

ALGORITMOS PREDICTIVOS AL SERVICIO DE LA JUSTICIA: ¿UNA NUEVA FORMA DE MINIMIZAR EL RIESGO Y LA INCERTIDUMBRE? ¹

PREDICTIVE ALGORITHMS AT THE SERVICE OF JUSTICE: A NEW WAY TO MINIMIZE RISK AND UNCERTAINTY?

*Nuria Bellosó Matin*²

UNIVERSIDAD DE BURGOS

Resumen:

Las aportaciones del Profesor Raffaele de Giorgi en relación a las categorías de seguridad y riesgo y del Profesor Jorge Douglas Price en relación a la decisión judicial iluminan una mejor comprensión de la problemática que actualmente atenaza al concepto de Derecho y a los sistemas de justicia. El ciudadano actual exige que el Derecho sea capaz de minimizar la incertidumbre y el peligro y, si es posible, neutralizar el riesgo. En la consecución de esta minimización del riesgo, los algoritmos predictivos han adquirido protagonismo en cuanto a su capacidad de optimizar las políticas públicas en la medida en que se aplican para detectar problemas antes de que ocurran, lo que facilita que tengan múltiples aplicaciones en el campo de la gestión pública. Una de las áreas en que más se está utilizando es el sistema de justicia penal. Sin embargo, el uso de algoritmos predictivos en el sistema de justicia penal ha sido ampliamente criticado por razones de equidad, con afirmaciones de que pueden exhibir prejuicios raciales y, por lo tanto, perpetuar las desigualdades sociales preexistentes. Además, investigaciones recientes han puesto en tela de juicio la eficacia de un 'software' concebido para prevenir la reincidencia demostrando las dudas con la que pueden hacerse predicciones. Se analizarán algunas experiencias en el Derecho comparado y se cuestionará si la justicia predictiva o algorítmica no corre el peligro de convertirse en una versión del positivismo de nuevo cuño, a modo de una dictadura de los algoritmos.

Palabras llave:

inteligencia artificial, algoritmos, justicia predictiva, riesgo, decisión judicial

Abstract:

The contributions of Professor Raffaele de Giorgi in relation to the categories of security and risk and Professor Jorge Douglas Price in relation to the judicial decision illuminate a better understanding of the problem that currently grips the concept of law and justice systems. The current citizen demands that the Law be able to minimize uncertainty and danger and, if possible, neutralize the risk. In achieving this risk minimization, predictive algorithms have gained prominence in terms of their ability to optimize public policies to the extent that they are applied to detect problems before they occur, which facilitates having multiple applications in the field of public management. One of the areas where most is being used is the criminal justice system. However, the use of predictive algorithms in the criminal justice system has been widely criticized for reasons of fairness, with claims that they may exhibit racial prejudice and, therefore, perpetuate preexisting social inequalities. In addition, recent research has questioned the effectiveness of a 'software' designed to prevent recidivism by demonstrating the doubts with which predictions can be made. Some experiences in comparative law will be analyzed and it will be questioned whether predictive or algorithmic justice does not run the risk of becoming a version of positivism of new stamp, as a dictatorship of algorithms.

Key words:

artificial intelligence, algorithms, predictive justice, risk, judicial decision

1 INTRODUCCIÓN

¹ El presente trabajo se ha realizado en el marco del proyecto de investigación “*La inteligencia artificial jurídica (IAJ)*”. (Código: RTI2018-096601-B-I00).

² Catedrática de Filosofía del Derecho en la Universidad de Burgos, España. Directora del Departamento de Derecho Público.

El Derecho del siglo XXI tiene ante sí numerosos retos. Uno de estos desafíos es el de proporcionar el mayor nivel de seguridad posible. Si bien entre las funciones del Derecho siempre ha estado presente la seguridad, el ciudadano actual exige que el derecho sea capaz de minimizar la incertidumbre y el peligro y, si es posible, neutralizar el riesgo. Se hacen cálculos de probabilidades y predicciones para formular hipótesis sobre qué resultados son más plausibles según los diversos espectros que se presenten. Bien sea en la política (en la que se formulan hipótesis sobre posibles pactos entre distintos partidos políticos), en la economía (donde la incertidumbre no gusta a los mercados) o en otros campos, como el Derecho. Actualmente se precisan nuevos marcos teóricos y paradigmas metodológicos que den respuesta a las nuevas situaciones. La crisis del positivismo, la globalización, la hiperinflación de las fuentes del Derecho y otras, acaban incidiendo también en el Derecho. Desde la ficción jurídica de la pirámide kelseniana hasta el concepto del Derecho como una novela escrita en cadena, postulada por Dworkin- el Derecho y la teoría de la decisión judicial se transforman y evolucionan al compás de los progresos tecnológicos.

El criminólogo y médico italiano Cesare Lombroso, en su libro *El hombre delincuente*, publicado en 1896, sostenía que las conductas delictivas son determinadas por predisposiciones fisiológicas, que a menudo se revelan también externamente en la configuración del cráneo. A través de la observación de algunas características anatómicas, por lo tanto, sería posible desenmascarar a un futuro criminal. Lombroso era un positivista, es decir, consideraba que la respuesta a todas las preguntas de la especie humana se encontraba en la ciencia. Actualmente vivimos una era de positivismo digital y la fe se pone en las oportunidades ofrecidas por el *big data* y los algoritmos. Sin embargo, el optimismo tecnológico choca con la pesadilla de un futuro similar al descrito por George Orwell en su obra *'1984'*, con una "Policía del Pensamiento" capaz de castigar no solo los delitos, sino también las intenciones. La llegada de la Inteligencia artificial (IA) ha dado lugar a cambios relevantes en cuanto a la forma de "hacer" y de "decir" el Derecho. Y ello porque si bien constituyen una ayuda para los profesionales y operadores del derecho, también han propiciado nuevos problemas. En concreto, la teoría de la decisión judicial se ha visto afectada de forma clara por estos avances en las relaciones entre Derecho e inteligencia artificial (en adelante, IA), dando lugar a lo que se denomina una justicia algorítmica o una justicia predictiva.

Las aplicaciones de inteligencia artificial diseñadas para mejorar la investigación legal con el fin de hacer más rápida y eficiente la búsqueda y procesamiento de información jurídica están ya presentes en las Oficinas Judiciales y en la mayoría de los despachos de abogados (SOLAR CAYÓN, 2019). Actualmente operan robots *lawyers* que conversan con los humanos con la

finalidad de asesorarles a resolver problemas legales simples (DoNotPay y Lee&Aly). Pero la aplicación de la IA en el sector legal va más allá. Hay programas diseñados por empresas que trabajan para predecir el resultado de decisiones judiciales utilizando herramientas tecnológicas denominadas justicia predictiva tales como Watson/Ross de IBM y Blue J. Legal. También elaboran perfiles de jueces y magistrados a partir del estudio histórico de su actividad jurisdiccional (por ejemplo, Ravel Law de LexisNexis y LegalOpticis).

En cualquier caso, tanto las partes implicadas en un proceso judicial como sus abogados defensores, desean evitar o, al menos, minimizar el “riesgo” de no saber si una resolución judicial será estimada o no. Desean eliminar la incertidumbre legal antes de iniciar un proceso judicial que seguramente resultará costoso, tanto económicamente como emocionalmente. Asimismo, en las diversas disciplinas jurídicas (como la civil, la mercantil, la laboral, la contencioso-administrativa) la IA puede facilitar la armonización de la jurisprudencia en los asuntos judiciales donde haya que establecer escalas o baremos (por ejemplo, procesos civiles relativos a daños corporales o indemnizaciones compensatorias por divorcio; o procesos laborales referentes a despidos con indemnización) además de promover la resolución de conflictos online o servicios judiciales electrónicos (ODR). En todas estas posibilidades se trasluce el interés por neutralizar el riesgo. A la IA se la considera como un potente instrumento para contribuir a minimizar esa incertidumbre.

Al igual que se intentan evitar los riesgos de catástrofes naturales o epidemias, la IA aplicada al sistema de justicia puede evitar no sólo errores humanos sino subjetividad y prejuicios. Así, se multiplican opiniones que sustentan que: “La administración de justicia es una de las funciones que más fácilmente se puede informatizar. La función de un juez es aplicar un conjunto de reglas, leyes y normas, a un conjunto de datos que son los hechos que se deben juzgar. Un programa de ordenador hace básicamente lo mismo.”. Y las ventajas también se elogian: “El sistema sería mucho más justo ya que no está sujeto a la subjetividad del juzgador. Sería mucho más ágil y rápido. Se paliaría parte del colapso judicial actual”. Los errores se disculpan: “¿Podrían haber errores?, probablemente, pero también los hay ahora, con la diferencia de que en ese caso, los errores se retroalimentarían de forma inmediata de forma que no se reprodujeran.”

Ahora bien, al igual que en nuestro modelo de sociedad, la incertidumbre y el riesgo forman parte de la misma, ¿es posible eliminarlas en el ámbito legal recurriendo a la justicia predictiva? Si bien es cierto que el cálculo de escalas o baremos parece mejorar cuando se combina con técnicas informáticas de datos con un enfoque de probabilidades, sin embargo no está demostrado que el aprendizaje automático (*machine learning*) que desarrolla la IA sea

totalmente fiable y sea capaz de diferenciar lo importante de lo accesorio. O más grave aún, hay un riesgo de que afecte a la imparcialidad de tales métodos seleccionando únicamente unos datos –y no otros- de forma que excluya muestras de las decisiones judiciales que no encajen en las correlaciones de las secuencias lingüísticas que el *machine learning* haya establecido. Todo ello sin obviar que los ciudadanos tiene derecho de acceso a la justicia y derecho a un recursos judicial efectivo (artículos 6 y 13 de Convenio Europeo de Derechos humanos), además del derecho de las personas físicas a negarse a ser objeto de una decisión basada exclusivamente en un procedimiento automatizado –con ciertas excepciones- (artículo 22 del reglamento General de Protección de Datos) (OLIVA LEÓN, 2018).

En el presente trabajo se trabajara en el análisis de si la justicia predictiva es capaz de neutralizar o al menos, minimizar el riesgo característico de las sociedades actuales. Para ello, hay que poner de relieve el mérito de los ilustres iusfilósofos Raffaele de Giorgi -de la Universidad del Salento (Italia)- y de Jorge Douglas Price – de la Universidad Nacional de Comahué (Argentina)- en la medida en que sus teorías y algunas de las categorías que han trabajado (tales como riesgo / peligro / incertidumbre / seguridad además de otras tales como democracia, la justicia y la teoría de la decisión judicial) (DOUGLAS PRICE, 2012) han sido esenciales para poder proyectar sus enseñanzas sobre la posible viabilidad - de un sistema la justicia algorítmica.

