

Inteligencia artificial en el proceso penal: análisis a la luz del Fiscal Watson*

Artificial intelligence in the criminal process: analysis in the light of Prosecutor Watson

*Liced Morales Higueta** , Santiago Agudelo Londoño*** ,*

*Mariana Montoya Raigosa**** , Ana María Montoya Vidales******

Fecha de Recepción: 6 de noviembre de 2021

Fecha de aceptación: 9 de diciembre de 2021

RESUMEN

El funcionamiento de la inteligencia artificial (IA) es desconocido por la mayoría de los operadores jurídicos. Aunque no parezca muy relevante conocer cómo funciona un computador a efectos de implementar su uso, sí resulta determinante develar cómo funcionan los algoritmos que sirven de apoyo al juez para tomar una decisión. En el proceso penal resulta decisivo investigar sobre los algoritmos aparentemente neutrales que permiten el funcionamiento de la IA; esto, en virtud de que la persecución penal es por naturaleza sesgada, por lo que un algoritmo corre el riesgo de

automatizar dichos sesgos y producir sentencias masificadas regidas por probabilidades. El Fiscal Watson emplea una pequeña porción de la información de los delitos reportados, para arrojar una probabilidad de que alguien en concreto pueda ser responsable penalmente. Al ser una herramienta patentada por una compañía privada, una persona tendrá muy pocas probabilidades de controvertir el porcentaje arrojado por el algoritmo.

Palabras clave: Fiscal Watson, heurísticos, inteligencia artificial, presunción de inocencia, sesgo de veracidad.

* Este artículo es producto de una investigación en derecho procesal correspondiente al Semillero de Derecho Procesal de la Corporación Universitaria Lasallista. Directora de investigación: Ana María López Pinilla. Coinvestigadores: Santiago Agudelo Londoño, Mariana Montoya Raigosa, Ana María Montoya Vidales y Liced Morales Higueta. Investigación terminada el 20 de julio de 2020.

** Estudiante de Derecho y miembros del Semillero en Derecho Procesal de la Corporación Universitaria Lasallista. limorales@ulasallista.edu.co

*** Estudiante de Derecho y miembros del Semillero en Derecho Procesal de la Corporación Universitaria Lasallista. saagudelo@ulasallista.edu.co

**** Estudiante de Derecho y miembros del Semillero en Derecho Procesal de la Corporación Universitaria Lasallista. anamontoya@ulasallista.edu.co

***** Estudiante de Derecho y miembros del Semillero en Derecho Procesal de la Corporación Universitaria Lasallista. mamontoya@ulasallista.edu.co

ABSTRACT

The way the Artificial Intelligence (AI) works is unknown by most of legal practitioners. Even if it does not seem very relevant to know how a computer works in order to use it, it is decisive to disclose how the algorithms that support judicial decisions works. In the Criminal Procedure it is decisive to investigate about the apparently neutral algorithms that allows the functioning of the AI; this is because criminal procedure is by nature biased, in consequence, an algorithm has the risk of automatize those biases and

produce mass sentences ruled mainly by probabilities. IBM Watson uses a small portion of the information on reported crimes to produce a result of probability that someone may be the author of the felony. As it is a patented tool by a private company, a person will have very little probabilities of complain the probability thrown by the algorithm against his innocence.

Keywords: *IBM Watson, Confirmation Bias, heuristics, Artificial Intelligence, Presumption of Innocence.*

1. INTRODUCCIÓN

En la actualidad se valora la inteligencia artificial (en adelante IA) porque se piensa que alcanzará grandes logros, permitirá la vida en Marte o la prolongará ilimitadamente; sin embargo, hay poca conciencia del impacto de la IA hoy, tanto así que, estando en la cuarta revolución, no se sabe que un algoritmo podría decidir sobre la división de los bienes tras un divorcio o enviar a alguien a la cárcel.

El avance de la IA inspira cada vez más a las personas a confiar en las bondades que esta trae consigo, pues ayuda a solventar diversas necesidades, incluyendo las jurídicas, aunque sin considerar los riesgos que vienen aparejados. Por su parte, el derecho ha sido, por antonomasia, una práctica humana en la que la labor del juez se erige como un pilar fundamental. La tradicional ciencia jurídica se ha encargado arduamente de dilucidar cómo razona y cómo debe razonar un juez para que sus decisiones sean la más exigua muestra del derecho; no obstante, el uso de la IA puede suponer un cambio radical en la forma en la que funcionan los procesos.

La IA supone retos y trae consigo riesgos a los derechos fundamentales. Su uso indiscriminado puede conducir a sentencias injustas o ilegítimas e inclusive puede conllevar la fosilización del derecho. Sin embargo, resulta iluso rechazar el uso de la IA cuando ya está instalada en la cotidianidad y el proceso judicial.

Este artículo se ocupará de analizar el impacto que podría tener la implementación masiva y sistemática de algoritmos patentados en el proceso penal, como es el caso del algoritmo de IBM Watson. Una apresurada implementación de esta tecnología puede fácilmente justificarse bajo la necesidad de economía procesal y de necesidad de una justicia pronta; no obstante, la rapidez de estas sentencias va en desmedro de su calidad.

Se puede evidenciar, a partir del caso Watson, cómo la práctica del derecho no escapa al estudio de los algoritmos, pues pronto la contraparte en un proceso podrá convertirse en una compleja aplicación altamente tecnificada, que no solo tiene una mejor memoria y capacidad de procesamiento de datos, sino que además goza de la presunción de veracidad derivada de ser una prueba científica.

