

INTELIGENCIA ARTIFICIAL EN LOS TRIBUNALES: EL JUEZ ELECTRÓNICO COMO FORMA DE EQUIDAD

<https://doi.org/10.55839/2318-8650RevParRPv33n1pa2-22>

Jorge Szeinfeld¹
Thomas Law²
Víctor Gabriel Rodríguez³
Sebastião Sérgio da Silveira⁴

RESUMEN

Los mecanismos de Inteligencia artificial hoy pueden expresarse en lenguaje natural, además de interpretar textos con gran habilidad. Así nace la tentación de crear un juez robot, en principio para acelerar el proceso decisorio en los tribunales. Aquí se defiende, de un lado, la existencia de un derecho a la decisión judicial humana. No obstante, la IA tiene el potencial de garantizar equilibrio en la toma de decisiones, señalando eventual imparcialidad de las sentencias.

Palabras-Clave: juez robot; inteligencia artificial; singularidad; imparcialidad; huellas digitales; criptografía.

ABSTRACT

Today's artificial intelligence mechanisms can express themselves in natural language and interpret texts with great skill. This gives rise to the temptation to create a robot judge, initially to speed up the decision-making process in courts. This paper argues, on one hand, for the existence of a right to a human judicial decision. Nonetheless, AI has the potential to ensure balance in decision-making, indicating potential impartiality in judgments.

¹ Docente e investigador de la Universidad Nacional de La Plata (UNLP). Profesor Titular de la Cátedra 1 de Derecho Político, Facultad de Ciencias Jurídicas y Sociales, y Titular Asociado a la Cátedra de Relaciones Internacionales y Comunicación, Facultad de Periodismo y Comunicación Social, UNLP. Profesor de posgrado en la Maestría en Inteligencia Estratégica Nacional Siglo XXI, de la que es también Secretario Académico. Profesor Invitado en FLACSO sede Ecuador, FLACSO Costa Rica, Universidad Autónoma de Madrid, España, Universidad Augusto José Celia, Madrid España, Universidad Nacional Autónoma de México, Universidad Metropolitana de México, Universidad de Sao Paulo y Universidad Estadual de Sao Paulo, Brasil. Actualmente es Director del Programa de Defensa y Seguridad, UNLP.

² Doctor en Derecho por la PUC/SP, Brasil. Cursa Pos Doctorado en FDRP/USP, Brasil. E-mail: contato@thomaslaw.adv.br

³ Doctor y Libre-Docente por la USP/Brasil. Director en Pos-Doctorado en FDRP/USP, Brasil. E-mail: victorgabriel@usp.br

⁴ Profesor Doctor de Derecho Procesal Penal en USP y UNAERP, Brasil. E-mail: ssilveira@unaerp.br

Keywords: robot judge; artificial intelligence; impartiality; digital fingerprints; cryptography

1. PUNTOS FAVORABLES A LA INTELIGENCIA ARTIFICIAL

La Inteligencia Artificial se impone como objeto de estudio de muchos juristas, preocupados con los cambios inevitables que vendrán a raíz de su introducción más plena en la sociedad. Problemas de derechos autorales, de imputación penal, de riesgos a las empresas, de relaciones laborales o incluso fiscales, todos eso son desafíos en que el derecho ya está totalmente metido. A final, los riesgos para la estabilidad de las relaciones son evidentes.

Sin embargo, es posible plantear ventajas del uso de la IA en el campo jurídico, sin que ella implique la distopía de que los ordenadores dominen el sistema de decisión judicial, comandando la vida humana. Para aceptar este nuestro planteamiento, que creemos original⁵, es necesario estar sensible a una triste realidad nada inusual de los tribunales latinoamericanos: la falta de equidad en las decisiones de las Altas Cortes, que no raramente las toman a partir de los vínculos de los magistrados con cuestiones ideológicas o, peor, de acuerdo con el sistema de poder⁶.

El tema del sesgo ideológico o personal de las decisiones de las Altas Cortes tiene sido denunciado en muchos países, y esa fragilidad democrática puede ser corregida por lo que se llama la Inteligencia Artificial Generativa.