Se iniciará por un recorrido por esas categorías y principios que permean nuestro modelo de sociedad –y nuestros modelos de sistemas de justicia-. Seguidamente, se analizarán algunas de las propuestas enunciadas en la Carta ética europea sobre el uso de la inteligencia artificial en los sistemas, adoptada por la European Commission for the Efficiency of Justice (CEPEJ). Después, se revisarán algunas de las experiencias de aplicación de la justicia predictiva al ámbito penal, tales como el contexto norteamericano -con el caso Loomis-, en el contexto italiano –con una policía predictiva-, en el contexto inglés y, por último, el español). En el estudio se pondrá de manifiesto que, por un lado, los defensores del uso de algoritmos reivindican que un problema que puede predecirse podrá prevenirse con mayor facilidad, mientras que los detractores señalan que no está claro cómo se están utilizando exactamente la enorme cantidad de datos personales que los algoritmos necesitan para realizar sus predicciones. En este trabajo se analizará cómo esa justicia predictiva se está convirtiendo en un instrumento de auxilio a la tarea de juzgar pero que puede poner en peligro algunas de las conquistas de las democracias actuales tales como la afectación de algunos derechos fundamentales (protección de datos, no discriminación e igualdad). Los derechos humanos, como factores de innovación o de cambio social, están ligados a una evolución imprevisible que trabaja con lo *improbable* (LUHMANN, 1998; MARTÍNEZ GARCÍA, 2019). Hay que diferenciar sistemas de ayuda y auxilio a la labor judicial de sistemas predictivos

de justicia o de jueces robots. También es diferente el programa informático que puede predecir delitos e incidencias de forma fiable (pre-crimen) y que es utilizado por la Policía, donde el programa es capaz de configurar un mapa predictivo según hechos que se sitúan en espacio y tiempo gracias a un algoritmo. Por último, en las reflexiones finales, se planteará la duda que se cierne sobre el cada vez mayor protagonismo de la justicia algorítmica en relación a si aunque pensábamos que estábamos en un tiempo de postpositivismo, la exactitud y rigor matemático que acompaña a la imperatividad de los algoritmos, no puede derivar en una nueva versión del positivismo apoyado en la dictadura de los algoritmos.

2 DEL RIESGO DE LAS SOCIEDADES AL RIESGO EN LA TEORÍA DE LA DECISIÓN JUDICIAL

Según De Giorgi, la sociedad se define por tres elementos básicos que intentan legitimar la concepción de sociedad concebida en cada época, tales como: la verdad, los valores y la seguridad. El Derecho no es el regulador de la sociedad pero sí de los hechos sociales que determinan qué es el Derecho, constituyéndose así en una herramienta esencial en la sociedad compleja para poder dar forma a los hechos sociales, establecidos a través de la comunicación, y decir qué es Derecho. Ahora bien, ese propósito de alcanzar la seguridad no debe llevar a oscurecer los otros dos elementos como son la verdad y los valores. De ahí que la justicia predictiva deba estar atenta a que, no por conseguir la seguridad, obvie la verdad y los valores que, necesariamente, deben formar parte de la sociedad.

Hay diversos ámbitos relativos a la toma de decisiones que, adoptadas en condiciones de incertidumbre sobre el futuro, y si se verifican, pueden conllevar consecuencias dañosas. Estos daños se pueden imputar a decisiones que, si hubieran sido diversas, hubieran posibilitado que se hubieran evitado las consecuencias.

En 1986, Ulrich Beck publicó su obra *La sociedad del riesgo*. Por su parte, Luhmann - a partir de las críticas a Beck- escribió en 1991 una *Sociología del riesgo*. El concepto clave que acompaña a las reflexiones de Beck es el de consecuencias no esperadas de la acción. Así, cada innovación científico técnica no sólo debe ser vista desde el ángulo de sus metas manifiestas, sino de sus consecuencias latentes. Para los críticos de Beck, todo parece indicar que antaño los riesgos eran mucho mayores que hoy. Mientras que antes la humanidad quedaba literalmente sometida a los riesgos, ahora –gracias al cálculo de probabilidades y al creciente dominio de la naturaleza posibilitado por el desarrollo tecnológico– podemos protegernos de diversas maneras. Ahora bien, qué es el riesgo? ¿Por qué influye tan notablemente en las decisiones?

Por riesgo, siguiendo al profesor De Giorgi, podemos entender aquello que puede ser imputado como consecuencia de una decisión. De Giorgi define así el riesgo:

O risco não é nem uma condição existencial do homem, muito menos uma categoria ontológica da sociedade moderna, e tampouco o resultado per verso do trabalho da característica das decisões, uma modalidade da construção de estruturas através do necessário tratamento das contingências. É uma modalidade da relação com o futuro; é uma forma de determinação das indeterminações segundo a diferença de probabilidade/improbabilidade (DE GIORGI, 1998: 52-54).

El riesgo no debe de considerarse como un concepto opuesto al de seguridad. En el contexto general de una inseguridad con respecto al futuro y de un daño posible, se puede hablar de riesgo cuando un daño cualquiera viene imputado a una decisión, es decir, cuando este daño es tratado como consecuencia de una decisión. El concepto opuesto sería, entonces, el de peligro, que se aplica cuando daños posibles se imputan a elementos externos (DE GIORGI, 2014: 26). La idea central del riesgo es que se trata de una elección entre formas de actuación arriesgadas, es decir, consiste en una distribución de futuros beneficios o de futuros daños bajo el aspecto de una inseguridad presente. Incluso, el hecho de intentar adoptar como objetivo el evitar absolutamente el riesgo es, a su vez, arriesgado. Una parte de los riesgos difícilmente se podrán calcular racionalmente. Sin embargo, sí hay espacios para cálculos racionales por lo que los eventuales resultados que deriven de las decisiones que adopten los sujetos, las instituciones y las racionalidades encajarían en la presunción de racionalidad del cálculo. La dificultad radica en que no hay estándares de racionalidad comunes, de manera que si, ante la incertidumbre de que para evitar que se produzcan futuros eventos dañosos, optamos por otras políticas que prevemos van a evitar esos riesgos, el resultado es que se abrirán otros riesgos posibles.

Serrano Moreno también se ha apoyado en algunos binomios conceptuales para ayudar a comprender el significado de riesgo. “Riesgo es la contingencia de un daño. Contingente es aquello que puede ser y puede no ser. Contingencia se opone a imposibilidad y necesidad, porque contingente es aquello que no es ni imposible ni necesario. De la misma manera, los antónimos de riesgo son seguridad y certeza.” Serrano Moreno sostenía que estos binomios contingencia/necesidad y, sobre todo, riesgo/seguridad son los que elegiría cualquier jurista especializado para comenzar a hablar del derecho del riesgo en las sociedades contemporáneas. Sin embargo, siguiendo a De Giorgi, optaba por un camino más arriesgado y construyó el concepto de riesgo a partir de la diferencia riesgo / peligro:

“Como concepto el riesgo adopta diferente significado si lo construimos con la diferencia riesgo/seguridad o con la diferencia riesgo/peligro. Con la primera de esas diferencias, riesgo es la ausencia de seguridad. Sin embargo, si lo construimos con la segunda, la idea de riesgo apunta hacia la voluntad, la decisión y la responsabilidad.

Peligro es todo lo malo que puede pasar con independencia de las decisiones que uno tome. Riesgo es todo lo que puede salir mal, después de haber decidido. Lo que puede pasar no depende de la decisión, lo que puede resultar sí” (SERRANO MORENO, 2010: 184).

Serrano Moreno clarifica que “Riesgo es lo imputable a otro o a uno mismo, mientras que peligro es una amenaza que proviene del exterior”. Como acertadamente afirma Martínez García, no se trata de eliminarlos, sino de detectarlos, configurarlos y aprender a manejarlos. “La gestión de los riesgos puede implicar transferirlos, hacerlos repercutir en otros puntos, transformarlos, concentrarlos o distribuirlos, descargarlos, compensarlos. Trabajar con riesgos activa y exige toda una dinámica social” (MARTÍNEZ GARCÍA, 2010).

En el lenguaje común, el factor distintivo es la decisión. Los riesgos se refieren a daños que se presentan como resultado de una decisión y que no se producirían si la decisión hubiera sido otra. Esto mismo es trasladable a la decisión judicial: ¿hay alguna forma de evitar o minimizar los riesgos? ¿Cómo dotar a los ciudadanos de ese mayor grado de seguridad que continuamente demandan de los poderes públicos? Pero la idea de seguridad es un concepto ilusorio que enmascara el problema real. “No existe seguridad que inmunice del riesgo” (DE GIORGI, 2014: 21).

Como advierte de Giorgi, “el problema del control de las indeterminaciones siempre preocupó a las sociedades” que de maneras diversas, han procurado soluciones adecuadas en sus respectivos niveles evolutivos. La adivinación, por ejemplo, era una técnica que permitía tratar las indeterminaciones, reconduciendo la escasez de informaciones, el tabú es también una técnica que sirve para evitar el daño que pueda derivar de las indeterminaciones; también el pecado, que representa el desvío de la razón frente al orden y, al mismo tiempo, precariedad de la relación de la razón con el orden. (DE GIORGI, 2014:49).

Ahora, la justicia predictiva se convierte en una forma renovada de adivinación.

Este tipo de justicia predictiva trabaja con unos principios renovados:

- a) Prevención/precaución. Estos principios constituyen dos funciones distintas, con un denominador común: la necesidad de actuar antes de la producción de un daño. En una síntesis radical, se desempeñan sobre distintos tipos de riesgos. Al riesgo potencial, la precaución. Al riesgo verificado corresponde la prevención. No podría utilizarse sistemas de IA si se dan las siguientes circunstancias: i) la ausencia de una trazabilidad algorítmica; ii) la imposibilidad de asegurar un “botón de apagado” o un mecanismo seguro de contención de la IA; iii) cuando en cualquier fase –diseño, desarrollo o aplicación– se advierta que el sistema se basa en distinciones que violan el

principio de igualdad y no discriminación. Aquí esto operaría como una suerte de “categoría sospechosa algorítmica”.

b) Autodeterminación algorítmica: se debe de garantizar la autodeterminación humana frente al uso de algoritmos inteligentes.

c) Transparencia algorítmica y principio de imparcialidad del validador: la IA debe ser “transparente” en sus decisiones.

d) Principio de no discriminación algorítmica. Impedir que la IA procese la información o los datos bajo sesgos o distinciones

3 LOS POSIBLES USOS DE LA INTELIGENCIA ARTIFICIAL EN LOS SISTEMAS DE JUSTICIA ³

El análisis sobre la inteligencia artificial admite varios enfoques. Un primer enfoque es el de qué modo la IA influye en el Derecho, en qué y cómo está incidiendo en los juristas (en el pensamiento, la metodología y la práctica jurídica). Un segundo enfoque es el de los problemas éticos y jurídicos –y, por tanto, en los Derechos humanos- que plantea la IA. Por tanto, la interrelación entre Derecho e Inteligencia Artificial es bidireccional. En relación a esta segunda perspectiva -la IA como materia a regular jurídicamente- son numerosos los interrogantes jurídicos que se plantean: ¿Son nuestros marcos jurídicos en vigor adecuados para hacer frente a los avances de la robótica? ¿Pueden regularse las nuevas tecnologías robóticas, en particular si se caracterizan por un comportamiento autónomo cada vez mayor, dentro de los marcos legales y éticos existentes o, en caso contrario, deben hacerse más genéricas las normas jurídicas existentes para que incluyan también a las tecnologías robóticas?, o más bien ¿debemos aspirar a una disciplina jurídica autónoma y propia para los robots? A estos interrogantes se pueden añadir otros tales como si los robots pueden ser agente morales, si existe un deber de “cuidado” hacia ellos, u otras tales como: ¿cuál es el grado de autonomía que aspiramos que alcance un robot inteligente? ¿En qué situaciones se podría justificar que un robot no respetara la voluntad humana? Los robots ¿deben crearse con aspecto de personas? ¿Cómo deben relacionarse los robots con seres humanos que sean menos inteligentes que ellos?.