2. DEFINICIÓN DE INTELIGENCIA ARTIFICIAL (IA)

Definir la IA supone un gran reto puesto que no hay consenso respecto al concepto de *inteligencia*. Entre las teorías más recurrentes se encuentra la triárquica de la inteligencia, desarrollada por el psicólogo estadounidense

Robert J. Sternberg (1985) y que recibe dicho nombre por la consideración del autor de tres tipos de inteligencia. En primer lugar, está la inteligencia analítica o componencial: se trata de la capacidad de captar, almacenar, modificar y trabajar con la información. En segundo lugar, se encuentra la inteligencia práctica o contextual: hace referencia a la capacidad de aquellos seres humanos que se adaptan fácilmente al entorno en el que habitan. Por último, se habla de inteligencia creativa o experiencial: se refiere a la habilidad que nos permite aprender a partir de la experiencia.

Siguiendo esta clasificación, se pueden esbozar tres tipos de IA. En primer lugar, la IA débil: aquella que resuelve problemas definidos con variables definidas; un ejemplo de esta es Deep Blue¹. En segundo lugar, la IA general: permite resolver tareas intelectuales por cualquier ser humano, es capaz de crear juicios y razonar; este tipo de inteligencia supera el Test de Turing²; por último, la IA fuerte: aquella que tiene conciencia de sí misma; no solo realiza procesos cognitivos, sino que puede aprender, resolver problemas, hacer juicios y tener conciencia propia; en la actualidad, este tipo de IA no existe por sus avanzadas características y solo puede encontrarse en el campo de la ciencia ficción.

3. INTELIGENCIA ARTIFICIAL Y ACOMPAÑAMIENTO AL JUEZ

El uso de la IA se ha hecho paulatinamente mayor conforme han avanzado las tecnologías, aunque solo hasta ahora, con ocasión a la Emergencia Sanitaria³, la administración de justicia se ha visto abocada a incursionar abruptamente en la utilización de IA para continuar con la prestación de sus servicios. A continuación, se exponen algunos ejemplos que ilustran la presencia de la IA en el derecho, tanto a nivel judicial como extrajudicial.

La incorporación de la IA en la administración de justicia comenzó poco a poco a través del uso de herramientas tecnológicas muy básicas (como los simples buscadores de jurisprudencia), llegando hasta la estandarización de condenas por medio de asistentes artificiales o la toma de decisiones por parte de algoritmos. Es el caso de China, donde funciona *Xiao Fa*, un asistente artificial que, traducido, significa “derecho pequeño”. “El primer piloto fue testado en el año 2004 en Shandong en temas penales. El sistema se diseñó para analizar

- 1 Es un ordenador de IBM que en 1997 logró vencer al campeón mundial de ajedrez después de un juego de seis partidos: dos victorias para IBM, uno para el campeón y tres empates.
- 2 Es una prueba de la habilidad de una máquina para exhibir un comportamiento inteligente similar al de un ser humano, de tal manera que, interactuando con ella en una conversación, alguien pueda determinar si su interlocutor es una máquina o una persona. Si luego de 5 minutos de conversación (el experimento original está pensado para conversaciones de texto) el humano no puede decir con certeza si su interlocutor es una máquina o una persona, la máquina aprobó el test.
- 3 Dispuesto en Colombia inicialmente por el Decreto 417 del 17 de marzo de 2020.

100 crímenes y la idea era estandarizar condenas y generar automáticamente borradores de sentencias”. Esto dio paso a la creación del “primer tribunal virtual o cibernético que se estableció en la ciudad china de Hangzhou en agosto de 2017. Después se abrieron salas similares en Pekín y Guangzhou. Estos llamados tribunales de internet son competentes para determinados asuntos relativos a las operaciones en red, comercio electrónico y propiedad intelectual” (The Technolawgist, 2020).

La intervención de la IA no se presenta solo como una IA débil que recopila información, sino que actualmente pueden encontrarse sistemas que apoyan al juez para dictar sentencia, o que ayudan a resolver un conflicto en el marco del escenario extrajudicial. En Estonia se han implementado pruebas piloto con jueces robot en procesos cuya cuantía no supera los 7000 euros. Su propósito es simplificar los servicios ofrecidos a la población. El gran reto que supone es lograr la aceptación de los usuarios, que creen que la calidad del servicio se reduciría si fuera prestado con ayuda de IA. Para estas pruebas, las partes aportarán los documentos requeridos en la plataforma y la IA decidirá sobre el caso. Es de resaltar la presencia humana en este proceso, pues la intervención de la IA no es absoluta, ya que su decisión estará sometida a revisión por un juez humano (The Technolawgist, 2019).

4. DIFERENCIAS EN LA TOMA DE DECISIONES DENTRO DE UN PROCESO PENAL CON PRESENCIA DE IA A LA LUZ DEL FISCAL WATSON

Actualmente en Colombia no existe una herramienta que sustituya al juez en la labor de tomar una decisión; la tarea de condenar o absolver a una persona en un proceso penal es una cuestión principalmente humana. A pesar de ello, no se puede desconocer que la injerencia de la IA puede resultar decisiva a la hora de fallar. La IA goza de una presunción de legitimidad que viene dada por la confianza en el método científico y en los resultados obtenidos a través de la razón.

