

## LA INTELIGENCIA ARTIFICIAL Y EL DERECHO: DOS PERSPECTIVAS

*Por Juliano Maranhão<sup>(\*)</sup>, Juliana Abrusio<sup>(\*\*)</sup>, Marco Almada<sup>(\*\*\*)</sup> y  
Nuria López<sup>(\*\*\*\*)</sup>*

### RESUMEN

Este trabajo señala la creciente utilización de las aplicaciones de inteligencia artificial en la sociedad (inteligencia artificial como objeto externo al derecho) y específicamente en el derecho (inteligencia artificial como objeto interno al derecho), con la consecuente necesidad y desafíos de una reglamentación del tema, un derecho de la inteligencia artificial. Discute la cuestión de la opacidad en la inteligencia artificial, la importancia de la explicabilidad, incluso delante de la reglamentación en protección de datos y sus límites técnicos. Más allá del desarrollo y de la innegable importancia de los principios éticos para las aplicaciones de inteligencia artificial, que llegan a algún consenso sobre la propia explicabilidad, la transparencia, la no discriminación y la responsabilidad, trae la necesidad de una disciplina específica y aplicable para cada sector para que sea efectiva, es decir, para que sea aplicable. Luego, elige como “sector” de análisis el derecho y las aplicaciones de inteligencia artificial como su objeto interno. Para tanto, trae dos perspectivas sobre los puntos positivos y negativos de la utilización de las técnicas de inteligencia artificial basada en los sistemas de conocimiento, y de las técnicas de aprendizaje automático en sus aplicaciones en el derecho a partir de su lógica de funcionamiento. Al final, concluye por la combinación de técnicas aprovechando las fortalezas de los enfoques presentados para construir sistemas más adecuados a las peculiaridades de las predicciones legales, de carácter normativo. Bajo esta perspectiva, los

---

(\*) Profesor asociado del Departamento de Filosofía y Teoría General del Derecho de la USP. Miembro del Comité Director de la Asociación Internacional de Inteligencia Artificial y Derecho. Director de Lawgorithm.

(\*\*) Doctora en Derecho por la PUC-SP. Profesora de la Universidade Presbiteriana Mackenzie. Directora del Instituto Legal Grounds for Privacy Design (Instituto LGPD).

(\*\*\*) Doctorando en Derecho en el Instituto Universitário Europeo. Mestre en Ingeniería Eléctrica por la Unicamp. Graduando en Derecho en la USP. Investigador de Lawgorithm

(\*\*\*\*) Doctora en Derecho por la PUC-SP. Investigadora del Instituto Legal Grounds for Privacy Design (Instituto LGPD).

profesionales del derecho deben hacer frente a los cambios tecnológicos, integrándose en equipos multidisciplinares para construir sistemas inteligentes que tengan efectos positivos y protejan los derechos e intereses legalmente protegidos.

*Palabras clave:* inteligencia artificial; derecho; explicabilidad.

### **ABSTRACT**

This paper highlights the growing use of artificial intelligence applications in society (artificial intelligence as an external object of Law) and specifically in Law (artificial intelligence as an internal object of Law), with the consequent need and challenges of a regulation of the subject, an Artificial Intelligence Law. It also discusses the issue of opacity in artificial intelligence, the importance of explicability, considering data protection regulations and its technical limits. Beyond the development and the undeniable importance of ethical principles for artificial intelligence applications, which reach some consensus on their own explicability, transparency, non-discrimination and responsibility, the paper raises the need for a specific and applicable discipline to each sector for the regulation to be effective, that is, to be applicable. Then, it chooses Law as its analysis "sector" and artificial intelligence applications as its internal object. Therefore, it brings two perspectives on the positive and negative issues of the use of artificial intelligence techniques based on knowledge systems and on machine learning techniques in their applications in Law based on its logic of functioning. In the end, it concludes for the combination of techniques, taking advantage of the strengths of the approaches presented to build systems more suited to the peculiarities of legal, normative predictions. Under this perspective, Law professionals must face technological changes, joining multidisciplinary teams to build intelligent systems that have positive effects and protect legally protected rights and interests.

*Keywords:* artificial intelligence; law; explicability.

## **I. INTRODUCCIÓN**

La inteligencia artificial es cada vez más relevante para el derecho por dos tendencias distintas y complementarias.

La primera es la creciente adopción de los sistemas inteligentes en varias aplicaciones, tanto en la toma de decisiones en los sectores públicos y privados cuanto en la construcción de sistemas vueltos al consumidor, como asistentes personales para las más diversas tareas cotidianas. Tal proliferación de inteligencias artificiales significa su involucramiento ubicuo en diversas relaciones sociales y económicas tuteladas por el derecho. En

este contexto, podemos hablar de un derecho de la inteligencia artificial, o sea, de la disciplina jurídica de los agentes digitales y de las implicaciones de su involucramiento en las relaciones jurídicas y en los conflictos de ellas decurrentes.

La segunda se debe a que la inteligencia artificial no es un objeto externo de la disciplina jurídica, sino una herramienta de aprendizaje para los profesionales del derecho, pero también una herramienta cada vez más utilizada por los profesionales del derecho. Aunque el empleo del aprendizaje automático haya presentado resultados extremadamente útiles para los abogados y para los tribunales, particularmente en lo que respecta al análisis predictivo, discutiremos que las correlaciones empíricas en las cuales están basadas enfrentan limitaciones en su capacidad explicativa, lo que compromete aplicaciones en el dominio jurídico, donde hay exigencia de justificación normativa de las decisiones. Por lo que creemos que la próxima generación de inteligencia artificial aplicada al derecho deberá incorporar modelos de representación de conocimiento jurídico a las herramientas de aprendizaje automático.

Este trabajo presenta los principales desafíos involucrados en la regulación de inteligencia artificial. Estas dos perspectivas están correlacionadas, puesto que la disciplina jurídica de la inteligencia artificial influenciará los sistemas construidos para finalidades jurídicas, además, como la tecnología puede volver posibles nuevos abordajes regulatorios. Por lo tanto, la actuación interdisciplinaria entre los profesionales del derecho, la ingeniería y la computación puede venir enhorabuena para garantizar que las aplicaciones jurídicas basadas en inteligencia artificial no lesionen o hasta efectiven los derechos e intereses protegidos por el derecho.

## II. LA REGULACIÓN DE LOS SISTEMAS INTELIGENTES

Un algoritmo es cualquier procedimiento de ordenador bien definido que posee algún valor agregado en la calidad de sus entradas (*inputs*), generando otros valores en la salida (*outputs*), de manera que puede ser considerada una herramienta para solucionar problemas <sup>(1)</sup>. Los algoritmos se utilizan en programas informáticos por diferentes organizaciones para la toma de decisiones y la asignación de recursos, basándose en grandes conjuntos de datos. Combinando cálculo, procesamiento y razonamiento, el *software* puede ser excepcionalmente complejo, codificando miles de variables en millones de puntos de datos <sup>(2)</sup>. De esta forma, las decisiones importantes sobre la vida

---

(1) CORMEN, Thomas H. — LEISERSON, Charles E. — RIVEST, Ronald L. — STEIN, Clifford, "Introduction to algorithms", The MIT Press, Cambridge, 2009, 3ª ed. ps. 5-6.

(2) CAPLAN, Robyn — DONOVAN, Joan — HANSON, Lauren — MATTHEWS, Jeanna, "Algorithmic accountability: a primer", Data & Society, 2018. Disponible en <https://datasociety.net/output/algorithmic-accountability-a-primer/>. Consultado en: 26/08/2018.

de las personas, como los procesos de selección de empresas y la concesión de créditos, son cada vez más provocadas por sistemas y algoritmos informáticos. Por tanto, existe la preocupación de cómo los programas informáticos basados en inteligencia artificial pueden limitar las oportunidades y, sobre todo, poner en riesgo los derechos fundamentales de los ciudadanos.

Entre los trabajos de la inteligencia artificial, uno de los principales focos de atención de la ley es la recogida y el tratamiento de datos personales para la elaboración de perfiles <sup>(3)</sup>. Determinar el perfil del individuo puede ser de gran valor para los anunciantes, las aseguradoras y el propio Estado. China, por ejemplo, ha utilizado una combinación de vigilancia que utiliza inteligencia artificial con el uso de enormes cantidades de datos personales para monitorear la vida y el comportamiento de las personas con todo detalle <sup>(4)</sup>.

Desde el punto de vista legal, debe existir la preocupación de permitir la posibilidad de impugnar y revisar decisiones basadas en inteligencia artificial, especialmente cuando se basan en técnicas computacionales opacas (*black box*) <sup>(5)</sup>. La contestabilidad y la posibilidad de revisión suponen la inteligibilidad, en términos humanos, es decir, la existencia de un conjunto de criterios determinantes que determinan sustancialmente una decisión.

No en vano, existen muchas iniciativas en el mundo académico jurídico y de la informática para pensar, reflexionar y proponer caminos sobre el

---

(3) Según el art. 4º [4] del Reglamento General de Protección de Datos de la Unión Europea (GDPR, en inglés), la elaboración de perfiles “es toda forma de tratamiento automatizado de datos personales consistente en utilizar datos personales para evaluar determinados aspectos personales de una persona física, en particular para analizar o predecir aspectos relativos al rendimiento profesional, situación económica, salud, preferencias personales, intereses, fiabilidad, comportamiento, ubicación o movimientos de dicha persona física”. En Brasil, la LGPD no define lo que se debe entender por perfiles, pero utiliza el concepto en dos ocasiones: en el art. 12 (datos anonimizados) y en el art. 20 (decisiones automatizadas).

