAL: ¿DE QUÉ HABLAMOS?	380
3.	TIPOLOGÍA DE LA INTELIGENCIA ARTIFICIAL: ¿PROBLEMAS O SOLUCIONES?	394
4.	ALGUNAS CUESTIONES SOBRE LA REALIDAD VIRTUAL	397
5.	CONCLUSIONES	405

Miguel Ángel Presno LineraCatedrático de Derecho Constitucional de la Universidad de Oviedo

PRÓLOGO

Pocos días antes de escribir estas líneas pudimos conocer el texto final del Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión Europea, así como el del Convenio Marco del Consejo de Europa sobre Inteligencia Artificial, Derechos Humanos, Democracia y Estado de Derecho.

Como es conocido, ambas normas son el resultado presente de un proceso de debate político, jurídico, institucional, tecnológico, económico y académico que se remonta a varios años atrás y en el que se asume que la inteligencia artificial (IA en lo sucesivo) “cambiará nuestras vidas, pues mejorará la atención sanitaria (por ejemplo, incrementando la precisión de los diagnósticos y permitiendo una mejor prevención de las enfermedades), aumentará la eficiencia de la agricultura, contribuirá a la mitigación del cambio climático y a la correspondiente adaptación, mejorará la eficiencia de los sistemas de producción a través de un mantenimiento predictivo, aumentará la seguridad de los europeos y nos aportará otros muchos cambios que de momento solo podemos intuir. Al mismo tiempo, la IA conlleva una serie de riesgos potenciales, como la opacidad en la toma de decisiones, la discri-

minación de género o de otro tipo, la intromisión en nuestras vidas privadas o su uso con fines delictivos” (Libro Blanco sobre la inteligencia artificial de la Comisión, de 19 de febrero de 2020).

Uno de los ámbitos afectados por la IA es el de la Administración de Justicia y, de manera singular, aunque obviamente no exclusiva, los procesos penales, como ya destacó la Resolución del Parlamento Europeo, 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)), donde se recuerda (párrafos 3 y 4) que, habida cuenta del papel y la responsabilidad de las autoridades policiales y judiciales y del impacto de las decisiones que adoptan con fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, el uso de aplicaciones de IA debe clasificarse como de alto riesgo en los casos en que tienen potencial para afectar significativamente a la vida de las personas y que toda herramienta de IA desarrollada o utilizada por las autoridades policiales o judiciales debe, como mínimo, ser segura, robusta, fiable y apta para su finalidad, así como respetar los principios de minimización de datos, rendición de cuentas, transparencia, no discriminación y *explicabilidad*¹.

- 1 Según la Recomendación sobre la Ética de la Inteligencia Artificial de la UNESCO, aprobada en la reunión del 9 al 24 de noviembre de 2021, “la explicabilidad supone hacer inteligibles los resultados de los sistemas de IA y facilitar información sobre ellos. La explicabilidad de los sistemas de IA también se refiere a la inteligibilidad de la entrada, salida y funcionamiento de cada componente algorítmico y la forma en que contribuye a los resultados de los sistemas. Así pues, la explicabilidad está estrechamente relacionada con la transparencia, ya que los resultados y los subprocesos que conducen a ellos deberían aspirar a ser comprensibles y trazables, apropiados al contexto. Los actores de la IA deberían comprometerse a velar por que los algoritmos desarrollados sean explicables. En el caso de las aplicaciones de IA cuyo impacto en el usuario final no es temporal, fácilmente reversible o de bajo riesgo, debería garantizarse que se proporcione una explicación satisfactoria con toda decisión que haya dado lugar a la acción tomada, a fin de que el resultado se considere transparente”.