No se desconoce que una prueba científica, en la mayoría de los casos, resulta más idónea para probar un hecho y disminuir el error⁴. Este caso supone una optimización del sistema judicial: a la vez que disminuye el riesgo de error, aumenta la eficiencia procesal. Por su parte, el proceso penal se ha visto altamente impactado por la incursión de la tecnología. El papel de la IA varía según

4 Esto se puede ver con el cambio de paradigma procesal que supuso la incorporación de la prueba de ADN en los procesos de filiación. Antes de la existencia de esta prueba, la filiación se determinaba principalmente a partir de los testimonios y de los indicios que llevaran a inferir que una persona era hija de otra.

el momento procesal al que se haga referencia, pudiéndose distinguir, *grosso modo*, el panorama que se expone a continuación.

4.1. La IA en la investigación de los hechos

La idea de que todos los delitos dejan una huella y que es posible rastrearla a partir de métodos científicos ha nutrido la literatura y la imaginación forense⁵, a la vez que ha demandado el uso de las nuevas tecnologías; esto ha convertido la ciencia forense en un laboratorio científico.

El uso de la tecnología en la fase de investigación es en principio de gran ayuda, pues hace posible identificar un mayor número de elementos materiales probatorios, procesarlos y almacenarlos para su posterior utilización. Estos métodos de investigación no son exclusivos de las fases preliminares de la investigación. Uno de los mayores impactos del uso de la tecnología está en el derecho penitenciario con el Proyecto Inocencia, fundado en 1992 para asistir a personas erróneamente condenadas y demostrar su inocencia con el uso de la tecnología del ADN. En el periodo comprendido entre 1989 y 2014, se documentaron 365 casos de exoneraciones por exámenes de ADN, incluyendo 20 casos de personas condenadas a pena de muerte. De esas cifras, se tiene que el 25% había aceptado su responsabilidad y al 11% realizaron allanamientos. Inclusive, esta tecnología ha ayudado a descubrir la verdad aun después de haber sido condenados a pena de muerte (The Innocence Project, s.f.).

En un sistema penal que privilegió el eficientismo de las sentencias obtenidas vía preacuerdos o allanamientos, la incorporación de la tecnología puede resultar determinante, especialmente porque permitiría fácilmente sustentar una sentencia y disminuiría las posibilidades de que un indiciado acuda fácilmente a los recursos tecnológicos para oponerse al valor suasorio de la prueba científica.

Otro ejemplo de lo anterior es Crime Center, un *software* diseñado para la policía de Estados Unidos que permite organizar un equipo de trabajo, asignando a cada funcionario una tarea de investigación que irá reportando a medida que la va cumpliendo. A su vez, toda la información se va centrando en una base de datos, la cual puede ir alimentándose con información de otras bases de datos que resulten relevantes para la investigación: por ejemplo, infracciones de tránsito, antecedentes penales, titularidad de un vehículo, entre otros (Crime Center Software, s. f.).

5 Esta situación puede ilustrarse con el llamado “efecto CSI”, en virtud del cual se espera de la investigación que haga uso de las herramientas de última tecnología disponible; a la vez que se aumenta la creencia en los postulados que son obtenidos a través de estos métodos (Goodman-Delahanty y Tait, 2016).

El IBM's Watson Debater es una herramienta a la que se le plantea un tema de debate; en respuesta, analiza los textos disponibles en la web sobre la materia, selecciona los argumentos que le parecen más sólidos —probablemente, entre otras razones, por ser los más repetidos— y los expresa en un lenguaje natural (Nieva, 2018, p. 30).

En el caso colombiano, la Fiscalía General de la Nación (FGN) cuenta con un *software* conocido como el Fiscal Watson. Según la entidad, este sistema puede definirse así:

Watson es un concepto que encapsula diversas herramientas y servicios, el cual ha sido elaborado por el fabricante IBM como parte de su oferta en estrategias de análisis cognitivo e inteligencia artificial. De manera particular, la Fiscalía General de la Nación ha implementado y desarrollado una serie de reglas y patrones mediante análisis gramatical y minería de texto, que permite de manera automática (i) realizar extracción de entidades e (ii) identificar nuevas entidades mediante técnicas de *machine learning* sobre todo el cúmulo de información que reposa en las bases de datos de la entidad. Este desarrollo tiene como finalidad fortalecer las labores de investigación realizadas al interior de la entidad, por parte de fiscales, investigadores y policías judiciales. (FGN, 2020)

Según la página oficial de la FGN, el sistema Watson comenzó a operar en Colombia a partir del 2018, fecha en la que se formuló imputación y “logró” que se dictara medida de aseguramiento en centro carcelario contra los coautores de 22 delitos (2018). Lejos de buscar claridad sobre el funcionamiento de Watson, puede decirse que esta herramienta opera a partir de la minería de datos, entendida como un proceso de identificación de una masa de datos recolectada para descubrir y extraer patrones.

Tal y como lo señaló la vicefiscal general María Paulina Riveros, Watson permite “esculcar” las bases de datos del Sistema Penal Acusatorio (SPOA) que del 2005 al 2018 contenía trece millones de denuncias, para correlacionarlas y hacer análisis de contextos sobre aspectos relevantes como *modus operandi*, características morfológicas, tipos de armas, entre otros (El Espectador, 2018). Cabe resaltar la respuesta oficial de la FGN: es esta entidad la que define la serie de reglas y los patrones que posteriormente serán identificados y extraídos por Watson.