(4) Para una visión más detallada de cómo China ha incorporado la sociedad de datos para sostener sus intereses políticos y económicos: LARSON, Christina, “Who needs democracy when you have data?”, *MIT Technology Review*, 20/08/2018. Disponible en <https://www.technologyreview.com/s/611815/who-needs-democracy-when-you-have-data/>. Consultado en 27/08/2018. Aun sobre el tema del Sistema de Crédito Social chino, por un programa de puntuación sobre sus ciudadanos, clasificándoles de acuerdo con las informaciones contenidas en un enorme banco de datos, alimentado con millones de informaciones personales (en su mayoría colectadas por Internet), tales como histórico de navegación, productos adquiridos, probabilidad de cumplimiento de obligaciones, etc. Ver GOMES, Rodrigo Dias de Pinho, “Big Data: desafios à tutela da pessoa humana na sociedade da informação”, *Lumen Juris*, Río de Janeiro, 2017, ps. 50-51. Por fin, es importante decir que la tutela de la privacidad y de la protección de datos en China no es comparable a la protección de datos en Occidente. En los últimos años, sin embargo, muchos casos fueron juzgados por los tribunales chinos sobre el tema e iniciativas legislativas han sido desarrolladas. Para un panorama de privacidad y protección de datos en China: ONG, Rebecca, “Recognition of the right to privacy on the Internet in China”, *International Data Privacy Law*, 3, v. 1, 2011, ps. 172-179. Disponible en <https://doi.org/10.1093/idpl/ipr008>. Accedido en: 10/10/2018.

(5) Jatinder Singh y coautores traen la dificultad en mantener el control sobre sistemas opacos: SINGH, Jatinder — WALDEN, Ian — CROWCROFT, Jon — BACON, Jean, “Responsibility and machine learning: part of a process”, SSRN, 2016, p. 7. Disponible en <https://ssrn.com/abstract=2860048>. Accedido en 18/10/2018.

tema. Entre otras iniciativas, se pueden mencionar Algorithm Watch <sup>(6)</sup> y, en Brasil, la asociación Lawgorithm <sup>(7)</sup>.

### II.1. La opacidad de los sistemas informáticos

La opacidad de los sistemas basados en el aprendizaje automático es una de las mayores fuentes de atención y preocupación en la actualidad, especialmente en lo que respecta a la posibilidad de impugnación, pero también con relación al riesgo de incorporar sesgos que se traduzcan en la construcción de perfiles o toma de decisiones discriminatorias, o incluso la posibilidad de toma de decisiones que desconozcan los valores humanos o no respeten los derechos fundamentales y la dignidad humana.

Tres factores contribuyen, en diversos grados, a que un sistema sea opaco: la complejidad de los modelos matemáticos involucrados, la dificultad para comprender las operaciones involucradas en el procesamiento de datos a gran escala y la falta de claridad en el contexto institucional del uso de estos sistemas <sup>(8)</sup>. La lógica de funcionamiento de los sistemas de inteligencia artificial involucra modelos computacionales de diversos grados de complejidad, que sirven de base para el desarrollo de programas informáticos para realizar una determinada tarea <sup>(9)</sup>. Comprender estos sistemas es, por tanto, una tarea que requiere conocimientos especializados <sup>(10)</sup>, pero incluso los expertos pueden tener dificultades para comprender todas las variables y factores involucrados en *big data* <sup>(11)</sup>, especialmente cuando se considera la presencia de secretos industriales y comerciales, o incluso de Estado, que pueden dificultar el acceso a los algoritmos por detrás de los sistemas informáticos o de las herramientas utilizadas para su implementación <sup>(12)</sup>.

(6) AlgorithmWatch (<https://algorithmwatch.org/de/>) es una iniciativa de los investigadores Lorena Jaume-Palasi, Lorenz Matzat, Matthias Spielkamp y Prof. Dr. med. Katharina Anna Zweig y es soportada por las autoridades de la media estatal de Hesse, Baviera, Baden-Württemberg, Renania-Palatinado, Sarre y Saxonia. Tiene foco en los algoritmos del gigante Google, así como en las elecciones del país.

(7) Lawgorithm es una asociación independiente, fundada por profesores de la ciencia de la computación, ingeniería, derecho, economía y filosofía de la Universidad de São Paulo, dedicada a la investigación sobre inteligencia artificial aplicada al derecho y a sus implicaciones jurídicas, económicas, sociales y culturales. Desde la perspectiva de la inteligencia artificial para el derecho (IA & Law), *Lawgorithm promueve la investigación sobre herramientas computacionales inteligentes capaces de traer más eficiencia a la actuación de profesionales del derecho y de generar informaciones sobre las actividades legislativas y jurisdiccionales. Analiza, bajo esa perspectiva, las nuevas cuestiones jurídicas traídas por la actuación de agentes digitales* (<http://www.lawgorithm.com.br/>). Consultado en 07/07/2020).

(8) BURRELL, Jenna, "How the machine 'thinks': Understanding opacity in machine learning algorithms", *Big Data & Society*, 1, v. 3, 2016, ps. 1-12.

(9) LEHR, David — OHM, Paul, "Playing with the Data: What Legal Scholars Should Learn About Machine Learning", *Davis Law Review*, v. 51, University of California, 2017, ps. 653-717.

(10) BURRELL, ob. cit., p. 4.

(11) BURRELL, ob. cit., ps. 4-5.

(12) BURRELL, ob. cit., ps. 3-4.

La presencia de uno o más de estos factores puede dificultar la identificación de lesiones o amenazas a derechos e intereses, sean individuales o colectivos, como resultado del uso de inteligencia artificial. Esta dificultad puede incluso servir como un instrumento para las empresas y otros controladores de datos, al mantener sus operaciones opacas, para evitar o confundir los esfuerzos regulatorios <sup>(13)</sup>. Como advierte el jurista estadounidense Frank Pasquale, las diversas formas de opacidad que rodean a los algoritmos nos dejan en la oscuridad con relación a las decisiones cruciales <sup>(14)</sup>. Para Pasquale, debería haber menos esfuerzo para concentrarse en tratar de controlar la recopilación de datos y más esfuerzo para regular el uso de esos datos, es decir, cómo las empresas y los Gobiernos están realmente implementando reglas para tomar decisiones, utilizando inteligencia artificial <sup>(15)</sup>. Es con relación a estos riesgos que los esfuerzos se dirigen, no solo a lo legal, sino también a la incorporación de criterios éticos en el desarrollo de esos sistemas.

## *II.2. Principios éticos e iniciativas regulatorias*

En los últimos años, en respuesta a las preocupaciones sobre el uso de la inteligencia artificial, especialmente las basadas en el aprendizaje automático, organismos gubernamentales, asociaciones de investigación y organizaciones privadas han elaborado documentos que proponen parámetros éticos para el desarrollo y aplicación de sistemas de inteligencia artificial. Según un estudio de Berkman Klein Center for Internet & Society <sup>(16)</sup>, existen al menos 36 conjuntos de principios publicados por Gobiernos, organizaciones internacionales, empresas y organizaciones del tercer sector, que son adoptados de maneras distintas —o con énfasis distintos— para el desarrollo de inteligencia artificial ética.

En estos documentos, existe convergencia en torno a los principios de transparencia (debe ser claro para el usuario que interactúa con un sistema artificial), explicabilidad (divulgación de información que permite al usuario comprender los criterios para la toma de decisiones), no discriminación (evitar que los sistemas corporativos tengan sesgos que puedan ofender los derechos fundamentales), no maleficencia (los sistemas de inteligencia artificial no pueden dañar a los seres humanos), responsabilidad y privaci-

---

(13) Además de BURRELL, ob. cit., ps. 3-4, también PASQUALE, Frank, "The Black Box Society: the secret algorithms that control money and information", Harvard University Press, Cambridge, 2015, p. 2.

(14) "[...] secrecy is approaching critical mass, and we are in the dark about crucial decisions. Greater openness is imperative" (PASQUALE, ob. cit., p. 3). En traducción libre: "[...] el sigilo se está aproximando de la masa crítica, y estamos en el oscuro sobre decisiones cruciales. Más apertura se hace imperativa".

(15) PASQUALE, ob. cit., p. 141.

(16) FJELD, Jessica *et al.*, "Principled Artificial Intelligence: A Map of Ethical and Rights-Based Approaches". Relatório técnico, Berkman Klein Center For Internet & Society, septiembre de 2018.

dad/protección de datos, aunque existen diferencias sobre el significado y sobre cómo implementar estos principios <sup>(17)</sup>.