Conforme a las *Directrices éticas para una IA fiable* del Grupo de expertos de alto nivel sobre inteligencia artificial de la Unión Europea, “la explicabilidad es cru-

Esa Resolución hace hincapié en que el enfoque adoptado en algunos países no pertenecientes a la Unión Europea en relación con el desarrollo, el despliegue y el uso de tecnologías de vigilancia masiva interfiere de manera desproporcionada con los derechos fundamentales y, por lo tanto, no debe ser seguido por la Unión; destaca que también deben regularse de manera uniforme en toda la Unión las salvaguardias contra el uso indebido de las tecnologías de IA por parte de las autoridades policiales y judiciales, y subraya el impacto del uso de herramientas de IA en los derechos de defensa de los sospechosos, la dificultad para obtener información significativa sobre su funcionamiento y la consiguiente dificultad para impugnar sus resultados ante los tribunales, en particular por parte de las personas investigadas (párrafos 7 y 10).

En suma, en la Resolución se considera esencial, tanto para la eficacia del ejercicio del derecho de defensa como para la transparencia de los sistemas nacionales de justicia penal, que un marco jurídico específico, claro y preciso regule las condiciones, las modalidades y las consecuencias del uso de herramientas de IA en el ámbito de las actuaciones policiales y judiciales, así como los derechos de las personas afectadas y procedimientos eficaces y fácilmente accesibles de reclamación y recurso, incluidos los recursos judiciales. Subraya, además, el derecho de las partes en un procedimiento penal a

cial para conseguir que los usuarios confíen en los sistemas de IA y para mantener dicha confianza. Esto significa que los procesos han de ser transparentes, que es preciso comunicar abiertamente las capacidades y la finalidad de los sistemas de IA y que las decisiones deben poder explicarse –en la medida de lo posible– a las partes que se vean afectadas por ellas de manera directa o indirecta. Sin esta información, no es posible impugnar adecuadamente una decisión. No siempre resulta posible explicar por qué un modelo ha generado un resultado o una decisión particular (ni qué combinación de factores contribuyeron a ello). Esos casos, que se denominan algoritmos de «caja negra», requieren especial atención. En tales circunstancias, puede ser necesario adoptar otras medidas relacionadas con la explicabilidad (por ejemplo, la trazabilidad, la auditabilidad y la comunicación transparente sobre las prestaciones del sistema), siempre y cuando el sistema en su conjunto respete los derechos fundamentales. El grado de necesidad de explicabilidad depende en gran medida del contexto y la gravedad de las consecuencias derivadas de un resultado erróneo o inadecuado”.

tener acceso al proceso de recopilación de datos y a las evaluaciones conexas realizadas u obtenidas mediante el uso de aplicaciones de IA; destaca la necesidad de que las autoridades de ejecución participantes en la cooperación judicial, al decidir sobre una solicitud de extradición (o entrega) a otro Estado miembro o a un tercer país, evalúen si el uso de herramientas de IA en el país solicitante podría manifiestamente comprometer el derecho fundamental a un juicio justo; pide a la Comisión que elabore directrices sobre cómo llevar a cabo dicha evaluación en el contexto de la cooperación judicial en materia penal; insiste en que los Estados miembros, de conformidad con la legislación aplicable, deben velar por la información de las personas que sean objeto de aplicaciones de IA utilizadas por parte de las autoridades policiales o judiciales (párrafo 14).

En esta línea, en el recientemente aprobado Reglamento de IA de la Unión Europea se admite que deben clasificarse como de alto riesgo determinados sistemas de IA destinados a la Administración de justicia y los procesos democráticos, dado que pueden tener efectos potencialmente importantes para la democracia, el Estado de Derecho, las libertades individuales y el derecho a la tutela judicial efectiva y a un juez imparcial. En particular, a fin de hacer frente al riesgo de posibles sesgos, errores y opacidades, procede clasificar como de alto riesgo aquellos sistemas de IA destinados a ser utilizados por una autoridad judicial o en su nombre para ayudar a las autoridades judiciales a investigar e interpretar los hechos y el Derecho y a aplicar la ley a unos hechos concretos. También deben considerarse de alto riesgo los sistemas de IA destinados a ser utilizados por los organismos de resolución alternativa de litigios con esos fines, cuando los resultados de los procedimientos de resolución alternativa de litigios surtan efectos jurídicos para las partes. La utilización de herramientas de IA puede apoyar el poder de decisión de los jueces o la independencia judicial, pero no debe sustituirlas: la toma de decisiones finales debe seguir siendo una actividad humana (considerando 61).