4.2. Medidas de aseguramiento

El escenario de las medidas de aseguramiento⁶ resulta más complejo, puesto que el análisis que realiza el juez no se limita a una valoración probatoria de la ocurrencia de un hecho, sino que se trata de una predicción de hechos que por su naturaleza no han ocurrido, ni se sabe si ocurrirán. El papel de la IA aquí es determinante, especialmente porque, como denuncia Nieva Fenoll (2018), existe una tendencia a “querer valorar el riesgo a través de variables estadísticas” (p. 62). A la hora de hacer predicciones con base en estadísticas, la IA puede resultar muy útil, principalmente porque no cae en los heurísticos, y en segundo lugar porque es menos susceptible de ceder ante presiones externas para decidir en uno u otro sentido.

Al común de las personas les resulta muy difícil pensar estadísticamente. Varios estudios han mostrado que los seres humanos tienen una confianza excesiva en lo que creen saber, y una ignorancia del mundo en el que viven, a la vez que subestiman el papel del azar en los acontecimientos (Kahneman, 2019, p. 27). Ante cuestiones difíciles, la mente toma atajos simplificadores, también conocidos como *heurísticos* (Tversky y Kahneman, 1974); y estos, aunque permiten estimar la probabilidad de un evento incierto, se basan en datos de validez limitados que pueden llevar a errores. Dentro de los sesgos más relevantes se encuentran:

- a. Heurístico de representatividad: aparece para responder preguntas como ¿cuál es la probabilidad de que el objeto A pertenezca a la categoría B? Este heurístico explica el uso frecuente de los estereotipos; estamos más propensos a creer que un objeto A pertenece a una categoría B si A cumple con el estereotipo del grupo, con independencia de la estadística de probabilidad
- b. Heurístico de disponibilidad: ocurre cuando un sujeto debe estimar la frecuencia de una clase o la plausibilidad de un desarrollo particular. Se evalúa la frecuencia de un evento por la facilidad con la que nos llega a la mente, los eventos que se recuerdan con más facilidad son aquellos que juzgamos como más probables. Este heurístico es el que hace parecer que la criminalidad ha aumentado cuando se puede pensar con facilidad en casos conocidos de víctimas de delitos recientes, sin que importe si la tasa real de criminalidad se haya visto afectada.

6 La medida de aseguramiento es un mecanismo preventivo y excepcional mediante el cual una persona es privada de alguno de sus derechos fundamentales con el objetivo de evitar un mal mayor, que de conformidad con el artículo 308 del Código de Procedimiento Penal puede ser: evitar la obstrucción a la justicia, evitar un peligro a la seguridad de la sociedad o de la víctima, evitar que el imputado no comparezca al proceso o se sustraiga del cumplimiento de una obligación.

- c. Heurístico de ajuste y anclaje⁷: se presenta cuando las personas deben tomar decisiones bajo incertidumbre, para lo cual utilizan la información disponible externamente como anclas, y luego ajustan este valor de acuerdo con sus propias estimaciones. Cuando la gente hace estimaciones comenzando con un valor inicial dado, este determina la respuesta final.

Merece la pena ilustrar este fenómeno con el ejemplo que traen los autores. A varios individuos se les pide estimar el porcentaje de países africanos que pertenecen a la ONU. Primero le pidieron a cada uno hacer girar una ruleta que indicaba un número entre 1 y 100 (el valor del ancla es aleatorio, lo que indica que no es necesaria una relación entre el valor por estimarse y el valor del ancla). A continuación les preguntaron si sus estimaciones eran superiores o inferiores, y por último, les pidieron que dieran su estimación numérica en porcentaje. Los resultados arrojaron que aquellos que obtuvieron el número 10 tuvieron una estimación mediana del 25%, mientras que los que obtuvieron el número 65 registraron una estimación mediana de 45%. Este estudio concluyó que el ancla (que era aleatoria) influyó en las estimaciones finales de los encuestados (Tversky y Kahneman, 1974).

El Fiscal Watson proporciona al juez una información que impactará la forma en la que se evalúa el material probatorio y se toma una decisión; en ese sentido, la información provista por Watson hará las veces de ancla a partir de la cual el juez ajustará su razonamiento. Como se analizará más detalladamente, el anclaje que provee Watson puede estar sesgado, lo que lo convierte en una herramienta no idónea para superar este sesgo.

No obstante lo anterior, la decisión de imponer o no una medida de aseguramiento corresponde principalmente a un juez, quien decide usualmente frente a un imputado que puede ver y escuchar directamente. Esta intermediación existente permite en cierta medida humanizar las decisiones judiciales y entablar un vínculo empático, concepto que puede medirse a través del Test de Asociación Implícita (TAI) desarrollado inicialmente por Greenwald, McGhee y Schwartz (1998). En 1952 la Orquesta Sinfónica de Boston utilizó un método “ciego” para seleccionar a los músicos, quienes audicionaron detrás de un telón de modo que el género no fuera un factor por tener en cuenta para la selección. Asimismo, muchas empresas adoptaron este método para realizar las contrataciones de sus empleados de forma más objetiva (Hewlett, Rashid y Sherbin, 2018). Aplicado al ámbito judicial, podría decirse que un juez que tome la decisión sin ver al indiciado podría evitar este sesgo de veracidad, pues el contacto directo entre ambos no deriva en las decisiones más acertadas, sino más bien intensifica las

7 Tal y como se ha evidenciado, el anclaje es el heurístico por excelencia para los jueces y magistrados. Algunos estudios han constatado el anclaje en la condena impuesta por el juez con relación a la petición del fiscal, y se estima que este impacto se presenta en un 58,3% de las sentencias (Fariña, Arce y Novo, 2002, p. 40).

probabilidades de error. La figura del *sesgo de veracidad* parte del supuesto de que los demás están diciendo la verdad, y si son concordantes con su lenguaje verbal, hay una tendencia a creerles (Gladwell, 2020).