Uno de los actores más activos en el ámbito de las propuestas éticas para la inteligencia artificial es la Unión Europea, que actúa a través de sus distintos órganos <sup>(18)</sup>. Las iniciativas en esta dirección se pueden ver en el trabajo del European Data Protection Supervisor <sup>(19)</sup>, del Parlamento Europeo <sup>(20)</sup>, y de la Comisión Europea <sup>(21)</sup>. En Brasil, el Ministerio de Ciencia, Tecnología e Innovación abrió una consulta pública sobre la elaboración de una Estrategia Nacional de Inteligencia Artificial, que tiene entre sus pilares el establecimiento de proposiciones éticas para sistemas inteligentes <sup>(22)</sup>.

El establecimiento de principios éticos para guiar el desarrollo de la inteligencia artificial es necesario para orientar la aplicación de la inteligencia artificial con fines socialmente positivos, pero conlleva algunos riesgos. Una primera dificultad está relacionada con el nivel adecuado del uso de la inteligencia artificial. Por un lado, existe la posibilidad de sobreuso, con el

(17) JOBIN, Anna — IENCA, Marcello — VAYENA, Effy, “The global landscape of AI ethics guidelines”, *Nature Machine Intelligence*, v. 1, 2019, ps. 389-399.

(18) Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions - Coordinated Plan on Artificial Intelligence (COM[2018] 795 final).

(19) En diciembre de 2015, el European Data Protection Supervisor ha creado el Ethics Advisory Group dedicado a las dimensiones éticas de la protección de datos en el actual contexto digital. En 2018, este Grupo ha producido un informe que sostiene que la nueva ética digital está basada en el derecho fundamental a la privacidad y a la protección de datos personales, y la comprende como crucial para la dignidad humana, base de la Carta de Derechos Fundamentales de la Unión Europea. BURGESS, Peter — FLORIDI, Luciano — POLS, Aurélie — van den HOVEN, Jeroen, “Towards a digital ethics”, Ethics Advisory Group, Bruselas, 2018.

(20) Ver KOENE, Ansgar, et al., “A governance framework for algorithmic accountability and transparency”, Parlamento Europeo, Bruselas, 2019. Este Informe desarrolla y propone opciones de regulaciones para la gobernanza de la transparencia algorítmica y de su responsabilización basada en el análisis de los aspectos sociales, técnicos y regulatorios.

(21) En el “A definition of AI: Main capabilities and scientific disciplines”, el High-Level Expert Group (AI HLEG) nombrado por la Comisión Europea ha propuesto una definición operacional de inteligencia artificial - “Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions —with some degree of autonomy— to achieve specific goals”. Además, describe conceptos relacionados como *machine learning* y *robótica*. El AI HLEG también presentó las “Ethics Guidelines on Artificial Intelligence”, con siete requisitos para una inteligencia artificial centrada en humanos. Por fin, la Comisión Europea publicó, en 2020, el *white paper* “On Artificial Intelligence - A European Approach to excellence and trust”, en el cual señala la reglamentación, especialmente en áreas de riesgo como salud, transporte, energía y parte de los servicios públicos, además de las aplicaciones que impacten a los trabajadores o que tengan identificación biométrica (reconocimiento facial). No hay una sugestión de reglamentación o prohibición de metodologías, sino que la indicación de la conveniencia de la adopción, por los programadores de IA, de transparencia interna: documentación del proceso de decisión del desarrollo del *software* (concepción, entrenamiento, lanzamiento, seguimiento), así como la importancia de evaluaciones del agregado de *outputs*, que sean fácilmente accedido en auditorías.

(22) Para el anuncio de la consulta pública, ver [https://www.mctic.gov.br/mctic.opencms/salaImprensa/noticias/arquivos/2019/12/MCTIC\\_lanca\\_consulta\\_publica\\_para\\_a\\_Estrategia\\_Brasileira\\_de\\_Inteligencia\\_Artificial.html](https://www.mctic.gov.br/mctic.opencms/salaImprensa/noticias/arquivos/2019/12/MCTIC_lanca_consulta_publica_para_a_Estrategia_Brasileira_de_Inteligencia_Artificial.html). Los autores de este artículo también participaron de la consulta, por la Asociación Lawgorithm, lo que puede ser accedido en <https://lawgorithm.com.br/estrategia-nacional-de-inteligencia-artificial/>.

uso de sistemas inteligentes en situaciones que pueden tener un impacto negativo en el ser humano; por otro, el temor a esta posibilidad puede llevar a la subutilización de la inteligencia artificial, impidiendo que la sociedad disfrute de los beneficios que se pueden extraer de estas tecnologías <sup>(23)</sup>.

Otra dificultad radica en el hecho de que los principios se formulan en discusiones *top-down* y reclaman universalidad. Esto hace que su contenido sea demasiado genérico y abstracto, lo que dificulta su aplicación. Por un lado, la definición de reglas generales en lugar de una regulación estricta favorece el desarrollo tecnológico y permite que la sociedad se familiarice con los sistemas inteligentes antes de decidir cómo deben regularse. Por otro lado, puede contribuir a lo que la literatura sobre regulación de la inteligencia artificial llama *ethics-washing*: el uso de principios vagos como sustituto de reglas que protegerían efectivamente los derechos e intereses individuales y colectivos que podrían verse afectados por el uso de sistemas inteligentes <sup>(24)</sup>.

Así, uno de los grandes desafíos en la regulación de la inteligencia artificial es identificar no solo los principios aplicables, sino los momentos en los que estos principios deben ser implementados por normas legales, así como los instrumentos legales más adecuados para esta regulación <sup>(25)</sup>.

Entendemos que el mejor camino para las discusiones éticas que pretenden la aplicabilidad es a través del análisis *bottom-up*, o sea, buscar el equilibrio reflexivo entre principios generales y casos concretos en sectores específicos. Es decir, tratar no de discutir los principios universales de la ética computacional o algorítmica, sino de diseñar principios específicos para diferentes sectores de aplicación: la ética algorítmica en el campo de la medicina, en el ámbito jurídico, en el ámbito comercial, etc. Al reflejar las prioridades existentes en cada dominio de aplicación, será posible construir sistemas inteligentes que se centren en las demandas humanas <sup>(26)</sup>.

(23) FLORIDI, Luciano, *et al.*, “AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations”, *Minds and Machines*, v. 28, 2018, ps. 689-707.

(24) WAGNER, Ben, “Ethics as an Escape from Regulation: From ethics-washing to ethics-shopping?”, en HILDEBRANDT, Mireille (ed.), *Being Profiling. Cogitas ergo sum*, Amsterdam University Press, Amsterdam, 2018. Para una defensa de la relevancia de la investigación ética sobre inteligencia artificial, ver BIETTI, Elettra, “From ethics washing to ethics bashing: a view on tech ethics from within moral philosophy”, en Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT\* ’20), Association for Computing Machinery, Barcelona, Spain, 2020, ps. 210-219.

(25) Ver MARANHÃO, Juliano — COUTINHO, Diogo R. “Melhor investir do que regular”. *Correio Braziliense*, 25/03/2019. Y para un ejemplo concreto de los riesgos involucrados en la reglamentación de tecnología es instructivo considerar la experiencia sudcoreana con internet banking, en la cual una ley que impuso una tecnología como representante del estado del arte expuso los usuarios coreanos de internet a una amplia gama de ataques en el ambiente online, como analizado por SWANDA, Gus, “The Deficiencies of a Westphalian Model for Cyberspace: A Case Study of South Korean Cyber Security”, *International Journal of Korean Unification Studies*, 2, v. 25, 2016.

(26) Comissão Europeia, *ob. cit.*; BRYSON, Joanna J. — THEODOROU, Andreas, “How Society Can Maintain Human-Centric Artificial Intelligence”, en TOIVONEN, Marja — SAARI, Eveliina (orgs.), *Human-Centered Digitalization and Services*, Springer Singapore, Singapore, 2019, ps. 305-323.



El equilibrio ideal dependerá no solo de las capacidades tecnológicas de los sistemas inteligentes, sino también de las peculiaridades de cada dominio de aplicación y de los requisitos de la ley. En el caso de la inteligencia artificial aplicada al derecho, por ejemplo, el problema de *black box*, lo comentado anteriormente aparece con más fuerza, ya que las decisiones judiciales están sujetas a una serie de requisitos que van más allá de la mera precisión predictiva, como la necesidad de que el contenido de la decisión esté racionalmente justificado. A continuación, exploraremos con más detalle los usos legales de la inteligencia artificial y el estado del arte en relación con el problema de hacer explicables las decisiones de estos sistemas.

### III. LA APLICACIÓN DE LA INTELIGENCIA ARTIFICIAL EN EL DERECHO

Los sistemas basados en inteligencia artificial son utilizados, como vimos anteriormente, para automatizar tareas que requerirían inteligencia humana para realizarlas. Entre estas tareas, cobra especial relevancia el uso de sistemas inteligentes para intentar predecir el comportamiento de las personas, lo que engloba distintos objetivos, como la identificación de las preferencias de consumo de alguien o la determinación de la probabilidad de que pague de vuelta un préstamo <sup>(27)</sup>.

Estos modelos, como se discutió anteriormente, pueden ser opacos para la comprensión humana. La opacidad puede ser el resultado de tres fuentes principales: la complejidad de los modelos matemáticos involucrados, la dificultad de comprender las operaciones involucradas en el procesamiento de datos a gran escala y la falta de claridad en el contexto institucional de uso de estos sistemas. Dado que las actividades jurídicas —sea en el ámbito judicial, sea en el ámbito administrativo, o en otros métodos de solución de controversias— están generalmente relacionadas con situaciones que pueden tener impactos significativos en las personas naturales y jurídicas, el ordenamiento jurídico impone una serie de restricciones que debe seguirse para el uso lícito de la inteligencia artificial como herramienta de soporte o automatización.