Y en el Anexo III de este Reglamento se consideran de alto riesgo, lo que los someterá a una serie de requisitos previstos en la Sección Segunda del Capítulo III, los “sistemas de IA destinados a ser utilizados por una autoridad judicial, o en su nombre, para ayudar a una autoridad judicial en la

investigación e interpretación de hechos y de la ley, así como en la aplicación de la ley a un conjunto concreto de hechos, o a ser utilizados de forma similar en una resolución alternativa de litigios”.

Por su parte, el Convenio Marco del Consejo de Europa prevé, entre otras cosas, que cada Parte velará para que, cuando un sistema de inteligencia artificial tenga un impacto significativo en el disfrute de los derechos humanos, las personas afectadas dispongan de garantías procesales, salvaguardias y derechos procesales efectivos, de conformidad con el derecho internacional y nacional aplicable. Igualmente, cada Parte procurará que se notifique a las personas que interactúen con sistemas de inteligencia artificial que están interactuando con dichos sistemas y no con un ser humano (artículo 15).



Pues bien, en este contexto normativo europeo que trata de hacer frente a los desafíos jurídicos de la inteligencia artificial se insertan los diferentes estudios que componen esta publicación en equipo que, como dice el profesor Mario Caterini en la “Introducción”, desarrolla y propone una contaminación entre *humanitas* y *techne* en las deliberaciones sobre la *quaestio iuris*, asignando al juez el papel eventual de refutar la “probabilidad” de la interpretación más favorable al delincuente sugerida por la inteligencia artificial, que, de lo contrario, tendría que actuar como *default option*. Se trata de una colaboración interdisciplinar hispano-italiana en la que participan, además del mencionado profesor Caterini, los profesores e investigadores Maria Princi, Valentina Aragona, Mariana Rocca, Alberto D’Aiello, Gabriele Polito, Morena Gallo, Pere Simón y Alfredo Abadías.

Este cuidado y exhaustivo volumen nos ofrece, como se indica en el subtítulo, una serie de reflexiones sobre la toma de decisiones judiciales; entre ellas, y no con ánimo exhaustivo, sino meramente ejemplificativo, qué es una duda razonable en la era de la IA, si podemos “confiar” en la aplicación de la IA en un ámbito tan delicado como es de la Administración de Justicia, el análisis de las posibles aplicaciones algorítmicas en los procesos penales a partir de sus ventajas e inconvenientes, el impacto de estos sistemas en el tradicional principio de legalidad penal, la cuestión de cómo influirá en la

hermenéutica penal la era robótica en la que nos encontramos, cuáles son los posibles usos de la inteligencia artificial a la hora de motivar los actos jurisdiccionales, cuál será la solución probabilística en el proceso algorítmico, qué cabe esperar del varias veces mencionado Reglamento europeo sobre Inteligencia Artificial y, finalmente, cuál será el impacto de la IA en la práctica y la profesión legales en su conjunto.

Estas y otras reflexiones, interesantes tanto para quien se acerca a ellas desde el campo académico como para quienes se aproximan desde cualquiera de los ámbitos profesionales vinculados a la Administración de Justicia y, en especial, la jurisdicción penal, resultan de extraordinaria pertinencia en el momento actual y es de agradecer que se nos ofrezcan de manera sistematizada, rigurosa y exhaustiva.

En Oviedo, a 25 de abril de 2024.

Mario Caterini

INTRODUCCIÓN

La sombra de un futuro en el que los algoritmos podrían dirigir el destino de las decisiones judiciales parece una realidad, si no exactamente inminente, siquiera remota. A pesar de los temores fundados y generalizados sobre el riesgo de una posible deshumanización de la justicia, tal aterrizaje podría no ser indomable si se guía por la ciencia jurídica para garantizar el respeto de los principios *in primis* constitucionales. En el contexto de esta transformación inminente –si se permite forzar una comparación– se puede asemejar la idea del juez-robot al proyecto utópico de inspiración ilustrada, que proponía un juez mera *bouche de la loi*, mediante la aplicación aséptica de la ley sin ningún esfuerzo interpretativo.