La humanidad de los jueces tiene un papel determinante a la hora de tomar una decisión; factores como el cansancio o el hambre pueden definir la decisión judicial⁸. Los jueces agotados tienden a decantarse por la decisión más fácil, que —contraria a la naturaleza excepcional que debería tener la prisión preventiva— suele ser negar la libertad (Kahneman, 2019, p. 64).

4.3. ¿Puede Watson corregir los sesgos?

Los escenarios de predicción de algún resultado —como lo es la decisión de imponer una medida de aseguramiento— resultan altamente propicios para la aplicación de *machine learning* o aprendizaje automático⁹. Al respecto, un estudio llevado a los juzgados de Nueva York demostró que las predicciones que hacen los jueces al evaluar el riesgo de fuga y otorgar o negar la libertad provisional son errores que pueden ser corregidos a través del lenguaje automático (Kleinberg et al., 2018).

Resulta tentador pensar que un algoritmo, a través de un complejo pero eficiente análisis de datos, pueda disminuir el margen de error en que incurren los jueces de forma sistemática. Un algoritmo no tendría los problemas de cambiar de decisión motivado por el cansancio o el hambre, tampoco incurriría en el sesgo de veracidad. Sin embargo, existen otros riesgos en la aplicación de la tecnología que pueden resultar igual de perjudiciales que los sesgos del juez, a saber: el sesgo del algoritmo.

Los modelos matemáticos tienen un alcance epistemológico limitado. Por un lado, pueden reproducir sesgos basados en prejuicios; como el precedente de la sentencia *State vs. Loomis* en el uso del algoritmo COMPAS para valorar el riesgo de reincidencia. Algunos de los puntos controversiales alegados por la defensa consistían en que se violaba el debido proceso puesto que la sentencia no podría ser individualizada porque se basaba en información sobre un conjunto

8 Un estudio llevado a cabo por Danziger, Levav y Avnaim-Pesso (2011) analizó la actividad judicial de seis jueces de Israel que recibían aleatoriamente distintos casos y debían definir si conceder o no la libertad condicional; observando las pausas para comer, pudieron concluir que durante unas dos horas antes de la próxima comida, la cantidad de libertades que eran otorgadas descendían de 65% hasta casi 0, y volvían a subir abruptamente después del descanso (p. 1).

9 El aprendizaje automático puede definirse como el campo de estudio de IA que permite a los algoritmos aprender sin estar expresamente programados para ello. Esto se hace a través de la recolección de nuevos datos, los cuales utiliza para modificar los resultados que arroja; para ello usa varias disciplinas como la estadística, la teoría de algoritmos, el análisis funcional y la optimización (Muñoz, 2014).

de personas mucho más amplio. En segundo lugar, se alegaba la imposibilidad de ejercer una defensa efectiva al desconocer los patrones de funcionamiento de COMPAS¹⁰, algoritmo que está patentado y, por tanto, la información de funcionamiento es de conocimiento exclusivo de sus propietarios.

Según los cálculos de la Fiscalía, en el SPOA se tiene, para el periodo del 2005 al 2018, un total de trece millones de denuncias. Partiendo de las declaraciones oficiales, ese número de denuncias compone la “clase de referencia”¹¹, sin que se pueda determinar la forma en la que se clasifican esas denuncias para conformar otras clases de referencia (por ejemplo, dividir las por delitos, número de autores o partícipes, lugar de los hechos, entre otros). Determinar cuál es la clase de referencia por emplear resulta un problema estadístico definitivo para alcanzar un resultado.

El ejemplo propuesto por Allen y Pardo (2007, p. 109) ilustra bien esta situación: suponga que un testigo vio un bus golpear un carro, pero no recuerda el color del bus; asuma que Blue Company tiene 75% de los buses de la ciudad y que la Red Company tiene el 25% remanente. La vista prevalente en la literatura legal del valor de la probabilidad estaría determinada por la proporción de Blue Company frente a la Red Co. Pero suponga que Red Co. tiene el 75% de los buses de todo el país, ahora la proporción se revierte. Y lo haría otra vez si la Blue Co. tuviera el 75% de los buses del Estado. Y se volvería a revertir si Red Co. tuviera el 75% de los buses recorriendo la calle donde ocurrió el accidente, y así sucesivamente.

En estadística, el problema de la clase de referencia es el problema de decidir qué clase usar cuando se calcula la probabilidad aplicable a un caso en particular. En el ámbito legal, el caso *Shonubi vs. EE. UU.* es paradigmático en el uso desafortunado de los modelos matemáticos. En este caso, Shonubi fue detenido en el aeropuerto neoyorkino de JFK llevando consigo 427,4 gramos de heroína. Se constató en el proceso que Shonubi había hecho en el pasado siete viajes. El Tribunal del Distrito condenó a Shonubi por el transporte promedio en gramos traficados por otros nigerianos volando desde el mismo aeropuerto y utilizando

10 En una reciente carta dirigida a la editorial científica Springer Nature, más de veinte investigadores solicitaron a dicha editorial no publicar un artículo que presentaba un sistema de reconocimiento facial que pretendía predecir si alguien es un criminal. Dentro de los argumentos esbozados se encuentra que los programas de aprendizaje automático no son neutrales; las agencias de investigación reflejan los incentivos y las perspectivas de aquellos en posiciones privilegiadas. Especialmente, hacen énfasis en que un algoritmo que persiga la criminalidad no podrá evitar ser racista porque la misma persecución del delito está sesgada. A pesar de que la carta se refiere a un artículo en específico, puede evidenciarse que se trata de una práctica sistemática en la cual hay un creciente desarrollo de tecnologías encaminadas a la incriminación ligera y tecnicizada de un grupo de personas que tradicionalmente ha sido vulnerada por las políticas criminales. <https://medium.com/@CoalitionForCriticalTechnology/abolish-the-techtoprisonpipeline-9b5b14366b16>. En virtud de esta petición, la revista *Springer* se abstuvo de realizar la publicación del artículo