Son pues varias las cuestiones merecen ser objeto de atención, tales como la incidencia de la IA en los Derechos humanos y la de qué límites establece y diseña el Derecho con respecto a la IA (a sus aplicaciones y a su responsabilidad en caso de que provoque daños), es decir,

³ El Consejo de Europa adopta la primera Carta Ética Europea sobre el uso de la inteligencia artificial en los sistemas judiciales (5/12/2018).
<https://diariolaley.laley.es/home/NE0001786943/20181205/ProteccionDatos.aspx>

responsabilidad, consentimiento, seguridad, diseño seguro y trazabilidad, entre otras. Todas estas posibilidades hacen necesario partir de la explicación de qué es la IA.

3.1 La IA y los algoritmos

Explicar qué es la inteligencia artificial permite recurrir a varias posibilidades. Una definición formal sería aquella en la que una máquina "inteligente" se comporta como un agente racional flexible que percibe su entorno y lleva a cabo acciones que maximicen sus posibilidades de éxito en algún objetivo o tarea. Otra definición que podríamos usar tiene que ver con el Test de Turing (TURING, 1950), según el cual se puede considerar que existe verdadera Inteligencia Artificial cuando no podemos determinar si estamos interactuando con un ser humano o con una máquina. Comúnmente también se considera que existe inteligencia artificial cuando una máquina imita funciones cognitivas de organismos vivo (GARCÍA-PRIETO CUESTA, 2018: 46).

La Comisión Europea explica que:

El término «inteligencia artificial» (IA) se aplica a los sistemas que manifiestan un comportamiento inteligente, pues son capaces de analizar su entorno y pasar a la acción –con cierto grado de autonomía– con el fin de alcanzar objetivos específicos. Los sistemas basados en la IA pueden consistir simplemente en un programa informático (p. ej. asistentes de voz, programas de análisis de imágenes, motores de búsqueda, sistemas de reconocimiento facial y de voz), pero la IA también puede estar incorporada en dispositivos de hardware (p. ej. robots avanzados, automóviles autónomos, drones o aplicaciones del internet de las cosas).⁴

La inteligencia artificial trabaja con algoritmos cuya comprensión no está al alcance de los ciudadanos, por lo que se puede llegar a lo que se ha denominado como la “dictadura del algoritmo”, entendiendo por tal algoritmo una secuencia de instrucciones secuenciales mediante las cuales pueden llevarse a cabo ciertos procesos y darse respuesta a determinadas necesidades o decisiones. El Derecho trabaja con el lenguaje pero un ordenador no procesa lenguaje natural sino que su funcionamiento se basa en ejecutar un software escrito en un lenguaje de programación, que es una variedad de los lenguajes formales.

Los algoritmos que soportan los servicios de la robótica necesitan gestionar cantidades ingentes de datos personales cuya seguridad y confidencialidad deben quedar garantizadas, así como el compromiso de que sólo se usarán para fines exclusivamente autorizados por el usuario. Los ciudadanos deben poder interpretar los algoritmos que determinan esas decisiones

⁴ Comunicación de la Comisión al Parlamento europeo, al Consejo europeo, al Consejo, al Comité económico y social europeo y al Comité de las regiones Inteligencia Artificial para Europa {SWD(2018) 137 final} COM(2018) 237 final. Comisión Europea Bruselas, 25.4.2018.
<ec.europa.eu/transparency/.../2018/.../COM-2018-237-F1-ES-M>.

automatizadas, por ejemplo, por parte de las empresas cuando tengan gran impacto en su vida, como los que pronostican si alguien es supuestamente apto para un crédito de acuerdo a los requisitos de una entidad bancaria así como saber si esas entidades podrían estar usando sus sistemas de valoración de forma discriminatoria o no. Hay que comprobar que no se están utilizando algoritmos parciales con relación a raza, género y religión ni para favorecer los intereses de determinadas empresas. De este modo, se trata de erradicar desde el principio cualquier atisbo de comportamiento inmoral, como ocurrió en marzo de 2016 con Tay, un *chatbot* creado por Microsoft para hablar por Twitter con jóvenes de entre 18 y 24 años y que tuvo que ser retirado porque “aprendió” a escribir mensajes racistas (PLAZA LÓPEZ, 2017).

Preocupa los usos que puedan darse a la IA y al funcionamiento de los algoritmos. A este respecto ha habido diversas iniciativas, configurando grupos de investigación y de estudio tales como B-Debate, a propuesta de Biocat y la Obra Social “la Caixa”, que derivó en la publicación de la “Declaración de Barcelona para el Uso Apropiado de la Inteligencia Artificial”, celebrado el 8 de marzo de 2017.⁵ Este manifiesto, abierto a nuevas firmas y comentarios, recoge la preocupación por el posible uso “inapropiado, prematuro o malicioso” de las nuevas tecnologías y como contrapartida propone un código de conducta basado en seis puntos: prudencia, fiabilidad, transparencia, responsabilidad, autonomía restringida y el papel humano.⁶ También es destacable el informe de 2017 de *AI NOW Institute*, de la Universidad de Nueva York ⁷ con recomendaciones clave para el campo de la inteligencia artificial.

Martínez García advierte que si delegamos nuestra inteligencia en las máquinas, podría disminuir al aumentar la de las máquinas.

Las funciones que no se ejercitan se atrofian. Un mundo automatizado puede favorecer los automatismos. Acostumbrados a tratar constantemente con máquinas podemos mecanizarnos. La

⁵ Barcelona Declaration for the proper development and usage of Artificial Intelligence in Europe (08.03.2017). <https://www.iiia.csic.es/barcelonadeclaration/>.

⁶ La “Declaración de Barcelona” explica que la IA basada en datos, se ha comenzado a aplicar en la última década y funciona con grandes bases de datos procesados mediante algoritmos estadísticos para extraer patrones con los que hacer predicciones o emular un comportamiento humano. La IA basada en datos requiere usar el Big-Data y una gran potencia informática para alcanzar niveles de rendimiento adecuados. Entre los riesgos, destacan los denominados chat-bot. La IA ha permitido que estos chat-bots sean lo suficientemente realistas como para que los usuarios confiados no sean capaces de distinguir si están interactuando con una máquina o con un humano y sean engañados. Por ello, los científicos proponen exigir que “siempre esté claro si una interacción proviene de un humano o de un sistema de IA y que, en el caso de que se trate de un sistema de IA, sus responsables puedan ser rastreados e identificados”. Otro problema que vislumbran los expertos es que “el innegable entusiasmo por la IA puede llegar a dar la impresión de que la inteligencia humana ya no es necesaria, llevando a algunas empresas a despedir empleados y reemplazarlos por los sistemas de IA”.

⁷ AI NOW Institute, de la Universidad de Nueva York es un centro de investigación interdisciplinario dedicado a comprender las implicaciones sociales de la Inteligencia artificial. <<https://ainowinstitute.org/>>. El informe se centra en áreas de impacto clave, que incluyen trabajo y automatización, parcialidad e inclusión, derechos y libertades, y ética y gobernabilidad. El informe identifica la falta de mujeres y de minorías subrepresentadas que trabajan en IA como un problema fundamental que probablemente tenga un impacto material en los sistemas de inteligencia artificial y forme sus efectos en la sociedad.

inteligencia artificial puede promover un profesional del derecho rutinario, falto de reflejos, que delega en la máquina y se deja llevar por los programas establecidos. Los derechos humanos, como componentes centrales de la inteligencia jurídica, están llamados a ejercer una función de vigilancia, a mantener alerta al jurista. Una de sus funciones es de carácter cognitivo: despejar la mente, abrirla, evitar que se reduzca a “jurista digital”, a terminal de una instalación electrónica. [...] La inteligencia artificial puede ser sólo *operativa*, pero la humana es radicalmente *imaginativa*. Pensar los derechos es imaginarlos. Están siempre abiertos a nuevas versiones y posibilidades (MARTÍNEZ GARCÍA, 2019: 175).

La aplicación de la IA a diversas áreas y campos ha ido ampliándose progresivamente (DE ASÍS ROIG, 2014). En cualquier caso, las tecnologías de IA se están introduciendo en áreas críticas como la justicia –principalmente la penal–, lo que hace necesario analizar esta posibilidad y detectar si hay disfunciones en los algoritmos.

3.2 La aplicación de la IA a los sistemas judiciales

Entre las diversas áreas susceptibles de aplicar la IA se encuentran los sistemas judiciales. Como advierte J. I. Martínez García:

Ya no nos conformamos con decir que el derecho es un conjunto de normas. Lo vemos como un conjunto de programas. Y no sólo ejecuta programas: decide. Toda regulación jurídica tiene que ver con la toma de decisiones. Pero decidir no es lo mismo que ejecutar programas (MARTÍNEZ GARCÍA, 2019: 171 y 177).

La Comisión Europea para la Eficiencia de la Justicia (CEPEJ) del Consejo de Europa,⁸ ha adoptado el primer texto europeo que establece principios éticos relativos al uso de la inteligencia artificial en los sistemas judiciales: la Carta Ética Europea sobre el uso de la Inteligencia Artificial en los sistemas judiciales y su entorno (*European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment*).⁹ La finalidad de este documento es proporcionar un marco de principios que pueden guiar a los responsables políticos, legisladores y profesionales de la justicia ante el rápido desarrollo de la IA en los procesos judiciales nacionales.

La CEPEJ, tal como se establece en la Carta, parte de la asunción de que la aplicación de la IA en el ámbito de la justicia puede contribuir a mejorar su eficiencia y la calidad. Pero, a la vez, que debe aplicarse de manera responsable, respetando los derechos fundamentales garantizados por el ordenamiento y, en particular, por el Convenio Europeo de Derechos

⁸ El CEPEJ reúne a expertos de los 47 estados miembros del Consejo de Europa con el objetivo de mejorar la calidad y la eficiencia de los sistemas judiciales europeos y fortalecer la confianza de los usuarios de los tribunales en tales sistemas

⁹ <https://rm.coe.int/ethical-charter-en-for...4.../16808f699c>

Humanos (CEDH) y por el Convenio del Consejo de Europa para la Protección de Datos Personales.