Tras las razones que a lo largo del tiempo han hecho anhelar un juez-automatizado y el papel que los algoritmos actuales pueden jugar en revertir esta aspiración, es necesario comprender el funcionamiento de los llamados sistemas expertos basados en el conocimiento en el mundo del derecho, deteniéndonos en las posibles implementaciones de la interpretación robótica basada en precedentes jurisprudenciales. Observando las posibles fricciones con el principio de legalidad penal, conviene destacar las luces y sombras de la IA, estudiando también las experiencias de otros ordenamientos jurídicos. La comparación, de hecho, revela que la adopción de sistemas de inteligencia artificial en el mundo del derecho, en particular del derecho

penal, al basarse en la elaboración de precedentes jurisprudenciales, puede resultar menos problemática en los sistemas de *common law* que en los de *civil law*, como el italiano, en los que el principio constitucional de legalidad, al menos formalmente, parecería frenar su uso.

A la luz de la creciente relevancia de la jurisprudencia – que de hecho ya no se limita a la función de mera interpretación, sino que en algunos casos se ha convertido en un auténtico creador de derecho –, la opción de apoyarse en un algoritmo judicial podría estar en consonancia con la tendencia observada entre los jueces a simplificar la motivación de las sentencias mediante el reconocimiento de autoridad a los precedentes jurisprudenciales. Esta propensión, sin embargo, podría traducirse en una atenuación más que apreciable de las garantías constitucionales, en particular penales, si no se recurre a la teoría del probabilismo.

La doctrina probabilística desafía la lógica meramente actuarial del “más probable que no” que ha marcado hasta ahora el desarrollo de los primeros modelos de toma de decisiones jurídico-cibernéticas, proponiendo el uso del diferente criterio valorativo-político de “más allá de toda duda razonable”. El enfoque probabilístico –que hunde sus raíces en el pensamiento aristotélico– forja el principio universal de que cuando el límite de lo obligatorio parece incierto, debe favorecerse la libertad frente al castigo. Según esta lógica, la razonabilidad de la duda, que legitima la libertad de acción, se produce cuando la hipótesis más favorable es probable, aunque la desfavorable sea más probable. El probabilismo, por tanto, se erige en instrumento de la certeza en la medida en que, sin transformar lo probable en cierto, delimita la frontera de lo lícito cuando existen razones serias, aunque no incontestables, a favor de la libertad.

La teoría del probabilismo, por tanto, puede representar la base filosófica, coherente con los principios constitucionales, sobre la que fundamentar el posible funcionamiento de los algoritmos de decisión en el sistema penal. Los sistemas de inteligencia artificial, en efecto, entre las posibles soluciones interpretativas resultantes de la elaboración de precedentes jurisprudenciales, deberían optar no por lo estadísticamente más frecuente o “más probable”, sino por lo más favorable al acusado, que también es “probable”, grave,

porque se basa en otros precedentes jurisprudenciales. Es precisamente la interpretación favorable al acusado la que parece aplicar mejor los principios políticos y garantistas dictados por nuestra Constitución.

La inteligencia artificial, por supuesto, no debería excluir al juez del proceso de toma de decisiones, sino acompañarle. Esto, por un lado, reduciría el tiempo de respuesta de la autoridad judicial y facilitaría así la duración razonable del juicio; por otro lado, contribuiría a una mayor previsibilidad de las decisiones y a la uniformidad de la jurisprudencia. El algoritmo, por tanto, debería a la vez convertirse en una guía para el juez y simplificar la motivación de la sentencia en caso de que el mismo juez decida adherirse a la sugerencia maquinística más favorable al delincuente. Cuando, por el contrario, el juez quisiera apartarse de esta indicación y optar por una solución interpretativa más desfavorable, la carga motivacional debería reforzarse explicando las razones de la “improbabilidad” o “inverosimilitud” de la opción hermenéutica más favorable al acusado.