11 La clase de referencia alude al valor de base que se toma como referencia para calcular la probabilidad en un caso en particular.

el mismo *modus operandi*. Esto arrojó con un 99% de probabilidad que el acusado había transportado al menos 2090 gramos de heroína. En la última instancia se revocó esta decisión, y Shonubi resultó condenado por la cantidad que transportaba cuando fue detenido (1973). A efectos legales, esto implica que la información que sirve de base determina el resultado al que pueda llegar un algoritmo. El sistema Watson se alimenta a partir del crimen reportado (lo que constituirá la clase de referencia), dejando de lado la criminalidad oculta, es decir, aquella que no llega a estar en conocimiento directo de las autoridades.

En Colombia se estima que la criminalidad oculta oscila entre el 70% y el 80% (Restrepo y Cuéllar, 2004). Otras fuentes que estudian los índices de impunidad, como la ONU o el Índice Global de Impunidad Colombia 2019, agrupan dentro de la impunidad tanto la criminalidad oculta (es decir, aquellos delitos que no son puestos en conocimiento de las autoridades) como la impunidad penal (se refiere a los delitos que ingresan al conocimiento de las autoridades pero que por diversos motivos no reciben una respuesta de fondo) (Le Clercq et al., 2019). Si, adicionalmente, tenemos en cuenta que en Antioquia la confianza en las instituciones de justicia es del 37%, y que el índice de acceso a la justicia es de 5,47 (Le Clercq et al., 2019, p. 32), como resultado se obtendrá que una clase de referencia recoge, en el mejor de los casos, el 30% de los posibles delitos por investigar.

Por otro lado, existe una tendencia de que los delitos que sean registrados en las estadísticas policiales no lo sean por su gravedad, sino por el hecho de que las víctimas cuenten con un conocimiento razonable de la identidad de los agresores que faciliten su captura por las autoridades (Restrepo y Cuéllar, 2004, p. 5). En este sentido, un estudio mostró que, de 222 expedientes analizados por terrorismo, en 35% no se identificó un sindicado. Por lo anterior, el sistema Watson tendrá poca o ninguna incidencia directa frente a la criminalidad oculta. Frente a la calidad de la información aportada, se hace necesario resaltar que el sistema Watson recibe esa información sin tener un criterio valorativo para filtrarla según calidad. En Estados Unidos se estima que la equivocada identificación por parte de los testigos es la principal causa de las sentencias erróneas, y en el 75% de las condenas que fueron revocadas con prueba de ADN, la causa fue el error en la identificación por parte de las víctimas o testigos (The Innocence Project, s.f.). Por esto, Watson sería una herramienta que agrupe muchos datos indistintamente de su veracidad y los reproduzca para fundar una sentencia.

5. RETOS DE LA APLICACIÓN DE LA INTELIGENCIA ARTIFICIAL

La IA ha permeado positiva y negativamente muchas áreas de la vida del ser humano a nivel práctico; sin embargo, a nivel teórico no se le ha dado suficiente relevancia puesto que las personas se encuentran cómodas y acostumbradas a esta. Lo cierto es que, aunque actualmente no están claros los alcances de la IA, es necesario definir sus límites. No se le puede dar autonomía frente a la vigilancia de los ciudadanos, la defensa militar o la última decisión en un juicio. Algunos de los retos de la aplicación de la IA se enuncian a continuación.

5.1. Legitimidad

Un rasgo propio de la aplicación de la IA dentro de la administración de justicia es la falta de confianza de los usuarios en estas herramientas. Resulta paradójico que las personas confíen a la tecnología su información más importante y privada, pero desdeñan de esta cuando se trata de aplicarla al proceso. Esto se debe a la empatía, pues del juez humano se espera que no aplique todo el rigor de las leyes, sino que las adapte a lo que sea justo en cada caso (Nieva, 2018). Si contemplamos este pensamiento generalizado, podríamos encontrar un amplio margen de interpretación por parte del juez, donde tiene mayor peso la discrecionalidad. Es difícil asimilar que quien juzga a las personas no sienta empatía hacia ellas, sino que se alimenta del poder de su base de datos. Es la empatía la que hace que el sistema de justicia genere mayor confianza. Paradójicamente, es esa falta de empatía la que permite que se eliminen sesgos.

Finalmente, en casos de algoritmos que se encuentran patentados, como el caso del Fiscal Watson que pertenece a IBM, el conocimiento del funcionamiento del programa está protegido por los derechos de autor. Esto significa que el sistema judicial que haga uso de este tipo de algoritmos vulnera el principio de publicidad al no dar a conocer a las partes el detallado funcionamiento del algoritmo; esto lleva aparejada la vulneración al derecho de defensa, toda vez que alguien que resulte señalado por el Fiscal Watson, difícilmente podrá acceder a las razones que llevan al algoritmo a señalar como posible responsable.

5.2. La dinamización del derecho

Desde una postura sociológica¹² podría afirmarse que el derecho se transforma y evoluciona por mor de los nuevos acontecimientos dados en la sociedad.