Una primera fuente de estas restricciones, como se mencionó anteriormente, es la regulación legal de la privacidad y protección de datos personales: los principios éticos mencionados anteriormente también se aplican a los usos legales de la inteligencia artificial, así como a las leyes que se basan en ellos. En concreto, la Ley General de Protección de Datos (LGPD, ley 13.709/2018) trae reglas relevantes que deben ser utilizadas en aplicacio-

---

(27) Ver GOMES, ob. cit., 2017, ps. 50-51.

nes, como la generación automática de contratos o la búsqueda automatizada por jurisprudencia, que están vinculadas a cuestiones legales <sup>(28)</sup>.

Además de estas preocupaciones generales, hay también cuestiones legales específicas. Un ejemplo es el requisito de que las decisiones judiciales se basen en el análisis de los aspectos fácticos y legales presentes en el proceso <sup>(29)</sup>. Ante este requisito, un sistema inteligente que se proponga automatizar una decisión judicial —o, de manera más realista para la tecnología actual, proporcionar contribuciones a un juez humano— debe ser capaz de proporcionar la justificación legalmente requerida. Sin embargo, la opacidad involucrada en el uso y construcción de sistemas inteligentes puede dificultar, o incluso hacer impracticable, la producción de este tipo de razonamientos.

La LGPD, así como su contraparte de la Unión Europea, el Reglamento General de Protección de Datos (RGPD o GDPR), incluye una forma del llamado derecho a explicación <sup>(30)</sup>. Según la LGPD, en su art. 20, §1º, el responsable del tratamiento de los sistemas que toman decisiones basadas únicamente en el tratamiento automatizado de datos personales debe proporcionar información sobre los criterios y procedimientos utilizados para la decisión automatizada. Esta formulación del derecho se acerca al paradigma de los sistemas basados en el conocimiento, que operan a partir de representaciones predefinidas del conocimiento disponible sobre el problema que pretenden resolver <sup>(31)</sup>.

Sin embargo, muchas de las aplicaciones de inteligencia artificial que ahora juegan un papel destacado en la imaginación popular se basan en otro paradigma: los sistemas de datos. Este enfoque, en los últimos años, se ha utilizado para construir sistemas de búsqueda de divergencias en fallos de la Corte Suprema Federal <sup>(32)</sup>, predicción de decisiones de la Corte Eu-

---

(28) La LGPD establece, en su art. 7º, VI, el ejercicio regular de derechos en el proceso judicial, administrativo o arbitral como una de las hipótesis para el tratamiento lícito de datos personales. El cap. IV de la ley (arts. 23 a 32) regla el tratamiento de datos automatizados en el poder público, lo que abarca el Judicial y el Ministerio Público, así como las instancias procesuales administrativas.

(29) *Vide, e. g., la exigencia de justificación en la Constitución, art. 93, IX, y como elemento esencial de la sentencia por el Código de Proceso Civil, art. 489, caput.*

(30) Para un panorama de la discusión sobre los límites del derecho a la explicación en la regulación de la Unión Europea, ver KAMINSKI, Margot, "The Right to Explanation, Explained", *Berkeley Technology Law Journal*, 1, v. 34, 2019. Para una discusión sobre los límites de la viabilidad de la explicación delante de las posibilidades tecnológicas, ver ALMADA, Marco, "Human intervention in automated decision-making: Toward the construction of contestable systems", Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law (ICAAIL 2019), ACM Press, Montreal, Canadá. Nueva York, 2019, seção 3.

(31) RUSSELL, Stuart J. — NORVIG, Peter, "Artificial Intelligence: a modern approach", Prentice Hall, Upper Saddle River, 2010, 3ª ed., p. 234.

(32) OLIVEIRA, Rafael Brito de, "Utilização de Ontologias para Busca em Base de Dados de Acórdãos do STF", dissertação (Maestría en Ciencias de la Computación), Universidade de São Paulo, São Paulo, 2017.

ropea de Derechos Humanos <sup>(33)</sup> y la aplicación de tests de propósito en la tributación internacional <sup>(34)</sup>, por ejemplo.

Los sistemas de datos no se construyen sobre la base de representaciones predefinidas del problema que están tratando de resolver. En cambio, operan según el aprendizaje automático <sup>(35)</sup>: antes de ser utilizados para su aplicación, estos sistemas se exponen a datos sobre el problema a resolver, y mediante procesos de formación estadística <sup>(36)</sup>. Así, extraen de la base de datos estándares y correlaciones que se generalizarán para resolver problemas futuros.

Estos métodos de predicción son empíricos, por ejemplo, en aplicaciones para el derecho; pueden extraer los resultados de una serie de decisiones judiciales y correlacionarlos con factores como el tipo de demanda, el monto involucrado y el tribunal en el que se juzga la demanda para predecir un resultado probable. Como su estructura se basa en las propiedades detectables a partir de los datos, los sistemas basados en *machine learning* no tienen en cuenta ninguna justificación normativa sobre cómo debe basarse la decisión en las características y los argumentos del caso.

Además, la complejidad matemática de los modelos utilizados para extraer las propiedades de los datos hace imposible, incluso para el observador técnico, especificar en detalle cómo operan estos sistemas. Cuando estas técnicas se aplican en *big data* <sup>(37)</sup>, la escala de las operaciones involucradas hace aún más complejo el escenario, dificultando exponer las decisiones tomadas por el sistema de una manera comprensible para los humanos, con premisas, criterios, argumentos y conclusiones accesibles <sup>(38)</sup>.

En esta sección, discutiremos los éxitos y las limitaciones de ambos paradigmas de aprendizaje automático. Por un lado, los *sistemas basados en el conocimiento* son capaces de representar una amplia gama de aplicaciones legales, así como de producir respuestas inteligibles, pero su construcción y uso requiere un esfuerzo que a menudo dificulta su operacionalización. Los *sistemas de aprendizaje automático*, por otro lado, pueden extraer corre-

(33) MEDVEDEVA, Masha — VOLS, Michel — WIELING, Martijn, “Using machine learning to predict decisions of the European Court of Human Rights”, *Artificial Intelligence and Law*, 2019.

(34) KUZNIACKI, Blazej, “The Artificial Intelligence Tax Treaty Assistant: Decoding the Principal Purpose Test”, *Bulletin for International Taxation*, 9, v. 72, 2018.

(35) HAYKIN, Simon, “Neural Networks and Learning Machines”, Prentice Hall, 2008, p. 34.

(36) El proceso de aprendizaje puede ser: supervisado, en donde el sistema aprende a asociar determinados valores de los datos de entrada a través de etiquetas proporcionadas; de refuerzo, cuando las respuestas del sistema son realizadas a partir de la recompensa que se le provee por la respuesta correcta; o no-supervisado, donde el aprendizaje se adecúa a una métrica que refleja propiedades de los datos y no un resultado deseado (HAYKIN, ob. cit., ps. 34-45).

(37) El proceso de aprendizaje puede ser: supervisado, en donde el sistema aprende a asociar determinados valores de los datos de entrada a través de etiquetas proporcionadas; de refuerzo, cuando las respuestas del sistema son realizadas a partir de la recompensa que se le provee por la respuesta correcta; o no-supervisado, donde el aprendizaje se adecúa a una métrica que refleja propiedades de los datos y no un resultado deseado (HAYKIN, ob. cit., ps. 34-45).

(38) Sobre ambas las cuestiones, ver BURRELL, ob. cit.

laciones estadísticas con menos estructura que los sistemas basados en el conocimiento, pero sus soluciones no son explicables a un nivel compatible con los requisitos legales ni se benefician del conocimiento legal.

### *III.1. Sistemas basados en el conocimiento jurídico*

Los sistemas basados en el conocimiento son sistemas inteligentes que hacen inferencias basadas en representaciones internas del conocimiento (39). Estas representaciones del conocimiento se construyen como supuestos, al diseñar un sistema informático, pero, una vez realizadas, el sistema puede actualizar su base de conocimiento a partir de la información que adquiere del entorno, aplicando las mismas reglas de inferencia para tratar situaciones percibidas.

En el caso de los sistemas basados en el conocimiento jurídico, este razonamiento opera a través de la estructura formal, que puede construirse a partir de diferentes lógicas, como las lógicas deónticas <sup>(40)</sup>, que tratan conceptos como obligaciones, prohibiciones y permisos, sea por operadores modales o mediante representaciones de conjuntos de normas, constitutivas, obligatorias o permisivas, como en las lógicas de entrada/salida, o incluso, y lógicas de argumentación derrotable.

La lógica defendible del argumento encontró una amplia aplicación en la representación del conocimiento jurídico, ya sea el razonamiento basado en precedentes (razonamiento basado en casos) o el razonamiento basado en reglas. Estas lógicas modelan el razonamiento como inferencias de argumentos a favor o en contra de una tesis establecida, incluida la información sobre la fuerza relativa de estos argumentos. Un argumento puede ser una estructura inferencial compleja, que vincula sus premisas a conclusiones a través de pasos intermedios detallados <sup>(41)</sup>. Estos argumentos, a su vez, pueden ser atacados en diferentes uniones y de diferentes formas, y se puede derivar una conclusión si es posible construir un argumento a favor de la conclusión que sea defendible frente a todos los argumentos que la atacan.