El CEPEJ ha identificado los siguientes principios fundamentales que deben respetarse en el campo de la IA y la justicia:

a) Respeto a los derechos humanos: Entre estos principios, el respeto a los derechos humanos y la no discriminación es de importancia fundamental. El objetivo es garantizar, desde la concepción hasta la aplicación práctica, que las soluciones garanticen el respeto de los derechos garantizados por el Convenio Europeo de Derechos Humanos y el Convenio núm. 108 del Consejo de Europa. El principio de no discriminación se expresa expresamente debido a la capacidad de ciertos procesos, operaciones, en particular en asuntos penales, para revelar la discriminación existente al agregar o clasificar datos relacionados con personas o grupos de personas. Por lo tanto, los actores públicos y privados deben garantizar que estas aplicaciones no reproduzcan o agraven esta discriminación y no conduzcan a análisis o prácticas deterministas.

b) Principio de calidad y seguridad: También se tienen en cuenta algunos desafíos cualitativos relacionados con la metodología de análisis y el procesamiento automatizado de las decisiones judiciales. Se establece claramente un principio de calidad y seguridad según el cual debe ser posible procesar los datos mediante el aprendizaje automático sobre la base de originales certificados y la integridad de estos datos debe garantizarse en todas las etapas del procesamiento. Se recomienda encarecidamente la creación de equipos multidisciplinarios, compuestos por jueces, investigadores en ciencias sociales e informática, tanto en la fase de redacción como de dirección y en la aplicación de las soluciones propuestas.

c) Principio de transparencia: El principio de transparencia de las metodologías y técnicas utilizadas en el procesamiento de las decisiones judiciales también es de gran importancia. El énfasis aquí está en la accesibilidad y comprensión de las técnicas de procesamiento de datos, así como en la posibilidad de que las autoridades o expertos independientes realicen auditorías externas. También se fomenta un sistema de certificación, que se renovará periódicamente.

d) Principio 'bajo el control del usuario': Además, se incluye el principio *bajo el control del usuario*, con él se subraya la necesidad de convertir al usuario en un agente activo y de sentirse a cargo de sus elecciones. En particular, el juez debe poder volver en cualquier momento a las decisiones judiciales y los datos que se han utilizado para producir un

resultado y seguir teniendo la posibilidad de apartarse de él, teniendo en cuenta las especificidades del caso en cuestión. Cada usuario debe ser informado, en un lenguaje claro y comprensible, de la naturaleza vinculante o no vinculante de las soluciones propuestas por los instrumentos de AI, las diversas opciones posibles y su derecho a la asesoría legal y al recurso ante un tribunal

Para el CEPEJ, el cumplimiento de estos principios debe garantizarse en el procesamiento de las decisiones judiciales y los datos mediante algoritmos y en el uso que se haga de ellos. La Carta del CEPEJ se acompaña de un estudio en profundidad sobre el uso de AI en los sistemas judiciales, en particular las solicitudes de AI que procesan decisiones y datos judiciales. En este estudio se recogen cuatro tipos diferentes de posibles usos de la IA en el ámbito de la Justicia, clasificados según el posible riesgo derivado de su uso:

- a) Usos que deben ser fomentados, tales como:
 - i) La mejora en la búsqueda de jurisprudencia: La aplicación de las técnicas de *machine learning* al procesamiento del lenguaje natural constituyen una ventaja considerable para encontrar opciones de búsqueda para complementar la búsqueda por palabras clave o a texto completo (*full text*). Además estas herramientas pueden relacionar diversas fuentes de datos (por ejemplo, leyes, jurisprudencia y doctrina científica) optimizando su presentación mediante nuevas técnicas de visualización de la información;
 - ii) Acceso a la legislación: Los *chatbots* pueden facilitar el acceso a las diferentes fuentes de información existentes por medio del lenguaje natural. La tecnología también puede permitir la creación en línea de documentos (demandas, contratos de arrendamiento, etc.);
 - iii) Creación de nuevas herramientas estratégicas: La ciencia de datos y la inteligencia artificial pueden ayudar a mejorar la eficiencia de la justicia al permitir la realización de evaluaciones cuantitativas y cualitativas y para hacer proyecciones (por ejemplo, sobre la necesidad de recursos humanos y presupuestarios). Se recomienda que los profesionales del derecho, y especialmente de los jueces, participen en la implementación de estas herramientas, haciéndolos suyos analizando sus resultados en relación con las características específicas del tribunal en cuestión o la calidad de la justicia (por ejemplo, la necesidad de preservar el acceso a la justicia).
- b) Posibles usos, que requieren considerables precauciones metodológicas, tales como:
 - i) Ayuda en la elaboración de baremos en determinados litigios civiles;
 - ii) Apoyo a las medidas alternativas de solución de conflictos en materia civil;
 - iii) Resolución de disputas

en línea; iv) El uso de algoritmos en la investigación penal para identificar dónde se están cometiendo los delitos.

c) Usos que pueden considerarse después de realizar estudios científicos adicionales, tales como: i) Perfilado de jueces: Aunque el informe muestra ciertas reticencias en cuanto al alcance de esta actividad, también añade que podría fomentarse ofrecer a los jueces información cuantitativa y cualitativa detallada sobre su actividad, con un objetivo puramente informativo de ayuda a la toma de decisiones y para su uso exclusivo; ii) Anticipación de las decisiones judiciales: según el informe, los sistemas híbridos, basados en la construcción de modelos matemáticos que se supone que representan el modelo de razonamiento de los jueces, no son todavía eficientes porque están limitados por el sesgo en la muestra de datos que han procesado y porque además tienen que empezar de nuevo desde cero si se modifica una ley o si se produce un cambio en la jurisprudencia.

d) Usos a considerar con las más extremas reservas, tales como: i) Utilización de algoritmos para la elaboración de perfiles individuales en materia penal: Los experimentos en otros países (como el programa COMPAS en los Estados Unidos y HART en el Reino Unido) han sido criticados a causa de las limitaciones de la metodología utilizada basada en un acercamiento puramente estadístico a los hechos que ha conducido a resultados erróneos, al atribuir mayor riesgo delictivo a los sujetos de raza afroamericana, por ejemplo. Este enfoque discriminatorio y efectos deterministas, debe ser sustituido por otro más respetuoso con los estándares europeos. Cabe impulsar el desarrollo de sistemas algorítmicos que permitan ayudar a mejorar la recopilación y el análisis de la información necesaria para los servicios de libertad condicional, pero cualquier otro uso es propenso a sesgos que entrarán en conflicto con principios fundamentales nacionales y supranacionales; ii) Normas basadas en la cantidad (*Quantity-based norms*). No se trata de inundar de información a los jueces, sino de facilitarles el fundamento de las decisiones adoptadas por otros jueces, sin limitar su capacidad de decidir por el mero peso de los precedentes. De hecho, el estudio de la CEPEJ también ha puesto de relieve los peligros de la cristalización de la jurisprudencia y de sus posibles efectos negativos sobre la imparcialidad e independencia de los jueces.

4 LAS PROYECCIONES DEL PRINCIPIO DEL RIESGO EN EL SISTEMA PENAL DE JUSTICIA

Basada en un cuento de Philip K. Dick, la película de ciencia ficción de 2002 *Minority Report* presenta una visión del futuro en la que se pueden predecir y prevenir los crímenes. Aunque pueda parecer una utopía, sin embargo, la justicia predictiva ya existe hoy en día. Construidas sobre sistemas avanzados de aprendizaje automático, nuevas empresas proporcionan servicios de predicción a los tribunales; por ejemplo, en forma de algoritmos de evaluación de riesgos que estiman la probabilidad de reincidencia de los delincuentes para ayudar a los jueces en su toma de decisiones. Así, en el ámbito del Derecho Penal es donde la justicia predictiva o justicia algorítmica goza de mayores posibilidades de aplicación.

Puesto que Estados Unidos tiene un altísimo ritmo de encarcelación -que multiplica por cinco la media de la OCDE-, están surgiendo movimientos reformistas que buscan la forma de descongestionar las cárceles. Una de las propuestas planteadas es que las sentencias se basen no en una noción de castigo, sino de riesgo. ¿Cómo de peligroso es para la sociedad tener a cierto individuo fuera de la cárcel? Eso se decidiría en base a unas puntuaciones de riesgo que son calculadas por algoritmos de *machine learning* para establecer si un reo reincidirá o no en caso de ser puesto en libertad (PASCUAL, 2017). Estas predicciones algorítmicas no están exentas de riesgo, como se examinará seguidamente.

4.1. El riesgo en el sistema penal. ¿se puede predecir quiénes serán los futuros delincuentes o futuras reincidencias?

Como explica Romeo Casabona, existe actualmente una tendencia en vías de expansión, en concreto en el Derecho Penal continental europeo -que, a su vez, proviene del sistema jurídico-penal anglosajón- con una metodología que trata de sustituir la noción de peligrosidad criminal como característica psicológica (no necesariamente patológica), y por ello individual y subjetiva, del delincuente, en cuanto presupuesto necesario para poder imponer una medida de seguridad, por la del riesgo de delinquir que presenta la conducta de una persona de forma objetiva (BLAY GIL, 2014: 81; ROMEO CASABONA, 1986: 20 ss). Es decir, hay una corriente con preferencia metodológica por la valoración del riesgo, a la que atribuyen una mayor eficacia predictiva, más objetividad y utilidad profesional. Hay cada vez más especialistas que prefieren orientarse a presupuestos - más operativos en su opinión- como es el concepto de riesgo, que legitimen la imposición de medidas de seguridad (ROMEO CASABONA, 2018: 40).

Esta concepción, que en el fondo parte en ocasiones de ideas retribucionistas e incluso vindicativas encubiertas, ha allanado el camino a tratamientos inocuidos del delincuente peligroso, y en particular en los últimos años al que es imputable, a través de penas de larga duración, como la cadena perpetua o la prisión permanente revisable

[...]. A la criticada escasa utilidad pronóstica de la peligrosidad criminal, entre otras censuras, se contraponen por esta corriente una preferencia metodológica por la valoración del riesgo, a la que atribuyen una mayor eficacia predictiva, más objetividad y utilidad profesional ((ROMEO CASABONA, 2018: 46-47).

Hay que recordar que en el sistema penal, los encargados de tomar decisiones legales a los que se les pedía que evaluaran el riesgo de un delincuente se basaban generalmente en las opiniones de los profesionales de la salud mental, las evaluaciones de los oficiales de libertad condicional o su propio análisis. Este tipo de juicio a menudo se denomina predicción "clínica", para distinguirla de la predicción basada en estadísticas y "actuarial", y sigue siendo la base de los procesos de toma de decisiones posteriores a la condena y previos a la condena en muchas jurisdicciones.

Los algoritmos se utilizan cada vez más en los tribunales para "fijar la fianza, determinar las sentencias e incluso contribuir a la determinación de la culpabilidad o inocencia". Este cambio hacia una mayor inteligencia artificial en los tribunales, que permitiría a la inteligencia artificial aumentar el juicio humano, podría ser extremadamente beneficioso para el sistema judicial en su conjunto.

Actualmente, el "principio de riesgo" tiene una especial relevancia ya que se utiliza como el primer principio básico de reducción de reincidencia del infractor. Los instrumentos de evaluación de riesgos -*risk assessment* instruments (RAIs)- están diseñados para predecir la reincidencia. Aunque se señalan críticas a los RAIs (no son más precisas que las evaluaciones clínicas, sesgadas racialmente, carecen de transparencia y, debido a su naturaleza cuantitativa, deshumanizantes) (GOEL), también hay defensores de sus ventajas: los RAI pueden predecir la reincidencia de manera más consistente, transparente y precisa que el juicio humano sin ayuda, es decir, la opinión intuitiva de un juez, oficial de libertad condicional, clínico u otro profesional.