12 Lawrence M. Friedman (2002), en su libro *Law in America*, muestra cómo a medida que la sociedad se va transformando, también se transforman o surgen nuevas leyes. Así pues, el autor afirma que la base del derecho es la sociedad.

Ferdinand Lassalle (1862) menciona que existe en un país “una fuerza activa e informadora, que influye de tal modo en todas las leyes promulgadas en un país determinado, que las obligaban a ser necesariamente, hasta cierto punto, lo que son y como son, sin permitirles ser de otro modo”. Esa fuerza activa e informadora son los factores reales del poder que rigen el seno de una sociedad determinada. Anteriormente, estos se referían a la monarquía, la aristocracia, la burguesía, los banqueros, la conciencia colectiva, la cultura general y, finalmente, la clase obrera; pues todos ellos conformaban una parte esencial del derecho. Pero los factores reales del poder tuvieron que evolucionar, transformarse, reinventarse para no quedar obsoletos, y actualmente estos son traducidos en las ramas del poder público, la democracia, la economía, las clases sociales, los ciudadanos, los medios de comunicación y, más recientemente, la tecnología.

Sobre la fosilización del derecho cuya aplicación queda delegada a la IA, aparecen dos posturas. La primera de ellas predica que un derecho dinámico necesita humanos, pues ellos pueden adaptar la norma a los cambios y a la realidad social, ya que están en la capacidad de modificar sustancialmente el derecho. La segunda postura sostiene que el derecho puede aplicarse a través de la IA, la cual, al contar con un sistema de aprendizaje automático¹³, puede aprender a partir de la información que va recibiendo (IBM, s.f.). Se puede decir que, si bien la aplicación de la IA puede permitir agilizar procesos jurídicos, esta no será la fuente de los cambios normativos, por lo que la injerencia humana resulta imprescindible para replantear el marco en el cual opera la IA¹⁴.

6. CONCLUSIONES

En lo atinente al uso de los algoritmos en la etapa de investigación en los procesos penales, concluimos que la utilización de estos, cuando están patentados, como es el caso del Fiscal Watson, vulneran el principio de publicidad al no poder ser conocidos por las partes los detalles de su funcionamiento. En consecuencia, representa una vulneración al derecho de defensa en la medida en que resulta más difícil defenderse de premisas desconocidas.

Si bien existe un margen de error judicial inherente a la naturaleza humana del juez, el Fiscal Watson no es una herramienta idónea para disminuir dicho error. Tanto porque se basa en una información limitada, por lo cual tampoco sirve para disminuir directamente la cifra oculta de la criminalidad, como

13 Véase nota 13 de este artículo.

14 Tal y como lo plantea Jordi Nieva (2018), la IA debe estar orientada a vencer el tremendo peso regresivo que suele tener entre los juristas la “tradicción”. La IA contribuye y debe contribuir a la evolución, y no debe servir para interrumpirla. Debe ser útil, asimismo, para descubrir y vencer los errores más frecuentes de nuestro pensamiento.

porque puede reproducir los sesgos producto de la programación del algoritmo, con el agravante de gozar de la presunción de veracidad por tratarse de una herramienta matemática.

El riesgo de que el derecho se fosilice radica en que la máquina no se replantea ni rectifica la decisión, sino que decide con la información de su base de datos; en cambio, el juez humano puede rectificar y cambiar de pensamiento y así dinamizar el derecho. El aprendizaje automático podrá permitir que la IA esté siempre actualizada a los cambios del derecho, pero difícilmente estos cambios nacerán de esta. En una sociedad democrática, la aplicación de la IA en el proceso judicial no puede hacerse a costa de las libertades ciudadanas. Es una función humana crear lo sustancial del derecho y su dinamización no dependerá de la IA, pues, por demás, carecería de legitimidad en Colombia para hacerlo.

7. REFERENCIAS

Ad Cordis (2020). Qué es ODR: Online Dispute Resolution. <https://adcordis.com/que-es-odr-online-dispute-resolution/>.

Crime Center Software (s.f.). *Police Case Management Center*. <https://crime-center.com/>

Danzinger, S., Levav, J. y Avnaim-Pesso, L. (2011). Extraneous Factors in Judicial Decisions. *Proceedings of the National Academy of Sciences of the United States of America*, 108(17), 6889-6892. <https://doi.org/10.1073/pnas.1018033108>

Digital Trends. (2015, 13 de julio). *Modria, un software para resolver divorcios y otras disputas jurídicas vía internet*. <https://es.digitaltrends.com/computadoras/modria-un-software-para-resolver-divorcios-y-otras-disputas-juridicas-via-internet/>.

El Espectador. (2018, 31 de agosto). *Así funciona Watson la inteligencia artificial de la Fiscalía*. <https://www.elespectador.com/noticias/judicial/asi-funciona-watson-la-inteligencia-artificial-de-la-fiscalia/>

deEstefan, S. (2018) ¿Qué es el Test de Turing (y, Google: qué has hecho)? Techcetera. <https://techcetera.co/que-es-el-test-de-turing/>

Fariña, F., Arce, R. y Novo, M. (2002). Heurístico de anclaje en las decisiones judiciales. *Psicothema*, 14(1). <http://www.psicothema.com/pdf/684.pdf>

Fiscalía General de la Nación (2018, 7 de noviembre). Con ayuda del sistema Watson la Fiscalía prepara 10 nuevas imputaciones por hurto. <https://www.fiscalia.gov.co/colombia/seccionales/con-ayuda-del-sistema-watson-la-fiscalia-prepara-10-nuevas-imputaciones-por-hurto/>.