Nuestra propuesta no implica la dispensa del papel que juega la inteligencia humana, con su capacidad única de relaciones abductivas, para la decisión judicial, pero intenta demostrar que no es difícil aplicar la IA como elemento limitador del efecto indeseado de esa abducción: la tendencia judicial a decidir de acuerdo a gustos personales, aún a sabiendas de que se desobedece a la tradición jurisprudencial de la Corte. Para comprenderlo, hacemos un

⁵ Este planteamiento ya fue hecho en primera vez en la obra *Argumentación Jurídica* (Valencia, Tirant lo Blanch, 2024) y aprimorada en otras publicaciones. Este texto es una revisión de esa posición, por los demás autores. Se hace en el marco de la revisión internacional a partir del programa de Pós Doutorado da FDRP-USP.

⁶ BEDOYA, Andrés Felipe Roncancio; RESTREPO MARULANDA, Mariana; AGUILERA IZAGUIRRE, Gustavo. La intervención estatal y el estado social de derecho: análisis de como las ramas del poder público materializan los derechos fundamentales. *Nuevo Derecho*, [s.l.], v. 28, n. 17, p. 1-14, 2021. Disponible en: <https://doi.org/10.25057/2500672x.1395>.

pequeño recorrido por lo que parece realidad y lo que es leyenda en la futura aplicación de la IA.

2. SENTENCIAS ARTIFICIALES

La Inteligencia Artificial (IA) ya está en nuestros tribunales. Está presente y en uso, aunque se aplica con bastante moderación, si consideramos el potencial tecnológico existente. Los *softwares* que filtran argumentos, identificar temas, repeticiones e incluso estructurar (lo que serán) decisiones de jueces y magistrados⁷. Según una interpretación literal de las peticiones, están en funcionamiento, pero con competencias restringidas, lo que significa también que vivimos una época de debate sobre los límites éticos de la IA en la vida forense cotidiana⁸. Hay que comprender cómo funciona la IA y compararla mínimamente con la toma de decisiones humana, para luego sacar conclusiones sobre sus posibles cambios en la retórica y la argumentación.

3. INTELIGENCIA GENERATIVA

En primer lugar, hay que darse cuenta de que no todo el *software* es Inteligencia Artificial. Buscar o, por analogía, extraer datos dentro de otros datos es una de las primeras capacidades de los ordenadores más incipientes, todavía del siglo XX. Este proceso de filtrado, por supuesto, ha mejorado mucho, pero puede no ser cualitativamente superior a un procedimiento de organización de datos.

Un ordenador solo se convierte en IA cuando aprende. Así, la afirmación de que la inteligencia artificial es “la ciencia de hacer máquinas capaces de realizar tareas que

⁷ Conviene señalar, en aras de la precisión del vocabulario, que nunca podrá decirse que el ordenador construye una decisión judicial, a menos que esté investido como juez. Lo que hace es producir textos que se adoptarán como decisiones.

⁸ Pastor, aunque muy crítico con la justicia basada en la inteligencia artificial, en algún momento aclara lo que muchos hemos intuido: que puede “superar la jerga de los parlamentos y las divugaciones de la jurisprudencia, ofreciendo a la vez un oasis de certidumbre para la orientación normativa de los ciudadanos”. PASTOR, Daniel, *¿Sueña el sistema penal con jueces electrónicos?*, En: *Derecho Penal y Comportamiento Humano*, Eduardo Demetrio Crespo (Director), Valencia: Tirant lo Blanch, 2022, p. 540. Por nuestra parte, entendemos que esta *incertidumbre* de las decisiones colectivas debe ser reparada por la propia lógica humana, de lo contrario no habrá lugar para toda la narrativa que hay detrás de cada caso, que no sólo debe introducirse, con límites, en cada proceso, sino también en el proceso de la decisión colectiva. También es necesario que el ser humano conozca el proceso que hay detrás de la decisión, todos los factores que se han tenido en cuenta y en qué medida. Esto, en un algoritmo complejo, es imposible.

requerirían inteligencia si fueran realizadas por humanos”⁹ solo puede entenderse como correcta si se acepta que la inteligencia está dotada de razonabilidad, creatividad y aprendizaje¹⁰. Encontrar *soluciones* fuera de la programación. Por lo tanto, es más fácil adoptar definiciones que asocien la inteligencia artificial con el propio proceso de aprendizaje automático (*machine learning*)¹¹.