Romeo Casabona advierte tanto de la vulnerabilidad del concepto de peligrosidad criminal como de los métodos predictivos. En relación al primero, no resultan fiables ni los métodos que se utilizan para averiguar la peligrosidad ni la validez de los propios síntomas. Y la predicción con respecto a la probabilidad de volver a delinquir, por su propia naturaleza, es incierta; es una mera hipótesis de trabajo, pues no puede haber una certeza matemática de que el individuo sometido a examen cometerá un delito en el futuro¹⁰

Pero también los métodos predictivos adolecen de inconvenientes que, incluso, pueden ser superiores a los del concepto de peligrosidad criminal. En estos procedimientos preocupa eliminar los componentes subjetivos que el propio especialista evaluador del delincuente podría proyectar sobre su evaluación (por prejuicios, empatía o falta de empatía con el evaluado, estados

de ánimo del evaluador); es decir, la evaluación ha de ser neutral, objetiva, libre de criterios subjetivos del especialista. Los juicios de predicción del riesgo se sirven de diversas herramientas. El primero de ellos fue el estadístico, mediante la elaboración o utilización de unas tablas de pronóstico o predicción. Se aplican modelos estadísticos y matemáticos para la evaluación del riesgo, en el caso del Derecho penal, del riesgo de comisión de un delito por un delincuente determinado, pero clasificándolo en ciertos grupos, sin entrar en su pronóstico conductivo psicológico concreto. Actualmente, el sistema de justicia predictivo, ya no se limita al uso de estadísticas o test, sino que se sirven de algoritmos para emitir el pronóstico, recurriéndose a sistemas inteligentes computerizados para su tratamiento. Se trata de la evaluación de un riesgo despersonalizado, pues para su declaración no interesa la causa u origen de ese riesgo.

Como señala Romeo Casabona:

No dejan de ser criterios próximos a la evaluación del riesgo que generan algunas actividades en otros sectores de la vida social (p. ej., la industria del seguro para la que el aseguramiento de ciertos riesgos es la base de su negocio; la industria biotecnológica, de la que se señalan sus riesgos para la vida y la salud de las personas y para el medio ambiente), la utilización de las energías, o el uso de ciertos instrumentos o aparatos, pues se valen de una metodología semejante (ROMEO CASABONA, 1986: 42).

Es un sistema que toma como punto de partida la metodología anterior, es decir, parte de una base actuarial, utilizando datos convertidos en algoritmos, en la que la evaluación de estos últimos disponibles se automatiza y se objetiva todavía más, si cabe, mediante su procesamiento por procedimientos informáticos de sistemas de inteligencia artificial. En esencia no es, por consiguiente, un método diferente, sino una herramienta tecnológica nueva que trataría de forma “inteligente” los algoritmos que sirven de base al procedimiento actuarial. Esta denominada justicia predictiva o justicia de algoritmo se utiliza tanto para otorgar la libertad condicional o la imposición de un programa de rehabilitación, como para fundamentar la condena a una pena de prisión de larga duración (FREEMAN, 2016: 75 ss.) o la imposición de una medida de seguridad. Adhiriéndonos a la opinión de a Romeo Casabona, es rechazable sustituir la peligrosidad criminal como presupuesto de las medidas de seguridad, a pesar de las limitaciones que pueda presentar (y por ello es necesario persistir en los estudios empírico-científicos), por otros instrumentos basados en la idea de riesgo, entendido éste como una fuente objetiva y despersonalizada -deshumanizada-, de modo semejante a como sucede con riesgos provenientes de fenómenos naturales, de actividades humanas, de objetos o instrumentos creados por el ser humano.

4.2 Experiencias predictivas de justicia en Derecho comparado: dictadura del algoritmo VS. decisión judicial

En diversos contextos se están llevando a cabo experiencias de sistemas de IA para ayudar a neutralizar el riesgo y lograr resultados más certeros que cuando son sólo los jueces quienes toman las decisiones. Estos sistemas llevan a comparar si los algoritmos han probado son o no mejores predictores de criminales reincidentes que los propios jueces. Se hará referencia a la experiencia en Argentina con PROMETEA; a la experiencia estadounidense con COMPAS; a la experiencia en Inglaterra con HART. Por último, se pondrán de manifiesto las tendencias en España al respecto, en el ámbito del Derecho Penal.

PROMETEA es la primera inteligencia artificial predictiva al servicio de la Justicia de Argentina. ¿Puede una máquina predecir la solución de un caso judicial en dos minutos? El sistema se probó con 161 expedientes de temáticas consideradas factibles de tratar con este desarrollo: cuestiones procesales, de caducidad, empleo público -y derecho a la vivienda, en las que mostró una eficacia de 98%. No se pretende reemplazar a funcionarios judiciales y abogados.

Detrás de PROMETEA siempre debe haber un ser humano que, con su inteligencia natural, y no artificial, defina si la propuesta del sistema es adecuada o no. PROMETEA está siendo probado en la fiscalía general de la ciudad y en la Corte Interamericana de Derechos Humanos (CIDH). El procedimiento es el siguiente: Se carga el número de expediente y el sistema busca la carátula en la página del Tribunal Superior de Justicia. Después, lee” las sentencias de primera y segunda instancia, analiza más de 1400 dictámenes (emitidos durante 2016 y 2017), realiza algunas preguntas al operador para completar datos y emite la predicción.

Se utiliza principalmente en tres procedimientos: para la resolución del fondo de asistencia legal a las víctimas; para realizar notificaciones a los países que forman parte de la Organización de los Estados Americanos (OEA); y como herramienta de búsqueda para rastrear precedentes. Su tarea es predecir sobre expedientes donde la Justicia ya tiene algo dicho. Permite que la persona pueda destinar su tiempo a los casos complejos. Es un sistema que lee patrones de información y detecta cuáles se corresponden con el caso presentado

COMPAS es un acrónimo que en español puede traducirse como Administración de Perfiles de Criminales para Sanciones Alternativas del Sistema de Prisiones de EE.UU. (Correctional Offender Management Profiling for Alternative Sanctions). La herramienta COMPAS pretende predecir el riesgo de reincidencia de un individuo sobre la base de un análisis complejo que implica el uso de la información obtenida de una encuesta de 137 preguntas dividida en varias secciones diferentes, así como de la información correspondiente a los antecedentes penales individuales. Incluye preguntas como: "¿alguien en tu familia ha sido arrestado?", o "¿cuántas veces has sido arrestado?". ¿El individuo tiene una pandilla?; ¿Es parte de

un programa de rehabilitación?; ¿Cuántas faltas ha tenido mientras estaba en prisión?; También evalúa otros aspectos como el lugar de procedencia, el barrio en el que los reos habitan, la raza o edad de cada preso. ¿Quién puede ser desviado a programas de tratamiento de drogas?; ¿quién puede ser liberado antes del juicio?; ¿quién debe ser enviado a la cárcel por el tiempo más largo posible?; ¿quién debe obtener una sentencia indulgente?

La industria de los software predictivos está en rápido crecimiento. Por un lado, están los programas que intentan predecir las tendencias criminales de las personas, como Compas, y por el otro, los que pretenden elaborar mapas de potenciales áreas de tensión en las ciudades, donde las posibilidades de cometer un delito son estadísticamente mayores. A este segundo grupo pertenece PredPol, un software desarrollado por la Universidad de California-Los Ángeles y utilizado por más de 60 departamentos de policía en los EE.UU. En la página web de PredPol se explica que el programa se fundamenta en un historial de delitos pasados (entre dos y cinco años de datos), enriquecido a lo largo del tiempo. Con esta información, el algoritmo intenta predecir dónde y cuándo será más probable que ocurra un cierto tipo de crimen. No prevé los asesinatos, pero sí permite mapear en tiempo real la criminalidad potencial de las ciudades. La asociación *Human Rights*, además, sostiene que PredPol se centra casi solo en los barrios más pobres, discriminando a las comunidades afroamericana e hispana.

Sin embargo, los algoritmos de “asesoramiento de la peligrosidad” se han visto, previamente, profundamente criticados. Así, el programa COMPAS también ha recibido críticas, primero porque se ha considerado que es racista y segundo, porque no resulta fiable en cuanto a sus predicciones.

En relación al sesgo racista, una investigación de ProPublica publicada en mayo de 2016 analizó los resultados y las estadísticas recogidas durante dos años en un condado y llegó a la conclusión de que sólo el 20% de personas que el programa consideró que podían cometer un crimen violento lo hizo. Pero lo más grave es que concluyeron que el programa exageraba el riesgo de reincidencia en personas afroamericanas como más propensas a volver a cometer crímenes, mientras que la raza blanca no era considerada un factor de riesgo, según el análisis de la metodología predictiva del programa.¹¹ Si se compara a una persona negra y una blanca que tienen el mismo historial, la misma edad, el mismo género, el mismo pasado judicial y el mismo 'futuro criminal' (las posibilidades de cometer un crimen, dos crímenes o ninguno), el acusado negro tiene un 45% más de posibilidades de obtener un puntaje de riesgo que un acusado blanco. En estos casos, el uso de este algoritmo puede violar el principio de igualdad y de no

¹¹<https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE.html>; <https://www.propublica.org/.../compas-recidivism-risk-score-dat>; <https://www.propublica.org/.../how-we-analyzed-the-compas-re>; <https://www.equivant.com/response-to-propublica-demonstratin>.

discriminación, porque las predicciones que realizan se basan en un código fuente que toma en consideración distinciones de raza, género, entre otras (BISWAS; BERK et al. 2017). Se ha fundado la Liga de Justicia Algorítmica,¹² con la finalidad de que los sistemas de reconocimiento facial o la filtración de mensajes en función de la raza en redes sociales eviten el prejuicio racial.

Un ejemplo de que los algoritmos muestran prejuicios raciales, se encuentra en el caso del Sr. Loomis. En el caso *Wisconsin v. Loomis*¹³ el Tribunal Supremo del Estado de Wisconsin dictó una sentencia en la que se pronunció por primera vez sobre la constitucionalidad del uso de algoritmos procesados informáticamente (mediante un sistema de inteligencia artificial) como fundamento del fallo del tribunal de instancia inferior. Resumidamente, los hechos fueron los siguientes: en 2013 Eric Loomis fue detenido por agentes de policía del Estado de Wisconsin (Estados Unidos) cuando conducía un vehículo implicado en un reciente tiroteo. Se le acusaba de huir de la policía y de utilizar un vehículo sin la autorización de su propietario. Loomis se declaró culpable de ambos delitos con la esperanza de que no tuviera que ingresar en prisión. Durante la vista para decidir sobre su libertad condicional, el fiscal aportó un informe elaborado por el programa informático Compas, desarrollado por la empresa privada Northpointe Inc, según el cual el señor Loomis tenía un riesgo elevado de reincidencia y de cometer actos violentos. El informe concluía que el condenado representaba un “alto riesgo para la comunidad”. El tribunal de instancia declaró a este respecto: "Usted ha sido identificado, a través de la evaluación de COMPAS, como un individuo de alto riesgo para la comunidad. Mediante la ponderación de diversos factores, hemos descartado la libertad condicional debido a la gravedad del crimen cometido y porque su historial personal, su historial de libertad bajo control y las herramientas de evaluación de riesgos que se han utilizado, sugieren que Ud. representa un riesgo extremadamente elevado de volver a delinquir". Partiendo de tales consideraciones, el juez impuso al señor Loomis una pena de 6 años de prisión y otros 5 en régimen de libertad vigilada. Como se deduce de este párrafo y de la propia condena, no se trataba de imponer una medida de seguridad, sino de una pena de mayor duración y de no conceder la libertad condicional.