Fiscalía General de la Nación, respuesta a derecho de petición, 9 de junio de 2020.

Friedman, L. (2002). *Law in America*. Nueva York: The Modern Library.

Gladwell, M., (2020). *Hablar con extraños. Por qué es crucial (y tan difícil) leer las intenciones de los desconocidos*. Audible. (Original publicado en 2020).

Goodman-Delahunty, J., Tait, D. (2006, 2 de diciembre). CSI Effects on Jury Reasoning and Verdicts. En D. Tait y J. Goodman-Delahunty (Eds.), *Juries, Science and Popular Culture in the Age of Terror*. https://doi-org.crai-ustadigital.usantotomas.edu.co/10.1057/978-1-137-55475-8_12

Greenwald, A. G., McGhee, D. E. y Schwartz, J. L. (1998). Measuring Individual Differences in Implicit Cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74(6), 1464-1480. <https://pdfs.semanticscholar.org/10cc/2d53ff8349d3432b8f822d58e4ddee3d475e.pdf>

Hewlett, S., Rashid, R. Y Sherbin, L. (2018). Cómo desmontar los sesgos y potenciar el valor. <https://www.harvard-deusto.com/como-desmontar-los-sesgos-y-potenciar-el-valor>.

IBM (s.f.) ¿Qué es Machine Learning? <https://www.ibm.com/co-es/analytics/machine-learning>

IBM 100-Deep Blue. <https://www.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/>

International Journal for Court Administration (2017, 11 de julio). Can artificial intelligence and online dispute resolution enhance efficiency and effectiveness in courts. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2999339

Kahneman, D. (2019). *Pensar rápido, pensar despacio*. Penguin Random House, Debolsillo.

Kleinberg J., Himabindu L., Leskovec J., Ludwig J. y Mullainathan, S. (2018). Human Decisions and Machine Predictions. *Quarterly Journal of Economics*, 133, 237-293. <https://doi.org/10.1093/qje/qjx032>

- Lara, L. (2012). La relación entre derecho y sociedad. <http://derechocomparado-uniandes.blogspot.com/2012/05/la-relacion-entre-derecho-y-sociedad.html>
- Lasalle, F. (2013). ¿Qué es una Constitución? SKLA. (Original publicado en 1862).
- Le Clercq, J., Chiádez, A., Gómez, D., Rodríguez, G. y Valencia, L. (2019). Índice Global de Impunidad Colombia 2019. La impunidad subnacional en Colombia y sus dimensiones IGI-Col 2019. Puebla: Universidad de las Américas. <https://www.casede.org/index.php/biblioteca-casede-2-o/justicia/480-indice-global-de-impunidad-de-colombia/file>
- Muñoz, A. (2014). Machine learning and optimization. https://www.cims.nyu.edu/~munoz/files/ml_optimization.pdf
- Nieva Fenoll, J. (2018). *Inteligencia artificial y proceso judicial*. Barcelona: Marcial Pons.
- Restrepo, E. y Cuéllar, M. (junio de 2004). Impunidad penal: mitos y realidades [Documento CEDE 2004-24]. <https://core.ac.uk/download/pdf/6689145.pdf>
- Ronald J. Allen and Michael S. Pardo. (2007). The Problematic Value of Mathematical Models of Evidence. *The Journal of Legal Studies*, 36(1), 107-140. <https://www.jstor.org/stable/10.1086/508269>
- Semillero de Derecho procesal @semillero derecho procesal (2020, 19 de mayo). ¿Si usted fuera a ser condenado preferiría escuchar su sentencia por parte de: juez humano o juez robot? https://www.instagram.com/s/aGlnaGxpZ2hoOjE3ODQzMzkyNDA5MDUyMzAo?igshid=1xlovywy49dd5&story_media_id=2312813828411513579_30109854623
- Sternberg, R. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Nueva York: Cambridge University Press.
- Superintendencia de Industria y Comercio. (s.f.) <https://www.sic.gov.co/noticias/la-herramienta-sic-facilita-celebra-dos-anos-solucionando-la-vida-de-consumidores-y-proveedores>
- The Innocence Project (s. f.). *DNA Exonerations in the United States*. <https://www.innocenceproject.org/dna-exonerations-in-the-united-states/>

The Problematic Value of Mathematical Models of Evidence. Ronald J. Allen and Michael S. Pardo.

The Technolawgist (2019, 12 de junio). Estonia se prepara para tener “jueces robot” basados en inteligencia artificial. <https://www.thetechnolawgist.com/2019/06/12/estonia-se-prepara-para-tener-jueces-robot-basados-en-inteligencia-artificial/>

The Technolawgist (2020, 28 de abril). China, el monstruo mundial en inteligencia artificial que utiliza cientos de jueces robot <https://www.thetechnolawgist.com/2019/12/13/china-el-monstruo-mundial-en-inteligencia-artificial-que-utiliza-cientos-de-jueces-robot/>

Tversky, A. y Kahneman D. (1975). Judgment under Uncertainty: Heuristics and Biases. En D. Wendt y C. Vlek (Eds.), *Utility, Probability, and Human Decision Making. Theory and Decision Library* (vol 11). Dordrecht: Springer.

Tversky, A. y Kahneman, D. (1974, 27 de septiembre). Judgment Under Uncertainty: Heuristics and Biases. *Science*, 185. <https://www2.psych.ubc.ca/~schaller/Psyc590Readings/TverskyKahneman1974.pdf> United States District Court, Eastern District of New York. (4 de agosto de 1995).