Esta estructura basada en argumentos, contraargumentos, refutaciones y supuestos está directamente alineada con el razonamiento jurídico, considerando que las decisiones judiciales se toman a partir de la valoración de los argumentos de las partes contrarias. Con eso, la lógica de la argumentación se ha aplicado con éxito para representar varios aspectos del razonamiento jurídico. Entre las aplicaciones exitosas de la lógica de la

---

(39) RUSSELL — NORVIG, ob. cit., p. 234.

(40) Ver HILPINEN, Risto — MCNAMARA, Paul, “Deontic Logic: a historical survey and introduction”, en GABBAY, Dov *et al.* (orgs.), *Handbook of Deontic Logic and Normative Systems*, College Publications, Londres, 2013.

(41) Desde un punto de vista técnico, tanto los pasos intermedios como la conclusión final deben ser autorizados por operaciones de una lógica deductiva o derrotable (no-monotónica).

argumentación jurídica en inteligencia artificial, tenemos modelos de razonamiento sobre la argumentación oral en los tribunales <sup>(42)</sup>, evidencias procesuales <sup>(43)</sup> y racionios con precedentes judiciales <sup>(44)</sup>. En particular, más recientemente, se han desarrollado modelos lógicos de argumentación en los que se produce un equilibrio de múltiples pros y contras <sup>(45)</sup> (factores del caso, razones, principios y valores). Además, algunos modelos formulan principios racionales sobre el desarrollo de precedentes en el tiempo y sobre la dinámica de construcción y alteración de conceptos en interpretaciones de leyes y precedentes judiciales <sup>(46)</sup>.

Gran parte de los avances en inteligencia artificial y el derecho (AI & Law) en la interpretación de conceptos giran en torno a la noción de “factor”, que surgió en dos programas informáticos pioneros en AI & Law: o sistema HYPO <sup>(47)</sup> e o sistema CATO <sup>(48)</sup>. Los factores son abstracciones o estereotipos de la descripción de un caso, que pueden favorecer (factores favorables) o perjudicar (factores contrarios) una conclusión legal. Por ejemplo, el factor “embarazo resultante de una violación” es un factor favorable, según la legislación brasileña, a la decisión de permitir o no sancionar el aborto. El factor “procedimiento no realizado por un médico” es un factor que conduce a la prohibición del aborto.

La atención prestada al papel de los valores y propósitos llevó a planteamientos de la interpretación jurídica como un problema de decisión, es decir, como una elección entre interpretaciones alternativas considerando la probabilidad de esa interpretación o de decisiones basadas en esa in-

(42) BENCH-CAPON, Trevor J. M. — PRAKKEN, Henry, “Using argument schemes for hypothetical reasoning in law”, *Artificial Intelligence and Law*, v. 18, 2010, ps. 153-174.

(43) Ver, e. g., BEX, Floris J. et al., “Towards a formal account of reasoning about evidence: argumentation schemes and generalisations”, *Artificial Intelligence and Law*, v. 11, 2003, ps. 125-165; y VERHEIJ, Bart, “Dialectical argumentation with argumentation schemes: an approach to legal logic”, *Artificial Intelligence and Law*, v. 11, 2003, ps. 167-195.

(44) Ver, e. g., BENCH-CAPON, Trevor J. M., “The missing link revisited: the role of teleology in representing legal argument”, *Artificial Intelligence and Law*, v. 10, 2002, ps. 79-94; e PRAKKEN, Henry — SARTOR, Giovanni, “A dialectical model of assessing conflicting arguments in legal reasoning”, *Artificial Intelligence and Law*, v. 4, 1996, ps. 331-368.

(45) Ver, e.g., HAGE, Jaap C., “Comparing alternatives in the law”, *Artificial Intelligence and Law*, v. 12, 2005, ps. 181-225; SARTOR, Giovanni, “The logic of proportionality: reasoning with non-numerical magnitudes”, *German Law Journal*, v. 14, 2013; PRAKKEN, Henry, et al., “A formalisation of argumentation schemes for legal case-based reasoning in ASPIC+”, *Journal of Logic and Computation*, 2013; y HORTY, John F., “Rules and reasons in the theory of precedent”, *Legal Theory*, v. 17, 2011, ps. 1-33.

(46) Ver, e. g., MARANHÃO, Juliano — SARTOR, Giovanni, “Value Assessment and Revision in Legal Interpretation”, Proceedings of the 17th edition of the International Conference on Artificial Intelligence and Law (ICAIL 2019), ACM Press, Montreal, Canadá. Nova York, 2019; y HORTHY, John F. — BENCH-CAPON, Trevor J., “A factor-based definition of precedential constraint”, *Artificial Intelligence and Law*, v. 20, 2012, ps. 181-214.

(47) Introduzido por ASHLEY, Kevin D., “Modeling Legal Argument: Reasoning with Cases and Hypotheticals”, MIT Press, Cambridge, 1990.

(48) Introduzido por ALEVEN, Vincent, “Using background knowledge in case-based legal reasoning: a computational model and an intelligent learning environment”, *Artificial Intelligence*, 150, 2003, ps. 183 ss.

interpretación, o incluso de las posibles consecuencias de esa interpretación en términos de promoción o degradación de los valores relevantes. En esos enfoques, la elección de una interpretación se basa en los efectos positivos y negativos que las decisiones potenciales pueden tener en relación con los valores u objetivos que se toman como relevantes, considerando las preferencias relativas entre esos valores u objetivos, y, en algunos modelos, considerando también la medida en que estos valores u objetivos se logran o se impactan. También existen vínculos entre estos enfoques cualitativos y las teorías de la decisión <sup>(49)</sup>, recientemente exploradas por Giovanni Sartor <sup>(50)</sup>.

Sistemas computacionales más recientes, como o VJAP <sup>(51)</sup>, desarrollados para su aplicación en temas de competencia desleal, en particular en el derecho del secreto comercial, buscan incorporar valores y propósitos perseguidos en la construcción de justificaciones de decisiones judiciales. Con ello, el sistema predice el desenlace del caso, es decir, la probable decisión judicial, mediante una medida de confianza derivada de los gráficos argumentativos, y genera textos con argumentos que justifican la predicción.

En todos estos enfoques, es posible reconstruir la lógica que utilizó el sistema para construir sus predicciones, lo que proporciona una justificación racional de las decisiones tomadas. Sin embargo, la construcción de estos sistemas no solo requiere el uso de muchos conocimientos específicos sobre el tema, sino que también existe la necesidad de extraer manualmente la información que utilizarán los algoritmos, lo que requiere, en la práctica, un trabajo sustancial de preprocesamiento para la construcción y uso de estos sistemas. Por esta razón, los sistemas basados en el conocimiento no se aplican hoy en día en la potencialidad que podría sugerir su sofisticación teórica.

### *III.2. Aplicaciones jurídicas del aprendizaje automático*

En los últimos años, se han empleado técnicas de aprendizaje automático para predecir el resultado de las decisiones judiciales basadas en textos legales. Un ejemplo es el modelo construido por Nikolaos Aletras y coautores <sup>(52)</sup>, que alcanzó un 79% de precisión al valorar si el Tribunal Europeo de

---

(49) KEENEY, Ralph L. — RAIFFA, Howard, "Decisions with Multiple Objectives", Wiley, New York, 1976.

(50) Ver, dentro otros trabajos del autor, "Fundamental legal concepts: A formal and teleological characterisation", *Artificial Intelligence and Law*, v. 21, 2010, ps. 101-142, e ob. cit., 2013.

(51) GRABMAIR, Matthias, "Predicting trade secret case outcomes using argument schemes and learned quantitative value effect tradeoffs", Proceedings of the 16th edition of the International Conference on Artificial Intelligence and Law (ICAIL 2017), ACM Press, Londres. Nueva York, 2017, ps. 89-98.

(52) ALETRAS, Nikolaos et al, "Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective", *PeerJ Computer Science*, v. 2, e93, 2016.

Derechos Humanos decidiría por la existencia de una vulneración de derechos en un determinado caso. En aplicaciones legales, los enfoques basados en *support vector machines* (SVM) han obtenido los mejores resultados, aunque los enfoques basados en el aprendizaje profundo estén ganando espacio <sup>(53)</sup>.

En general, estos modelos basan sus predicciones en elementos textuales (características) que revelan estándares de decisiones. La construcción de un modelo de aprendizaje automático para el procesamiento del lenguaje natural, que incluye el procesamiento de textos legales, generalmente tiene cuatro pasos: 1) la compilación de un corpus de textos relevantes para el campo de aplicación; 2) el preprocesamiento de los textos de este corpus, para dejarlos en un formato que puedan consumir los algoritmos de procesamiento del lenguaje natural; 3) la anotación de estos textos, por medios automáticos o manuales, para asignar etiquetas apropiadas (por ejemplo, para decir si se ha proporcionado un recurso); y 4) entrenar el modelo que llevará a cabo las predicciones deseadas <sup>(54)</sup>. Todos estos pasos pueden beneficiarse del uso de conocimientos específicos del dominio de la aplicación como una forma de mejorar el tratamiento computacional de los textos analizados, como lo hacen los sistemas como CLAUDETTE, desarrollado en la Universidad de Bologna, que detecta cláusulas abusivas en documentos que describen o solicitan el cumplimiento de las políticas de privacidad de los sitios web <sup>(55)</sup>.