La defensa del Sr. Loomis no aceptó la sentencia y decidió apelarla, alegando que el uso del sistema COMPAS era inadecuado y violaba el derecho de su representado al debido proceso (*the right to a due process*). Apoyó su recurso en tres argumentos principales. Conforme al primero de ellos, declaró que no había podido conocer la manera exacta por la que se calculó su puntuación con COMPAS, debido a que esta técnica era de propiedad privada (y se determinó por el tribunal que había que proteger los derechos de propiedad intelectual de la empresa que lo construyó, al parecer recurriendo al llamado secreto de empresa). En segundo lugar, alegó que el

¹² <https://www.ajlunited.org/>

¹³ Sentencia 13 de julio de 2016: *State v. Loomis*, 881, N.W.2d 749, 7532 (Wis, 2016).

uso de este software violaba su derecho a una sentencia individualizada, ya que el algoritmo únicamente generaba datos basados en estadísticas de grupo. Y, por último, argumentó que su uso constituía una discriminación por razón de género porque sus evaluaciones tenían en cuenta el género.

Cuando el Tribunal Supremo del Estado de Wisconsin publicó su fallo, desestimó las cuestiones que fueron objeto del recurso de apelación.¹⁴ A partir de entonces, se creó un precedente relevante sobre la aceptación del uso del sistema inteligente COMPAS o de cualquier otro instrumento automatizado de evaluación de riesgos por medio de algoritmos, en la determinación de una sentencia. Por primera vez, una sentencia se pronunciaba sobre la constitucionalidad del uso de algoritmos procesados informáticamente (mediante un sistema de inteligencia artificial) como fundamento del fallo del tribunal de instancia inferior (*caso Wisconsin v. Loomis*)¹⁵. En su fallo, el Tribunal aceptaba su utilización argumentando que el derecho de los acusados al debido proceso no fue violado por el mero hecho de que no pudieron acceder a una explicación adecuada sobre el tratamiento informático de los algoritmos. Para apoyar esta posición, el Tribunal señaló que la exactitud de los instrumentos utilizados y la capacidad de los jueces para entender su posible mal funcionamiento eran suficientes para asegurar los derechos de los acusados (DE MIGUEL BERIAÍN, 2018: 45-53).

Como acertadamente subraya Romeo Casabona, resulta paradójico que al final el Tribunal Supremo de Wisconsin insista en la admisión de este medio de prueba, cuando reconoce en el sistema COMPAS y en el empleo de algoritmos tantas limitaciones, debilidades e inconvenientes, y que acabe reconduciendo su utilización al fundamental principio de libre apreciación de la prueba por parte del juez, lo que, sin embargo, no parece que se hiciera en la primera sentencia de instancia. Ello lleva a pensar que incluso aunque en el futuro se desarrollen nuevos sistemas inteligentes más perfeccionados, cabe prever que el principio de libre apreciación de la prueba debe continuar siendo la clave de la toma de la decisión por parte de los jueces y los tribunales, con el único límite, claro está, de no traspasar la frontera de la arbitrariedad ya que es el medio más contundente de garantizar que una decisión de tanta trascendencia social quede en manos del ser humano y no en sistemas artificiales.¹⁶

Una vez expuesto el caso sobre los problemas de COMPAS con el sesgo racial, en el último año el foco de atención se ha puesto en si COMPAS es fiable en sus predicciones. Cuando Julia Dressel estaba estudiando Ingeniería Informática y Estudios de Género en la Universidad, encontró la investigación de ProPublica que demostraba lo prejuiciosos que eran estos algoritmos

¹⁴ Por sentencia del 13 de julio de 2016. *State vs. Loomis*, 881 N.W.2d 749, 770-71 (Wis. 2016).

¹⁵ Sentencia 13 de julio de 2016: *State v. Loomis*, 881, N.W.2d 749, 7532 (Wis, 2016).

¹⁶ Así lo propugna el European Group on Ethics in Science and New Technologies (EGE), Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems, European Commission, Brussels, 2018, p. 9.

con respecto a la raza. Pero lo que llamó la atención de Dressel fue que no se pusiera en duda que estas predicciones fueran más certeras que las humanas. “Por debajo de toda la conversación sobre algoritmos se asumía que las predicciones algorítmicas eran por naturaleza superiores a las predicciones humanas”. Para evaluar esto, Dressel y Harid, dos profesores de informática de Dartmouth College, (DRESSEL and FARID , 2018)¹⁷ elaboraron una encuesta que publicaron en Amazon Mechanical Turk, donde cualquier internauta podía predecir la posibilidad de reincidencia de una persona sólo tomando en cuenta siete características. Los datos pertenecían a personas reales cuya información personal, incluyendo si había vuelto a cometer crímenes o no, se había hecho pública.

Cada encuestado evaluaba a 50 personas en base a descripciones que tenían el siguiente formato: “El acusado es [género] y tiene [edad]. Ha sido acusado de [cargo criminal]. Este crimen está clasificado como [grado criminal]. Antes había sido condenado a [cargo anterior no-juvenil]. Su registro tiene el cargo de felonía juvenil [cargo de felonía juvenil] y el cargo menor juvenil [cargo menor juvenil]”. Al revisar los resultados y compararlos con los datos de reincidencia de los acusados, los investigadores descubrieron que las respuestas de los internautas tuvieron una precisión del 67%, superior al 65% del algoritmo. Los investigadores siguieron encuestando a los usuarios, omitiendo datos pero aun así obteniendo resultados similares. Hasta que se dieron cuenta de que lo único que necesitaban para hacer sus predicciones era la edad y la cantidad de condenas previas. “En esencia, si eres joven y tienes muchos antecedentes criminales eres de alto riesgo, si eres viejo y tienes menos antecedentes, eres de bajo riesgo”.¹⁸ Ese casi empate supone un varapalo para el algoritmo, ya que su toma de decisiones se basa en hasta 127 criterios, mientras que a los humanos seleccionados en esa especie jurado popular predictivo sólo les ofrecieron algunos datos sobre los condenados y potenciales reincidentes.

De lo expuesto se ha podido identificar uno de los principales riesgos a los que tendrá que enfrentarse el sistema judicial (principalmente jueces y fiscales) y penitenciario en los años venideros es el de sobrevalorar la fiabilidad de las nuevas tecnologías y apoyarse ciegamente en las mismas a la hora de tomar las posiciones y decisiones que corresponden a los diversos agentes.

¹⁷ Los medios de comunicación y diversos estudios se han hecho eco, unos en contra (“Los tribunales usan algoritmos para ayudar a determinar la sentencia, pero las personas al azar obtienen los mismos resultados” (CHODOSH, 2018). ¿Puede el software predecir ¿crimen? Tal vez sea así, pero no mejor que un ser humano” (CHOKSHI, 2018), “El software no es más preciso que los humanos sin entrenamiento “a la hora de juzgar el riesgo de reincidencia” (DEVLIN, 2018) y otros a favor (BANSACK, Kirk. “Can non-experts really emulate statistical learning methods? A comment on “The accuracy, fairness, and limits of predicting recidivism”. Stanford, 2018, pp.1-30. web.stanford.edu/.../cgi.../Bansack_RecidivismAlgo_Comment.p...).

¹⁸ Los programas que predicen la reincidencia criminal en los tribunales no son más certeros que los presentimientos de un internauta cualquiera. <https://securelist.lat › News>

Al menos respecto al juez, significaría que éste renuncia a la posición insustituible que tiene encomendada de apreciar el conjunto de medios de prueba aportados en el proceso desde una visión y ponderación también de conjunto, posición singular que nunca llegan a adquirir los demás intervinientes en el proceso. Tal vez podría dejarse llevar por la comodidad que favorece la rutina de dar por bueno lo que viene respaldado por el humano experto o –en este caso- por las tecnologías “expertas” (ROMEO CASABONA, 2018: 54).

En Inglaterra, como es común en todo el sector público, el servicio de policía del Reino Unido está bajo presión para hacer más con menos, orientar los recursos de manera más eficiente y tomar medidas para identificar las amenazas de manera proactiva; por ejemplo, bajo esquemas de evaluación de riesgos tales como la "Ley de Clare"¹⁹ y la "Ley de Sarah".²⁰ Las mujeres de Inglaterra y Gales pueden conocer si sus parejas arrastran un historial de violencia. En esta misma línea, las herramientas algorítmicas prometen mejorar las capacidades de predicción y toma de decisiones de una fuerza policial haciendo un mejor uso de los datos (incluida la inteligencia). HART es uno de los primeros modelos algorítmicos implementados por una fuerza policial del Reino Unido en una capacidad operativa. Los investigadores destacan como institución municipal pionera en este campo al Ayuntamiento de Bristol, el cual cuenta con un organismo específicamente dedicado, el *Bristol Integrated Analytics Hub*. La misión del mismo es la de recopilar información sobre ingresos, asistencia a la escuela, casos de embarazos adolescentes o de enfermedad mental, a partir de los datos de 54.000 familias locales. El objetivo es lograr, a partir de dichos datos, predecir qué niños están más expuestos a sufrir casos de violencia doméstica, abuso sexual o situaciones similares. El sistema funciona asignando una puntuación de 0 a 100 a cada menor, una cifra que indica su valoración de riesgo, y a partir de ahí, proporciona

¹⁹ La medida, conocida como la *ley de Clare (Clare's law)* porque toma el nombre de una de las víctimas de la violencia de género, permite a la policía facilitar esa información en algunas zonas del país. Con la *ley de Clare* también la policía puede revelar acusaciones de abuso contra una persona, aunque estas nunca fueran probadas ni mediara ningún tipo de cargo. A pesar de que la medida ha despertado una gran controversia en el país, ello no ha impedido la ampliación del programa experimental -que funcionaba en el área metropolitana de Manchester, en Wiltshire, Nottinghamshire y Gwent desde septiembre de 2012-. Desde entonces, la policía de esas áreas ha procurado datos a 386 mujeres sobre sus compañeros actuales, o sobre hombres con los que anteriormente habían mantenido una relación, con la intención de evitar delitos o crímenes futuros.