Existe entonces una gran cantidad de programación avanzada, en lo que se denomina *machine learning*, el proceso por el cual la máquina puede resolver continuamente problemas, crear nuevos caminos, que la mente humana no sería apta a encontrar por sí sola. O lo sería, pero no tan rápido. En el campo de la ingeniería, el funcionamiento de las máquinas, la logística, los vehículos autoconducidos, que ya son una práctica visible; los ordenadores que juegan al ajedrez han demostrado ser capaces de vencer a *cualquier* ser humano; o tecnologías más avanzadas que, tratando de imitar el cerebro animal, como las *artificial neuronal networks* (ANN) promover el *deep learning*, a potenciar su propio proceso de aprendizaje, ya sea a partir de las experiencias de los demás o de sus propios “errores”¹². Hoy en día, procesos muy avanzados de creación de textos, incluyendo un alto grado de

⁹ “AI is the science of making machines capable of performing tasks that would require intelligence if done by humans.” MINSKY, Marvin. *Society of mind*. Amsterdam: Simon and Schuster, 1988, p. 64.

¹⁰ La definición de la Britannica es la siguiente: “The term is frequently applied to the project of developing systems endowed with the **intellectual** processes characteristic of humans, such as the ability to reason, discover meaning, generalize, or learn from past experience”. In: <https://www.britannica.com/technology/artificial-intelligence>

¹¹ Para conceptos actualizados de Inteligencia Artificial y su clasificación, en que no nos vamos a detener de momento, véase: ARRIETA, Alejandro Barredo et al. Explainable artificial intelligence (xai): concepts, taxonomies, opportunities and challenges toward responsible ai. *Information Fusion*, [s.l.], v. 58, p. 82-115, 2020. Disponible en: <https://doi.org/10.1016/j.inffus.2019.12.012>.

¹² El hecho de que un ordenador pueda mejorar un procedimiento que ya ha realizado no autoriza directamente a calificar de error el procedimiento anterior, menos eficaz. Discutirlo, sin embargo, implicaría entrar en la definición filosófica del error, que es muy interesante incluso dentro del Derecho penal, pero no procede aquí, por razones puramente espaciales. El ajedrez, por sus combinaciones y, especialmente, por su potencial de cálculo y las diferencias entre razonamiento humano y de máquina es siempre ejemplo de machine learning. Véase, como ejemplo: DAVID, Omid E.; NETANYAHU, Nathan S.; WOLF, Lior. *Deepchess: End-to-end deep neural network for automatic learning in chess*. In: *Artificial Neural Networks and Machine Learning–ICANN 2016: 25th International Conference on Artificial Neural Networks*, Barcelona, Spain, September 6-9, 2016, Proceedings, Part II, v. 25. Springer International Publishing, 2016. p. 88-96.

creatividad, están a la mano de cualquier usuario¹³ y ya hay noticias de que se han redactado sentencias utilizando estos instrumentos¹⁴.

Quién no apuesta por el potencial del *deep learning* para el futuro próximo estará, en nuestra opinión, muy alejado del análisis de lo que ya está ocurriendo en el ámbito tecnológico, no solo en la capacidad de las máquinas que se están construyendo, como los ordenadores cuánticos, con potencial para escapar al sistema binario de funcionamiento, sino también en lo que ya se está realizando, como los ordenadores capaces de crear poesía¹⁵.

Otra cuestión, bien distinta, es si creemos o no que una inteligencia artificial superpoderosa pueda superar a la humanidad, emancipándose de ella, como ya predijo Asimov en su obra de 1956. Allí, las máquinas ya se encontraban en un dilema frente a las leyes de la robótica, que, en definitiva, les impedían anular en modo alguno la voluntad animal.