Sin embargo, incluso con la ayuda del conocimiento experto, los patrones encontrados por un sistema inteligente no son a menudo los que un humano observaría al realizar la misma tarea. Esto se debe a que la predicción normativa, propia del jurista, no se basa en la correlación empírica entre hechos, sino en la valoración de los factores (características) del caso a la luz de las normas jurídicas aplicables, o de los argumentos a favor o en contra de un determinado reclamo. Por lo tanto, es difícil para el jurista humano comprender cuáles son los eventos empíricos que el sistema consideró relevantes y las correlaciones utilizadas que resultaron en los *outputs* del sistema.

Un ejemplo puede ser encontrado en el trabajo de Verma *et al.* <sup>(56)</sup>, quienes desarrollaron un sistema con aproximadamente un 75% de precisión

---

(53) CONTISSA, Giuseppe *et al.*, "CLAUDETTE meets GDPR: automating the evaluation of privacy policies using Artificial Intelligence". Relatório Técnico para BEUC, 2018.

(54) ECKHART DE CASTILHO, Richard, "A Legal Perspective on Training Models for Natural Language Processing", LREC 2018, Miyazaki, Japão, 7-12 de maio de 2018.

(55) CONTISSA *et al.*, *ob. cit.*, 2018.

(56) VERMA, S. — PARTHASARATHY, A. — CHEN, D., "The Genealogy of ideology: predicting agreements and Persuasive memes in the U.S. Courts of Appeals", Proceedings of the Sixteenth International Conference on Artificial Intelligence and Law (ICAIL 2017), ACM Press, Londres. Nova York, 2017.

para predecir cuándo diferirían los jueces de los tribunales de apelaciones de EE. UU. La hipótesis sería que la divergencia estaría determinada por diferencias de carácter ideológico; sin embargo, los factores que más correlacionaron en la determinación del resultado fueron: (i) la posición en que se sintieron los jueces en el juicio; (ii) el tamaño de los votos, y (iii) el número de citas de precedentes. Obviamente, en el análisis legal, factores como estos son absolutamente irrelevantes y la predicción jurídica de divergencia está vinculada a la evaluación de condenas sobre tesis y principios jurídicos aceptados por los jueces analizados.

Así, la investigación en inteligencia artificial ha buscado el desarrollo de sistemas de inteligencia artificial inteligibles (*explainable artificial intelligence*: xAI), tanto mediante la producción de sistemas capaces de explicar de forma más sencilla el funcionamiento de otros sistemas como mediante la construcción de sistemas capaces de lograr un buen desempeño a partir de mecanismos internos que incorporen la representación del conocimiento jurídico.

### *III.3. Perspectivas de la inteligencia artificial aplicada en el derecho*

#### *III.3.a. Abordajes híbridos y el problema de la explicabilidad de las decisiones*

A pesar de sus éxitos, ambos paradigmas de inteligencia artificial encuentran límites a su aplicación en materia legal. En el caso de los sistemas basados en el conocimiento, el principal obstáculo es la dificultad que implica codificar los datos de un caso para que tengan una forma que el sistema pueda comprender y procesar, trabajo que hoy en día es mayoritariamente realizado por humanos. Para remediar esta dificultad, existen exitosas líneas de investigación orientadas al desarrollo de sistemas para la identificación automática de factores jurídicamente relevantes en textos legales, utilizando modelos espaciales vectoriales<sup>(57)</sup>, y de detección automática y estructuración de argumentos<sup>(58)</sup>. Si bien estas tecnologías aún necesitan alcanzar un mayor grado de madurez, ya contribuyen a reducir el trabajo humano requerido para la adopción de sistemas basados en el conocimiento jurídico.

De la misma forma que el uso de técnicas de aprendizaje automático puede servir para hacer factible el uso práctico de sistemas basados en el

---

(57) ASHLEY, Kevin — FALAKMASIR, Mohammad Hassan, “Utilizing Vector Space Models for identifying legal factors from text”, Proceedings of Jurix Conference, 2017.

(58) Ver PALAU, Raquel Mochales — MOENS, Marie-Francine, “Argumentation Mining: the detection, classification and structuring of Arguments in text”, Proceedings of ICAIL 2009, Barcelona, 2009, así como PATHAK, Arkanath — GOYAL, Pawan — BHOWMICK, Plaban, “A two-phase approach towards identifying argument structure in Natural Language”, Proceedings of the 3rd Workshop on Natural Language Processing Techniques for Educational Applications, Osaka, 2016, ps. 11-19.



conocimiento, las representaciones del dominio que brindan pueden ser útiles para explicar cómo la inteligencia artificial legal llega a sus conclusiones. En el ámbito del derecho, la IA inteligible es de especial importancia, ya que cualquier acto o decisión judicial o administrativa solo tiene validez jurídica en la medida en que pueda estar legalmente justificada <sup>(59)</sup>. En este dominio, el proceso y el contenido de la justificación son tan relevantes como el resultado.

Como ejemplos de proyectos que buscan dar explicaciones a las decisiones tomadas, por ellos mismos, en el ámbito legal, tenemos el proyecto NWO Forensic Science <sup>(60)</sup>, coordinado por Bart Verheij, que busca desarrollar sistemas para generar explicaciones sobre redes bayesianas y modelos probabilísticos de análisis de evidencia procesal, a través de escenarios y argumentos para que los modelos sean comprensibles para los abogados <sup>(61)</sup>.

La necesidad de explicación no es solo un requisito para que los sistemas inteligentes se utilicen en aplicaciones legales, sino que también es un requisito para que las predicciones de inteligencia artificial sean efectivas. Dado que una decisión judicial a menudo no es una consecuencia inmediata de “estimular los hechos” <sup>(62)</sup> que se puede detectar directamente a partir del texto, una predicción de decisiones judiciales no debe limitarse a aspectos empíricos, sino también tener en cuenta la dinámica argumentativa que debe reconstruirse para que un sistema pueda realizar una predicción adecuada. De esta forma, la incorporación de técnicas computacionales para el tratamiento de argumentos jurídicos a las técnicas de procesamiento del lenguaje jurídico existentes permitirá que los sistemas sean capaces de identificar los argumentos presentes en un documento legal y, a partir de ahí, predecir el resultado en un documento legal. Al mismo tiempo, más precisa y explicable para los observadores humanos.

El camino hacia la integración de modelos de predicción es, por un lado, el desarrollo de programas capaces de identificar factores o argumentos relevantes en documentos legales, de tal forma que puedan generar premisas

---

(59) En el derecho brasileño, la Constitución (art. 93, IX) y el Código de Proceso Civil (art. 489, II) tratan la justificación como elemento esencial de la sentencia. En Europa, la Corte Europea de Derechos Humanos reconoce la presencia de una justificación como esencial para que exista la posibilidad de recurso (“Fomin vs. Moldova” [2011], Nro. 36755/06, § 31) y para que las partes puedan tener sus quejas oídas, como parte de las garantías de un proceso justo (“Suominen vs. Finland” [2003] Nro .37801/97, § 37). Sobre el papel de las razones judiciales como mecanismo de control de sistemas automatizados, ver STRANDBURG, Katherine J., “Rulemaking and Inscrutable Automated Decision Tools”, *Columbia Law Review*, 7, v. 119, 2019, ps. 1851-1886.

(60) <http://www.ai.rug.nl/~verheij/nwofs/>.

(61) Ver VERHEIJ, Bart *et al.*, “Arguments, Scenarios and Probabilities: Connections Between Three Normative Frameworks for Evidential Reasoning”, *Law, Probability & Risk*, v. 15, 2016, ps. 35-70, así como VLEK, Charlotte S., “Building Bayesian Networks for Legal Evidence with Narratives: a Case Study Evaluation”, *Artificial Intelligence and Law*, 4, v. 22, 2014, ps. 375-421.

(62) ALETRAS *et al.*, *ob. cit.*, 2016.

para sistemas computacionales basados en lógicas legales y argumentales. Por otro lado, el desarrollo y perfeccionamiento de sistemas de argumentación y construcción de justificaciones basados en factores y argumentos para que puedan procesar argumentos y factores extraídos automáticamente (por aprendizaje automático) de textos y documentos legales. Con esto, es posible desarrollar modelos de argumentación que pueden, a partir de textos legales con poco o sin ningún tratamiento especial, ofrecer no solo predicciones empíricas, basadas en correlaciones estadísticas, sino también predicciones normativas, extraídas de la fuerza argumentativa del caso con relación a los precedentes, así como la construcción de justificaciones jurídicas basadas en argumentos inteligibles.

### III.3.b. Derecho computable

Otra línea de desarrollo de la inteligencia artificial legal es hacer que los estándares existentes sean computables. Dada la ubicuidad de los sistemas de IA, la imposibilidad de la supervisión humana de todas las posibles decisiones y acciones por parte de agentes digitales inteligentes ya no es factible, incluso en los casos en que estas decisiones se explican adecuadamente. Entonces, se vuelve imperativo que los sistemas inteligentes, especialmente aquellos empleados en las aplicaciones más sensibles legalmente, puedan ajustar sus acciones de acuerdo con las leyes.