Sobre la base de investigaciones anteriores, en conclusión, este trabajo en equipo desarrolla y propone una contaminación entre *humanitas* y *techne* en las deliberaciones sobre la *quaestio iuris*, asignando al juez el papel eventual de refutar la “probabilidad” de la interpretación más favorable al delincuente sugerida por la inteligencia artificial, que, de lo contrario, tendría que actuar como *default option*.

Mario CateriniUniversidad de Calabria

CAPÍTULO I**DERECHO, ALGORITMOS Y JUSTICIA
PENAL: LA DUDA RAZONABLE EN LA
ERA DE LA INTELIGENCIA ARTIFICIAL**

SUMARIO: **1.** INTRODUCCIÓN. **2.** LA CERTEZA DEL *IUS DICERE* "ARTIFICIAL": ¿HA LLEGADO EL FUTURO?. **3.** LOS ALGORITMOS PREDICTIVOS EN EL ÁMBITO JURÍDICO: MODELOS Y DEFECTOS DE FUNCIONAMIENTO. **4.** LA "FACTUALIZACIÓN" CIBERNÉTICA DEL DERECHO. **5.** LA IMPLEMENTACIÓN ALGORÍTMICA DE LA REGLA DE LA LENIDAD. **6.** CONCLUSIONES. **7.** REFERENCIAS BIBLIOGRÁFICAS.

1. INTRODUCCIÓN

En este capítulo se examina el papel emergente de la inteligencia artificial en el ámbito de la justicia penal, no solo como una herramienta para esclarecer los hechos, sino también para interpretar y resolver cuestiones legales complejas. Se argumenta que, si bien los sistemas expertos actuales se basan en precedentes jurídicos, en un Estado de Derecho, no pueden simplemente aplicar reglas estadísticas, sino que deben adherirse al estándar

de “más allá de toda duda razonable”, incluso en la interpretación de la ley. Ante jurisprudencia contradictoria, estos algoritmos deberían favorecer la interpretación más favorable al acusado, y los jueces solo podrían apartarse de esta interpretación si refutan razonadamente la solidez del precedente más favorable. Este enfoque subraya la importancia de equilibrar la eficiencia de la inteligencia artificial con los principios fundamentales de justicia y equidad en el sistema legal.

En la era actual de avances tecnológicos, la integración de la inteligencia artificial en el sistema judicial está en constante evolución. La capacidad de utilizar algoritmos para analizar evidencia, interpretar leyes y ayudar en la toma de decisiones judiciales plantea tanto oportunidades como desafíos significativos. Este trabajo surge en un momento crucial en el que es necesario explorar cómo la inteligencia artificial puede influir en el funcionamiento de la justicia penal y cómo asegurar que se apliquen los principios fundamentales del derecho en este contexto.

La intersección entre el derecho, los algoritmos y la justicia penal es un tema de creciente relevancia e interés tanto en el ámbito jurídico como en el tecnológico. La posibilidad de que la inteligencia artificial tenga un papel cada vez más importante en la interpretación y aplicación de la ley plantea preguntas fundamentales sobre la equidad, la transparencia y la protección de los derechos individuales en el proceso judicial. Este tema despierta un debate intenso sobre cómo garantizar que la inteligencia artificial opere dentro de los límites éticos y legales, manteniendo la integridad del sistema judicial.

Los objetivos del presente trabajo consisten en analizar el impacto actual y potencial de la inteligencia artificial en el sistema de justicia penal; examinar los principios fundamentales del derecho que deben guiar el desarrollo y la implementación de algoritmos en el ámbito judicial; investigar cómo los sistemas de inteligencia artificial pueden integrarse de manera efectiva para mejorar la eficiencia y la equidad en el proceso judicial; proporcionar recomendaciones para garantizar que la inteligencia artificial en la justicia penal se adhiera a estándares éticos y legales, incluyendo el principio de “más allá de toda duda razonable”; contribuir al debate académico y jurídico sobre

el uso de los sistemas algorítmicos, no sólo al debate intelectual existente en Italia, sino muy especialmente al actual debate español e internacional, que cuenta con grandes y notables contribuciones de autores como Simón Castellano (2021, 2022 y 2023), Miró Llinares (2022) o Cotino Hueso (2022).