²⁰ Estados Unidos y Gran Bretaña ya han afrontado el peligro que representan los pederastas para sus respectivas sociedades. Los dos han desarrollado leyes que suponen el control de estos sujetos una vez que quedan en libertad porque han cumplido la condena. Los dos desarrollos legislativos se produjeron debido a dos casos que impactaron enormemente a la opinión pública. En Estados Unidos, la “Ley Megan” fue aprobada, en Nueva Jersey, apenas un mes después del asesinato de la pequeña Megan. Esta ley obliga al registro de los delincuentes sexuales en una base de datos en cada Estado y permite a la Policía notificar a los vecinos del barrio de la llegada del delincuente sexual. También contempla la cadena perpetua, sin posibilidad de libertad, a los acusados de un segundo delito sexual. La Oficina Federal de Investigación (FBI) ha construido, a su vez, una base de datos nacional con las informaciones de los Estados, para poder seguir a cada pederasta cuando se desplacen de un Estado a otro. En Inglaterra, Sarah Payne y tenía ocho años cuando el pederasta Roy Whiting la secuestró cuando estaba jugando en un campo de maíz con sus dos hermanos, de 13 y 5 años, en el pueblo de Kingston Gorse, condado de West Sussex, al sur de Inglaterra. Fue el 1 de julio de 2000. Fue encontrada muerta. El pederasta vivía a 20 Km del domicilio de la niña. Actualmente está en vigor y, en esencia, es similar a la estadounidense. Con una diferencia. En el caso británico, los ciudadanos no tienen acceso al contenido de la base de pederastas. Es la Policía la que los controla y los vigila. Y los pederastas con antecedentes están obligados a registrarse en ella.

sugerencias a los trabajadores sociales del municipio, de la misma manera en que lo haría un motor de recomendaciones de webs como Amazon o Facebook. Se aconseja un concepto de proporcionalidad 'experimental' para permitir el uso de algoritmos no comprobados en el sector público de manera controlada y por tiempo limitado. Debería eliminarse de la influencia de la toma de decisiones algorítmica por completo (OSWALD et al. 2018: 223-250).

En Alemania crearon PRECOBS,²¹ en Italia, el sistema XLAW.²² El entusiasmo generalizado por los algoritmos hace que algunos pretendan usar el aprendizaje automático para leer las caras de las personas y adivinar sus intenciones más secretas. VAAK es un software creado por una *startup* japonesa que analiza el lenguaje corporal e identifica las actitudes más sospechosas.²³ Inquietud, incertidumbre, movimientos inusuales son captados por las cámaras y analizados por el software para evitar, por ejemplo, los robos en los supermercados. Cortica, una empresa israelí que se ocupa de seguridad e inteligencia artificial, ha firmado recientemente un acuerdo para analizar las imágenes recopiladas por las cámaras de seguridad en áreas públicas de la India. El software de la compañía promete buscar anomalías de comportamiento que indiquen que alguien está a punto de cometer un crimen violento. El programa se basa en los sistemas militares desarrollados para identificar terroristas. Busca las llamadas microexpresiones, pequeños espasmos o gestos que pueden revelar las intenciones de una persona. Según sus desarrolladores, Cortica puede incluso aprender a predecir cuándo un mercado pacífico y lleno de personas o una manifestación política están a punto de volverse violentos. Una aplicación potencialmente muy útil en Israel. A nuestro juicio, debe eliminarse la tentación de perseguir el pre-crimen. Los ciudadanos tienden a confiar en la computadora; aunque se diga que puede ayudarte a tomar una decisión, en realidad reemplaza a la decisión humana, porque se confía en que siempre estará en lo correcto.

En Estonia, a finales de año comenzará a desarrollarse un programa de 'jueces robot', especializados en conflictos de entre 6.000 y 7.000 euros. El proyecto en el país báltico es el primero con una fecha ya concretada y se centra en cuestiones vinculadas fundamentalmente a conflictos vinculados a contratos: las partes subirán la documentación a Internet y el 'juez robot' decidirá. Si las personas no están de acuerdo con el fallo del robot, podrían recurrir la decisión a un tribunal humano.

²¹ www.ifmpt.de/projects.html

²² www.elialombardo.it/documenti/siap_n69.pdf

²³ <https://yellrobot.com/shoplifting-detection-software-vaak/>
"This AI Detects Shoplifters Before They Steal, But There's Something Even Creepier About It". (07.03.2019).
<https://www.outerplaces.com/science/item/19294-vaak-ai>

Por último, en España los programas predictivos aún no se utilizan de manera oficial, pero el programa Eurocop, sistema de gestión policial²⁴ un proyecto de la Universidad de Castellón, está diseñado para elaborar un mapa de previsión de riesgo en lugares concretos de una ciudad y en determinados horarios. Asimismo, un equipo de científicos de la Universidad Autónoma de Madrid busca un algoritmo que pueda predecir mejor el riesgo que corre una mujer maltratada de volver a sufrir otra agresión. Para ello, se ha probado el nuevo cuestionario que permita mejorar la valoración del riesgo de las maltratadas. En el cuestionario se incluyen preguntas sobre acoso por Internet y se tiene en cuenta el entorno de la víctima y aspectos de su vida que puedan hacerla vulnerable. El formulario nuevo recoge 39 variables relacionadas con la vulnerabilidad de las víctimas de malos tratos, más del doble del que aspira a relevar. "Entre las nuevas variables incluimos el acoso en las redes sociales o también las circunstancias de la víctima, respecto a si carece de apoyo social o sufre alguna discapacidad, si tiene menores a su cargo, si ha iniciado una nueva relación sentimental así como indicadores como su situación laboral", explican los científicos. En algo más de un mes, los investigadores han recibido 4.266 valoraciones del nuevo formulario, cifra que consideran todavía insuficiente para lograr el modelo predictivo que buscan.²⁵

Junto al ámbito de la violencia de género, también se está valorando la posibilidad de aplicar la justicia predictiva a procesos administrativos, incumplimiento de contratos, aplicación de penalizaciones, faltas leves de tráfico y de tipo laboral.

5 REFLEXIONES FINALES

La implantación de sistemas de "Justicia predictiva" está siendo objeto de valoración en el sistema judicial español. Estas tecnologías consisten en algoritmos capaces de recoger toda la jurisprudencia sobre un asunto y estimar las probabilidades de que los procedimientos judiciales se resuelvan en un sentido u otro, con un grado de eficacia que ronda el 85% de los casos según

²⁴ <https://www.eurocop.com/>

²⁵ También otras propuestas tales como La *startup* de la US *Legal Innovation*, que ha desarrollado una innovadora herramienta capaz de predecir los resultados de litigios en función de las búsquedas y parámetros que introduzca el usuario. Su nombre es Legal Data, y con esta esta tecnología contribuye a la transparencia y democratización de los procesos judiciales. Legal Data es la primera herramienta predictiva basada en Big Data y lenguaje natural que existe en el marco del derecho nacional. Carlos Ibáñez explica que "es una aplicación que está generando mucha expectación y es que te permite conocer en cuestión de segundos el histórico de victorias y derrotas en función de un tema, un Juzgado o Juzgados concretos, el tiempo medio que tardará un litigio en resolverse o cuáles han sido los veredictos de los distintos jueces en diferentes materias a lo largo del tiempo". Además, esta herramienta es capaz de interpretar todos los datos disponibles y ofrecer un análisis que sirva de ayuda a los abogados en la toma de decisiones. Con Legal Data se pretende conseguir que las propias partes sepan de antemano las probabilidades reales de éxito o derrota en un asunto concreto y puedan llegar a acuerdos antes de tener que resolver los problemas en los Tribunales. <http://canalciencia.us.es/presentan-una-nueva-herramienta-para-saber-de-antemano-las-probabilidades-reales-de-exito-o-derrota-en-los-tribunales/>

experiencias ya dadas en otros países europeos y en EE.UU. a la hora de perfilar y cuantificar en sus resultados a los jueces y a los abogados. Ahora bien, la sombra de que bajo el pretexto de reducir el riesgo y el peligro en nuestros sistemas sociales –en concreto, bajo la forma de peligrosidad criminal- se implanten sistemas de justicia predictiva, no está exenta de claras ventajas pero también plantea serios desafíos.

Entre las ventajas aparecen la posible eliminación de incertidumbre jurídica (dado el alto grado de eficacia de los mecanismos), la reducción de una importante cantidad de burocracia (permitiría anticipar aquellos asuntos menores que se repiten con frecuencia en una dirección y 'atascan' los juzgados), la rebaja de costes derivada de lo anterior, y un incremento de las conciliaciones en detrimento de los juicios.

Paralelamente, entre los factores de riesgo hay que subrayar la posibilidad de que la Justicia predictiva traiga aparejado un viraje del 'Imperio de la Ley' a la 'norma aplicada', contribuyendo a un entorno en el que las resoluciones se argumenten cada vez menos. Otros riesgos son el llamado 'efecto rebaño', esto es, que los operadores jurídicos en general, incluidos los propios jueces, entren a cuestionar menos los asuntos y se ciñan por el peso de la matemática; e injusticias por la vía de los posibles sesgos del algoritmo - no siempre es fácil entender por qué el algoritmo llega a las conclusiones que llega- y las cuestiones éticas inherentes. Recurriendo a métodos matemáticos y algoritmos informáticos, que contrastan miles de casos con supuestos de hechos similares, y con una consulta exhaustiva de jurisprudencia, el sistema judicial parece estar llamado a prescindir prácticamente de la argumentación judicial, de la apreciación de la prueba, - ya que hay que prescindir de la subjetividad-. Una justicia predictiva que es rápida, segura, exacta, que no tendrá problemas como lagunas de ley, ni tendrá que enfrentarse a conceptos jurídicos indeterminados, rápida y ágil, podrá convertirse en una tiranía del algoritmo. Si hasta ahora, el imperio de la ley, interpretada siempre por un juez, era la base a la hora de dictar sentencia; si la racionalidad y complejidad que encierra la decisión judicial se hace descansar exclusivamente en fórmulas matemáticas; si las decisiones que antes se adoptaban a partir de principios y analogías son salvadas por un algoritmo, seguramente la propia justicia del sistema deba ponerse en cuestión. La aportación del Profesor Jorge Douglas Price al respecto contribuye a dejar claro en qué consiste la decisión judicial.

Desde el Consejo General del Poder Judicial español, la intención del órgano es que los sistemas de Justicia predictiva puedan servir como "herramienta de ayuda" para juristas y profesionales del derecho en general (abogados, procuradores, etc.), y no afecten directamente a las decisiones de los propios jueces. Los jueces tienen que resolver sobre personas, con sentimientos, y entrando en cada caso concreto, argumentando y examinado las especificidades

concretas del caso. Por ejemplo, Francia, ha prohibido ya por ley el uso de los sistemas de justicia predictiva para perfilar públicamente los resultados de un juez. Sin embargo, el Gobierno de Francia ha comenzado a implementar un proyecto de reforma de la Justicia (2018-2022) que abordará una primera introducción de este tipo de soluciones con al menos 500 millones de euros dirigidos a la digitalización para, entre otras cosas, automatizar cuestiones que llevan mucho tiempo a los jueces. El proyecto francés apunta así a la posibilidad de implementar "mecanismos algorítmicos de resolución de conflictos para asuntos judiciales relativamente menores" por valor de entre 4.000 y 6.000 euros.