Una primera solución a este problema es la adopción de técnicas de diseño de sistemas informáticos que aseguren su funcionamiento compatible con la ley. Este es el caso de los enfoques de protección de datos *by design*<sup>(63)</sup> e de técnicas por ella inspiradas, como *contestability by design*<sup>(64)</sup> o *ethics by design*<sup>(65)</sup>. Estos enfoques, en particular los destinados a proteger la privacidad, ya se adoptan con diversos grados de éxito, pero se enfrentan a dos dificultades principales: la posibilidad de conflictos entre los valores que

---

(63) El art. 25 del GDPR establece la necesidad de que los sistemas adopten, como estándar y desde el momento de su concepción, medidas para la protección de datos personales de sus usuarios. En Brasil, esa disposición encuentra amparo en el art. 46, §2º. Hay amplia literatura sobre el tema, por ejemplo: “Guidelines 4/2019 on Article 25 on Data Protection by Design and by Default”, European Data Protection Board, 2019, Bruselas (EDPB Guidelines).

(64) ALMADA, ob. cit.; con respecto a la incorporación, en el proyecto, de la posibilidad de contestar decisiones automatizadas, ver también BAYAMLIOGLU, Emre, “Contesting Automated Decisions”, *European Data Protection Law Review (EDPL)*, v. 4, 2018, ps. 433-446, y HILDEBRANDT, Mireille, “Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning”, *Theoretical Inquiries in Law*, 1, v. 20, 2019.

(65) DIGNUM, Virginia — BALDONI, Matteo — BAROGLIO, Cristina *et al.*, “Ethics by Design: Necessity or Curse?”, Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society. ACM, New Orleans - USA 2018, ps. 60-66.

estos enfoques intentan implementar <sup>(66)</sup> y la falta de recursos conceptuales y técnicos para anticipar problemas que puedan tener un impacto real <sup>(67)</sup>.

Frente a tales dificultades, necesario el desarrollo de sistemas inteligentes que sean capaces de procesar normas legales o morales <sup>(68)</sup>. A medida que los sistemas de inteligencia artificial adaptan su comportamiento a las circunstancias, dicho procesamiento no puede realizarse solo mediante la programación del regimiento, sino que requiere la construcción de representaciones formales de las normas éticas o legales relevantes para un dominio de aplicación determinado. A partir de estas representaciones, los sistemas inteligentes podrán razonar y aplicar reglas éticas y legales a la hora de elegir su curso de acción, considerando las peculiaridades del contexto en lugar de limitarse a determinaciones *ex ante*.

#### IV. CONCLUSIONES

La difusión del uso de sistemas inteligentes tiene el potencial de transformar la práctica de la abogacía, no solo al traer nuevos temas a ser considerados en las profesiones legales, sino también por la automatización de las actividades legales, comenzando por aquellas que involucran trabajo repetitivo y luego se vuelve más sofisticado. Estos cambios requerirán, además, un cambio en el perfil del profesional del derecho, que deberá ser capaz de hacer frente al nuevo escenario social y a las nuevas tecnologías.

En cuanto a las aplicaciones de la inteligencia artificial al derecho, este proceso de adaptación implica tanto el uso de herramientas de inteligencia artificial por parte de los profesionales del derecho como el uso de sus conocimientos para construir sistemas informáticos capaces de desarrollar su actividad al procesar datos de manera compatible con la ley. Hacer frente a los cambios tecnológicos significa sobre todo operar en equipos interdisciplinarios, que aprovechen las habilidades de abogados, informáticos y

---

(66) Ver, dentre outros, ALMADA, Marco — ATTUX, Romis, “Ethical design of social simulations”, Anais do 3º Workshop sobre Aspectos Sociais, Humanos e Econômicos de Software (WASHES 2018). SBC, Porto Alegre, 2018; WARDMAN, Jamie K. — LÖFSTEDT, Ragnar, “Anticipating or Accommodating to Public Concern? Risk Amplification and the Politics of Precaution Reexamined”, *Risk Analysis*, 9, v. 38, 2018, ps. 1802-1819; FEENBERG, Andrew, “Subversive rationalization: Technology, power, and democracy”, *Inquiry*, 3-4, v. 35, 1992, ps. 301-322; VAN DE POEL, Ibo, “Core Values and Value Conflicts in Cybersecurity: Beyond Privacy Versus Security”, en CHRISTEN, Markus — GORDIJN, Bert — LOI, Michele (orgs.), *The Ethics of Cybersecurity*. Springer International Publishing, Cham, 2020, ps. 45-71. (The International Library of Ethics, Law and Technology).

(67) Sobre as ferramentas e dificuldades envolvidas na antecipação de efeitos tecnológicos, ver BREY, Philip A. E., “Anticipatory Ethics for Emerging Technologies”, *NanoEthics*, 1, v. 6, 2012, ps. 1-13; FLORIDI, Luciano — STRAIT, Andrew, “Ethical Foresight Analysis: What it is and Why it is Needed?”, *Minds and Machines*, 1, v. 30, 2020, ps. 77-97.

(68) Ver, en este sentido, el proyecto europeo CompuLaw: <https://cordis.europa.eu/project/id/833647>. A partir del 2019, la Facultad de Derecho de la Universidad de São Paulo ofrece disciplinas de posgrado sobre derecho computable.

otros profesionales para construir sistemas inteligentes que tengan efectos positivos y protejan los derechos e intereses legalmente protegidos en juego.

## V. REFERENCIAS BIBLIOGRÁFICAS

- CORMEN, Thomas H. — LEISERSON, Charles E. — RIVEST, Ronald L. — STEIN, Clifford, “Introduction to algorithms”, The MIT Press, Cambridge, 2009, 3ª ed., ps. 5-6.
- CAPLAN, Robyn — DONOVAN, Joan — HANSON, Lauren — MATTHEWS, Jeanna, “Algorithmic accountability: a primer”, *Data & Society*, 2018. Disponible en <https://datasociety.net/output/algorithmic-accountability-a-primer/>. Accedido en: 26/08/2018.
- LARSON, Christina, “Who needs democracy when you have data?”, *MIT Technology Review*, 20/08/2018. Disponible en <https://www.technologyreview.com/s/611815/who-needs-democracy-when-you-have-data/>. Accedido em 27/08/2018.
- GOMES, Rodrigo Dias de Pinho, “Big Data: desafios à tutela da pessoa humana na sociedade da informação”, *Lumen Juris*, Río de Janeiro, 2017, ps. 50-51.
- ONG, Rebecca, “Recognition of the right to privacy on the Internet in China”, *International Data Privacy Law*, 3, v. 1, 2011, ps. 172-179. Disponible en <https://doi.org/10.1093/idpl/ipr008>. Accedido en: 10/10/2018.
- SINGH, Jatinder — WALDEN, Ian — CROWCROFT, Jon — BACON, Jean, “Responsibility and machine learning: part of a process” SSRN, 2016, p. 7. Disponible en <https://ssrn.com/abstract=2860048>. Accedido en: 18/10/2018.
- BURRELL, Jenna, “How the machine ‘thinks’: Understanding opacity in machine learning algorithms”, *Big Data & Society*, 1, v. 3, 2016, ps. 1-12.
- LEHR, David — OHM, Paul, “Playing with the Data: What Legal Scholars Should Learn About Machine Learning”, *Davis Law Review*, v. 51, University of California, ps. 653-717, 2017.
- PASQUALE, Frank, “The Black Box Society: the secret algorithms that control money and information”, Harvard University Press, Cambridge, 2015, p. 2.
- FJELD, Jessica *et al.* “Principled Artificial Intelligence: A Map of Ethical and Rights-Based Approaches”. Relatório técnico, Berkman Klein Center For Internet & Society, septiembre de 2018.
- JOBIN, Anna — IENCA, Marcello — VAYENA, Effy, “The global landscape of AI ethics guidelines”, *Nature Machine Intelligence*, v. 1, 2019, ps. 389-399.
- Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions - Coordinated Plan on Artificial Intelligence (COM(2018) 795 final).