¿Es realmente posible superar los límites de la incertidumbre real de un sistema penal, a través de la búsqueda de su certeza “artificial”? Ésta es la pregunta básica que se plantea en las páginas siguientes. Esa certeza que se persigue en la aplicación de las normas, de hecho, podría orientar al jurista a mirar favorablemente también el ingreso de la inteligencia artificial en el sistema penal. El uso de algoritmos, siguiendo esta lógica, podría conducir a una mayor predictibilidad de las decisiones y del denominado derecho viviente, lo cual permitiría pronunciamientos libres de ideologías, orientaciones políticas y sentimientos, a menudo ocultos detrás de las sentencias del juez “humano”.

El recurso a la selección del precedente judicial para resolver casos similares –típico del common law, pero cada vez más arraigado en la práctica de los sistemas de civil law–, podría ser, por tanto, una operación encomendada al machine learning. A partir de estas premisas, queremos mostrar cuánto, en la actualidad, los avances logrados por la inteligencia artificial y las correlativas ventajas que éstas traerían en el derecho penal, corren el riesgo de dejar en un segundo plano el respeto de los mismos principios de matriz constitucional y supranacional que supuestamente mejor salvaguardarían.

El derecho penal, como es bien sabido, se basa en garantías particulares a favor del imputado y un “moderno” sistema basado en el uso de la inteligencia artificial no podría dejar espacio a formas de elusión del mismo, incluida la “duda razonable”, corolario del favor rei, que se ha consolidado en sistemas avanzados, tanto de civil law como de common law¹. En EE.UU.,

1 Para un análisis del estándar en clave comparativa y sobre el vínculo que une el principio de duda razonable al de presunción de inocencia, sea permitido el reenvío a: Caterini (2015). Sobre el tema de la duda razonable en la doctrina estadounidense, aquí nos limitamos a recordar, entre otros, Whitman (2008); Sheppard (2003), Dripps (1987), Waldman (1959).

por ejemplo, tal estándar encuentra su fundamento en la Carta Constitucional y, más precisamente, en la Due Process Clause contenida en la enmienda XIV², por lo que se enmarca dentro de los corolarios del debido proceso de la propia Constitución estadounidense, mientras que, regresando al contexto europeo, hoy también podemos hacer referimiento al art. 6 de la Directiva 2016/343 /UE, el cual dispone que «cualquier duda sobre la culpabilidad debe resolverse en favor del sospechoso o imputado»³.

El aporte, por tanto, tiene como objetivo analizar la compatibilidad del estándar de duda razonable con el uso de la IA en un sistema penal, bajo el particular aspecto de la interpretación de las normas. ¿Cómo se concilia el respeto de tal estándar si el algoritmo está programado para indicar cuál es el precedente más consolidado y, por tanto, estadísticamente más utilizado en casos análogos? Los sistemas cibernéticos, en efecto, pueden seleccionar el precedente jurisprudencial aplicable al caso concreto y así ofrecer al juez la “solución” por la que optar, pero hasta la fecha, no pudiendo explicar las razones jurídicas que subyacen a la “decisión”, el criterio selectivo resulta anclado a datos puramente estadísticos o probabilísticos derivados de la “consolidación” de precedentes judiciales.

Al ilustrar brevemente las aplicaciones concretas de la IA en el sistema penal que hoy se registran, este ensayo reconocerá el potencial de los algorit-

2 Fourteenth Amendment, US Constitution: «... nor shall any state deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws». La primera sentencia con la que la Suprema Corte reconoció el fundamento constitucional del BARD standard contando entre los corolarios del debido proceso, se trata de la sentencia *In re Winship*, 397 U.S. 358, 364 (1970), en <https://supreme.justia.com/cases/federal/us/397/358/>: «Due Process Clause of the Fourteenth Amendment protects the accused against conviction except upon proof beyond a reasonable doubt». Sobre esta sentencia, entre otras tantas, Allen (1980), Dripps (1987).

3 Sobre la validez de esta disposición de la Directiva 2016/343/UE, sea permitido el reenvío a Caterini (2019, pág. 330 ss.).