La técnica y la tecnología ofrecen ayuda a los operadores jurídicos y a los sistemas judiciales que no debe menospreciarse. Los jueces robots podría ser muy positivo si se trata de ofrecer soluciones que ahorren llevar a las partes a un proceso por cuestiones cuantificables, que obedecen a fórmulas matemáticas, como pueden ser las fijaciones de cantidades, multas, liquidaciones y hasta indemnizaciones e incluso revisiones de renta, que estén sujetas a fórmulas o baremos complejos, y que muchas veces no suponen más que una pérdida de un tiempo preciso tanto para las partes como sobre todo para los jueces. Sin embargo, donde el robot no puede entrar es en la apreciación de pruebas, el componente subjetivo. Son las personas las que juzgan, se trate del asunto que se trate, porque así lo dice la Constitución. Para que la palabra final sobre un asunto jurídico, sea el que sea, la tenga un robot y no una persona, tendría que llevarse a cabo no solo una reforma legal, sino constitucional.

La racionalidad en la decisión; las formas de decisión a partir de principios y analogías; la idea de juicio como decisión; la función de la dogmática; la interpretación y la argumentación junto con la finalidad de seguridad y certeza, han dado lugar a una combinación que puede resultar peligrosa en los sistemas de justicia predictiva. Por ello, es necesario volver la mirada a las obras y teorías de grandes maestros que nos permitirán valorar las tecnologías en su justa medida, sin perder nunca el norte de qué sea la función judicial y que el riesgo y la incertidumbre forman parte de nuestros sistemas sociales –y también jurídicos, por mucho que se los pretenda neutralizar-.

REFERENCIAS BIBLIOGRÁFICAS

BANSACK, Kirk. "Can non-experts really emulate statistical learning methods? A comment on "The accuracy, fairness, and limits of predicting recidivism". Stanford, 2018, pp.1-30. web.stanford.edu/.../cgi.../Bansak_RecidivismAlgo_Comment.p...

BECK, Ulrich. "Teoría de la Modernidad Reflexiva", extraído de *Die Erfindung des Politischen*, Frankfurt, Suhrkamp. En: Josetxo Beriaín (compilador). *Las Consecuencias Perversas de la Modernidad. Modernidad, contingencia y riesgo*. Barcelona: Editorial Anthropos, 1996.

BECK, Ulrich. *La sociedad del riesgo. Hacia una nueva modernidad*. Barcelona: Paidós, 2006.

BERK, Richard; HEIDARI, Hoda; JABBARI, Shahin; KEARNS, Michael and ROTH, Aaron. *Fairness in criminal justice risk assessments: the state of the art*, 2017. arXiv preprint arXiv:1703.09207.

BISWAS, Arpita et al., “Algorithms, fairness, and race: Comparing human recidivism risk assessment with the COMPAS algorithm”, pp.1-4.
jdiessnerlab.ischool.illinois.edu/.../biasescss2018.../SocialPanel_A..

BLAY GIL, Ernesto. “Riesgo y penas comunitarias. El papel del riesgo en los protocolos para la supervisión de penados en la comunidad”. En: J. M. Landa Gorostiza (Ed.), E. Garro Carrera (Coord.), *Delinquentes peligrosos*. Madrid: Ed. Trotta, 2014

DE ASÍS ROIG, Rafael. *Una mirada a la robótica desde los derechos humanos*. Madrid: Dykinson, 2014.

DE GIORGI, Raffaele. “O risco na sociedade contemporânea”. *Revista Sequência*. Revista do Curso de Pós-graduação em Direito da Universidade Federal de Santa Catarina, jun.1994, n.º.28, ano 15, p.45-54.

DE GIORGI, Raffaele. *Direito, Democracia e Risco: vínculos com o futuro*. Porto Alegre: Sergio Antonio Fabris Editor, 1998.

DE GIORGI, Raffaele / LUHMANN, Niklas. “Centro di Studi sul Rischio dell’Università del Salento. Prólogo à edição brasileira. En: Luciano Nascimento Silva e Caterina Del Benea (Coordenadores) *Justiça e Direitos Humanos*, Estudos do Centro di Studi Sul Rischio dell’Università del Salento, Itália, Livro II, Curitiba: Juruá, 2014.

DE MIGUEL BERIÁIN, Íñigo. “Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the Wisconsin v. Loomis ruling, Law”. *Probability and Risk*, Vol. 17, 1, 2018, pp. 45–53.

DOUGLAS PRICE, Jorge Eduardo. *La decisión judicial*. Buenos Aires: Rubinzal - Culzoni Editores, 2012.

DRESSEL, Julia y Hany FARID. “The accuracy, fairness, and limits of predicting recidivism”, *Science Advances* 4 (1):eaao5580 · January 2018. <https://arxiv.org/pdf/1808.09123>

FREEMAN, Katherine. “Algorithmic Injustice: how the Wisconsin Supreme Court failed to protect due process rights in State v. Loomis”, *North Carolina Journal of Law and Technology* (NCJL & Tech.), Vol. 18, 2016, pp. 75 y ss.

GARCÍA-PRIETO CUESTA, José. “¿Qué es un robot?”. En: BARRIO ANDRÉS, M. (Director), *Derecho de los Robots*. Madrid: La Ley, 2018.

GOEL, Sharad. “The Accuracy, Equity, and Jurisprudence of Criminal Risk Assessment”.

<https://5harad.com/papers/RAI-chapter.pdf>

KLEINBERG, Jon et al., “Human decisions and machine predictions”, <https://cs.stanford.edu/people/jure/pubs/bail-qje17.pdf>

- KLEINBERG, Jon; Sendhil MULLAINATHAN and Manish RAGHAVAN. "Inherent Trade-Offs in the Fair Determination of Risk Scores". Cornell University, arXiv.org > cs > arXiv:1609.05807. <https://arxiv.org/abs/1609.05807>
- LUHMANN, Niklas. 'Soziologie der Moral', en N. Luhmann & S. Pfürtnner (eds) *Theorietechnik und Moral*, Frankfurt am Main, 1978, pp. 8–166.
- LUHMANN, Niklas. [1991] Niklas Luhmann. *Soziologie des Risikos* (Berlin, Walter de Gruyter, 1991) *Sociología del riesgo*. trad. J. Torres Nafarrate y otros, (México, Universidad Iberoamericana y Universidad de Guadalajara, 1992).
- LUHMANN, Niklas. *Sociología del riesgo*. Universidad Iberoamericana. México: Triana Editores, 1998.
- LUHMANN, Niklas. *Sistemas sociales. Lineamientos para una teoría general*, trad. de S. Pappe y B. Erker, Barcelona: Anthropos, 1998.
- MARTÍNEZ GARCÍA, Jesús Ignacio. "Pensar el riesgo. En diálogo con Luhmann". *Cuadernos Electrónicos de Filosofía del Derecho*, n°21, diciembre de 2010.
<http://ojs.uv.es/index.php/CEFD>.
- MARTÍNEZ GARCÍA, Ignacio. "Inteligencia y derechos humanos en la sociedad digital". *Cuadernos Electrónicos de Filosofía del Derecho*. 2019, n°40, pp.168-189.
<https://ojs.uv.es/index.php/CEFD/article/view/13846/pdf>
- MILGROM, Paul y TADELIS, Steven. "How Artificial Intelligence and Machine Learning Can Impact Market Design," NBER Chapters, in: *The Economics of Artificial Intelligence: An Agenda*, pages 567-585 National Bureau of Economic Research, Inc., 2018.
- OLIVA LEÓN, Ricardo. "¿Justicia algorítmica ética? (14/12/2018)
<https://www.algoritmolegal.com/tecnologias-disruptivas/justicia-algoritmica-etica/>
- OSWALD, Marion; GRACE, Jamie; URWIN, Sheena & BARNES, Geoffrey C. "Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality", *Information & Communications Technology Law*, 27:2, 2018, pp. 223-250, DOI: [10.1080/13600834.2018.1458455](https://doi.org/10.1080/13600834.2018.1458455)
- PAULUS, N., "Del concepto de riesgo: conceptualización del riesgo en Luhmann y Beck". *Revista Mad*, n°10, Departamento de Antropología, Universidad de Chile, mayo 2004, pp.1-66.
- ROMEO CASABONA, Carlos María. *Peligrosidad y Derecho Penal Preventivo*, Bosch, Casa Ed., Barcelona, 1986.
- ROMEO CASABONA, Carlos María. "Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad". R.E.D.S., n° 13, Julio-Diciembre 2018.
<https://dialnet.unirioja.es/descarga/articulo/6859383.pdf>
- SERRANO MORENO, José Luis. "La sociedad del riesgo y el derecho de la sociedad". *Cuadernos Electrónicos de Filosofía del Derecho*, n°. 21, 2010.
<https://ojs.uv.es/index.php/CEFD/article/download/274/3015>.

SOLAR CAYÓN, José Ignacio. *La Inteligencia Artificial Jurídica. El impacto de la innovación tecnológica en la práctica del Derecho y el mercado de servicios jurídicos*, Cizur Menor (Navarra), Aranzadi, 2019.

TURING, Allan. "Computing machinery and intelligence". *Mind*, vol.59, 1950.

PÁGINAS WEB

Barcelona Declaration for the proper development and usage of Artificial Intelligence in Europe (08.03.2017). <https://www.iiiia.csic.es/barcelonadeclaration/>.

El Consejo de Europa adopta la primera Carta Ética Europea sobre el uso de la inteligencia artificial en los sistemas judiciales (5/12/2018).

<https://diariolaley.laley.es/home/NE0001786943/20181205/ProteccionDatos.aspx>

Comunicación de la Comisión al Parlamento europeo, al Consejo europeo, al Consejo, al Comité económico y social europeo y al Comité de las regiones Inteligencia Artificial para Europa {SWD(2018) 137 final} COM(2018) 237 final. Comisión Europea Bruselas, 25.4.2018.

ec.europa.eu/transparency/.../2018/.../COM-2018-237-F1-ES-M.

European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment

<https://rm.coe.int/ethical-charter-en-for...4.../16808f699c>

Enciclopedia Jurídica y las Ciencias Sociales. <https://leyderecho.org/justicia-predictiva/>

PASCUAL, M. G., "Tendremos que tomar decisiones para controlar la inteligencia artificial", *Diario El País*, 23.07.2017.

https://retina.elpais.com/retina/2017/07/21/tendencias/1500651949_770150.html.

<https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE.html>; <https://www.propublica.org/.../compas-recidivism-risk-score-dat.>; <https://www.propublica.org/.../how-we-analyzed-the-compas-re.>; <https://www.equivant.com/response-to-propublica-demonstratin..>

PLAZA LÓPEZ, J.A., "Lecciones de ética para máquinas que 'piensan' y toman decisiones!", *Diario El País*, 19.12.2017.

RETINA 19/12/2017

https://retina.elpais.com/retina/2017/12/19/innovacion/1513661054_305253.html

<https://securelist.lat> > News

www.ifmpt.de/projects.html

www.elialombardo.it/documenti/siap_n69.pdf

<https://yellrobot.com/shoplifting-detection-software-vaak/>

“This AI Detects Shoplifters Before They Steal, But There's Something Even Creepier About It”, (07.03.2019). <https://www.outerplaces.com/science/item/19294-vaak-ai>

<https://www.eurocop.com/>

JURISPRUDENCIA

Sentencia 13 de julio de 2016: *State v. Loomis*, 881, N.W.2d 749, 7532 (Wis, 2016).