- BURGESS, Peter — FLORIDI, Luciano — POLS, Aurélie — van den HOVEN, Jeroen, “Towards a digital ethics”, Ethics Advisory Group, Bruxelas, 2018.
- KOENE, Ansgar, *et al.*, “A governance framework for algorithmic accountability and transparency”, Parlamento Europeo, Bruxelas, 2019.
- FLORIDI, Luciano, *et al.*, “AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations”. *Minds and Machines*, v. 28, 2018, ps. 689-707.
- WAGNER, Ben, “Ethics as an Escape from Regulation: From ethics-washing to ethics-shopping?”, en HILDEBRANDT, Mireille (ed.), *Being Profiling. Cogitas ergo sum*, Amsterdam University Press, Amsterdam, 2018.
- BIETTI, Elettra, “From ethics washing to ethics bashing: a view on tech ethics from within moral philosophy”, en *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT\* '20)*, Association for Computing Machinery, Barcelona, Spain, 2020, ps. 210-219.
- MARANHÃO, Juliano — COUTINHO, Diogo R., “Melhor investir do que regular”, *Correio Braziliense*, 25/03/2019.
- SWANDA, Gus, “The Deficiencies of a Westphalian Model for Cyberspace: A Case Study of South Korean Cyber Security”, *International Journal of Korean Unification Studies*, 2, v. 25, 2016.
- BRYSON, Joanna J. — THEODOROU, Andreas, “How Society Can Maintain Human-Centric Artificial Intelligence”, en TOIVONEN, Marja — SAARI, Eveliina (Orgs.), *Human-Centered Digitalization and Services*, Springer Singapore, 2019, Singapore, ps. 305-323.
- KAMINSKI, Margot, “The Right to Explanation, Explained”, *Berkeley Technology Law Journal*, 1, v. 34, 2019.
- ALMADA, Marco, “Human intervention in automated decision-making: Toward the construction of contestable systems”, *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law (ICAIL 2019)*, ACM Press, Montreal, Canadá. Nova York, 2019, seção 3.
- RUSSELL, Stuart J. — NORVIG, Peter, “Artificial Intelligence: a modern approach”, Prentice Hall, Upper Saddle River, 3ª ed., 2010, p. 234.
- OLIVEIRA, Rafael Brito de, “Utilização de Ontologias para Busca em Base de Dados de Acórdãos do STF. Dissertação (Maestría en Ciencias de la Computación)”, Universidade de São Paulo, São Paulo, 2017.
- MEDVEDEVA, Masha — VOLS, Michel — WIELING, Martijn, “Using machine learning to predict decisions of the European Court of Human Rights”, *Artificial Intelligence and Law*, 2019.

- KUZNIACKI, Blazej, “The Artificial Intelligence Tax Treaty Assistant: Decoding the Principal Purpose Test”, *Bulletin for International Taxation*, 9, v. 72, 2018.
- HAYKIN, Simon, “Neural Networks and Learning Machines”, Prentice Hall, 2008, p. 34.
- HILPINEN, Risto — MCNAMARA, Paul, “Deontic Logic: a historical survey and introduction”, en GABBAY, Dov *et al.* (orgs.), *Handbook of Deontic Logic and Normative Systems*. College Publications, Londres, 2013.
- BENCH-CAPON, Trevor J. M. — PRAKKEN, Henry, “Using argument schemes for hypothetical reasoning in law”, *Artificial Intelligence and Law*, v. 18, 2010, ps. 153-174.
- BEX, Floris J. *et al.*, “Towards a formal account of reasoning about evidence: argumentation schemes and generalisations”, *Artificial Intelligence and Law*, v. 11, 2003, ps. 125-165; e VERHEIJ, Bart, “Dialectical argumentation with argumentation schemes: an approach to legal logic”, *Artificial Intelligence and Law*, v. 11, 2003, ps. 167-195.
- BENCH-CAPON, Trevor J. M., “The missing link revisited: the role of teleology in representing legal argument”, *Artificial Intelligence and Law*, v. 10, 2002, ps. 79-9; y PRAKKEN, Henry — SARTOR, Giovanni, “A dialectical model of assessing conflicting arguments in legal reasoning”, *Artificial Intelligence and Law*, v. 4, 1996, ps. 331-368.
- HAGE, Jaap C., “Comparing alternatives in the law”, *Artificial Intelligence and Law*, v. 12, 2005, ps. 181-225; SARTOR, Giovanni, “The logic of proportionality: reasoning with non-numerical magnitudes”, *German Law Journal*, v. 14, 2013; PRAKKEN, Henry, *et al.*, “A formalisation of argumentation schemes for legal case-based reasoning in ASPIC+”, *Journal of Logic and Computation*, 2013; y HORTY, John F., “Rules and reasons in the theory of precedent”, *Legal Theory*, v. 17, 2011, ps. 1-33.
- MARANHÃO, Juliano — SARTOR, Giovanni, “Value Assessment and Revision in Legal Interpretation”, *Proceedings of the 17th edition of the International Conference on Artificial Intelligence and Law (ICAIL 2019)*, ACM Press, Montreal, Canadá. Nova York, 2019; y HORTHY, John F. — BENCH-CAPON, Trevor J., “A factor-based definition of precedential constraint”, *Artificial Intelligence and Law*, v. 20, 2012, ps. 181-214.
- ASHLEY, Kevin D., “Modeling Legal Argument: Reasoning with Cases and Hypotheticals”, MIT Press, Cambridge, 1990.
- ALEVEN, Vincent, “Using background knowledge in case-based legal reasoning: a computational model and an intelligent learning environment”, *Artificial Intelligence*, 150, 2003, ps. 183 ss. KEENEY, Ralph L. — RAIFFA, Howard, “Decisions with Multiple Objectives”, Wiley, New York, 1976.

- SARTOR, Giovanni, "Fundamental legal concepts: A formal and teleological characterisation", *Artificial Intelligence and Law*, v. 21, 2010, ps. 101-142, 2010, y ob. cit., 2013.
- GRABMAIR, Matthias, "Predicting trade secret case outcomes using argument schemes and learned quantitative value effect tradeoffs", *Proceedings of the 16th edition of the International Conference on Artificial Intelligence and Law (ICAIL 2017)*, ACM Press, Londres, Nova York, 2017, ps. 89-98.
- ALETRAS, Nikolaos *et al.*, "Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective", *PeerJ Computer Science*, v. 2, e93, 2016.
- CONTISSA, Giuseppe *et al.*, "CLAUDETTE meets GDPR: automating the evaluation of privacy policies using Artificial Intelligence". Relatório Técnico para BEUC, 2018.
- ECKHART DE CASTILHO, Richard, "A Legal Perspective on Training Models for Natural Language Processing", *LREC 2018*, Miyazaki, Japão, 7-12 de mayo de 2018.
- VERMA, S. — PARTHASARATHY, A. — CHEN, D., "The Genealogy of ideology: predicting agreements and Persuasive memes in the U.S. Courts of Appeals", *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Law (ICAIL 2017)*, ACM Press, Londres. Nova York, 2017.
- ASHLEY, Kevin — FALAKMASIR, Mohammad Hassan, "Utilizing Vector Space Models for identifying legal factors from text", *Proceedings of Jurix Conference*, 2017.
- PALAU, Raquel Mochales — MOENS, Marie-Francine, "Argumentation Mining: the detection, classification and structuring of Arguments in text", *Proceedings of ICAIL 2009*, Barcelona, 2009.
- PATHAK, Arkanath — GOYAL, Pawan — BHOWMICK, Plaban, "A two-phase approach towards identifying argument structure in Natural Language", *Proceedings of the 3rd Workshop on Natural Language Processing Techniques for Educational Applications*, Osaka, 2016, ps. 11-19.
- STRANDBURG, Katherine J., "Rulemaking and Inscrutable Automated Decision Tools", *Columbia Law Review*, 7, v. 119, 2019, ps. 1851-1886.
- VERHEIJ, Bart *et al.*, "Arguments, Scenarios and Probabilities: Connections Between Three Normative Frameworks for Evidential Reasoning", *Law, Probability & Risk*, v. 15, 2016, ps. 35-70.
- VLEK, Charlotte S. "Building Bayesian Networks for Legal Evidence with Narratives: a Case Study Evaluation", *Artificial Intelligence and Law*, 4, v. 22, 2014, ps. 375-421.
- BAYAMLIOGLU, Emre, "Contesting Automated Decisions", *European Data Protection Law Review (EDPL)*, v. 4, 2018, ps. 433-446.

- HILDEBRANDT, Mireille, “Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning”, *Theoretical Inquiries in Law*, 1, v. 20, 2019.
- DIGNUM, Virginia — BALDONI, Matteo — BAROGLIO, Cristina *et al.*, “Ethics by Design: Necessity or Curse?”, *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, ACM, New Orleans, 2018, ps. 60-66.
- ALMADA, Marco — ATTUX, Romis, “Ethical design of social simulations”, *Anais do 3º Workshop sobre Aspectos Sociais, Humanos e Econômicos de Software (WASHES 2018)*, SBC, Porto Alegre, 2018.
- WARDMAN, Jamie K. — LÖFSTEDT, Ragnar, “Anticipating or Accommodating to Public Concern? Risk Amplification and the Politics of Precaution Reexamined”, *Risk Analysis*, 9, v. 38, 2018, ps. 1802-1819.
- FEENBERG, Andrew, “Subversive rationalization: Technology, power, and democracy”, *Inquiry*, 3-4, v. 35, 1992, ps. 301-322.
- VAN DE POEL, Ibo, “Core Values and Value Conflicts in Cybersecurity: Beyond Privacy Versus Security”, en CHRISTEN, Markus — GORDIJN, Bert — LOI, Michele (orgs.), *The Ethics of Cybersecurity*, Springer International Publishing, Cham, 2020, ps. 45-71 (The International Library of Ethics, Law and Technology).
- BREY, Philip A. E., “Anticipatory Ethics for Emerging Technologies”, *NanoEthics*, 1, v. 6, 2012, ps. 1-13. FLORIDI, Luciano — STRAIT, Andrew, “Ethical Foresight Analysis: What it is and Why it is Needed?” *Minds and Machines*, 1, v. 30, 2020, ps. 77-97.