Un ordenador cuántico que pueda operar aprendiendo tan rápido que supere nuestra lógica de elegir sí o no, es decir, que supere la base binaria de lo legal-ilegal, o lo justo-injusto que es, de Kelsen a Luhmann¹⁶, la más lógica de las resoluciones de conflictos es algo no muy alejado de nuestra realidad. Esta distopía, ahora sí, asusta: los procesos de *machine learning* por los ordenadores cuánticos. Estos superan la lógica del sí o no, de lo binario tal

¹³ Aunque sabemos que esta observación quedará obsoleta a un ritmo muy rápido, la herramienta de lenguaje artificial más abierta para la creación de textos es GPT-4, que efectivamente es capaz de crear textos con un gran grado de "corrección" estilística y, si se puede decir así, técnica. Además, está aprendiendo en escala exponencial, recogiendo informaciones de la red. De momento, muchísimas respuestas son sesgadas, pero tal vez el trabajo de comparación mejore ese aspecto. ChatGPT se lanzó oficialmente, por OpenAI, en 2023, y estamos seguros de que otras similares o más avanzadas ya estarán funcionando cuando se lea este libro. La existencia de estas aplicaciones da mayor concreción a los debates, incluso éticos, por los que nos hemos movido en este trabajo. Otras herramientas como "Shortly" o "Write up" son más tradicionales en el mercado.

¹⁴ "A judge in Colombia used ChatGPT to make a court ruling, in what is apparently the first time a legal decision has been made with the help of an AI text generator—or at least, the first time we know about it. Judge Juan Manuel Padilla Garcia, who presides over the First Circuit Court in the city of Cartagena, said he used the AI tool to pose legal questions about the case and included its responses in his decision, according to a [court document](https://www.vice.com/en/article/k7bdmv/judge-used-chatgpt-to-make-court-decision) dated January 30, 2023". (<https://www.vice.com/en/article/k7bdmv/judge-used-chatgpt-to-make-court-decision>).

¹⁵ Véanse las conclusiones de KÖBIS, Nils; MOSSINK, Luca D. Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry. *Computers in human behavior*, 2021, vol. 114, p. 106553.

¹⁶ La solución legal/ilegal de Luhmann, basado en Maturana, e piensa entonces en una máquina que hace un *output* binario. Si pensamos en esta máquina como un ordenador que, en lugar de bits (0 ó 1), funciona con *cubits* (0 y 1), cambiaríamos toda la lógica jurídica, en el sentido de la toma de decisiones. Y supondríamos, como una obra de ciencia ficción, que existe una Ley que dice que algo es y no es al mismo tiempo: que sería la teoría de la Argumentación siendo tomada *ex machina*.

como lo conocemos; por lo tanto, si le dan la capacidad de aprehender el mundo, tendrán una comprensión del universo mejor que la que tiene el hombre, tal vez solucionando plenamente el mundo multidimensional que la ciencia sólo está empezando a imaginar, con hipótesis plausibles, pero solo virtualmente calculadas. Un escenario que, aunque interesante, está fuera de nuestro manejo.

Comprender argumentación no significa alejarse de la tecnología, sino todo lo contrario. Solo como mera técnica expositiva es que tomamos, acá, camino ortodoxo: primero, repasamos las claves del proceso humano de toma de decisiones. Luego las comparamos con el proceso de decisión artificial, que ya se utiliza.

4. CLAVES DE LA DECISIÓN HUMANA: LIBERTAD Y QUERER

El proceso humano de toma de decisiones es objeto de estudio de muchas ciencias y áreas del conocimiento. El derecho ha de ocuparse fundamentalmente de él, porque las acciones del individuo (y, por tanto, su forma previa de decisión sobre cómo actuar) son su objeto –y medio– de trabajo. Así, gran parte del derecho civil depende de un acuerdo de *voluntades*; también el derecho electoral, muy claramente, tiene (o debería tener) como núcleo la capacidad humana de decidir libremente. Sin embargo, en el derecho penal, la cuestión es más delicada, ya que la decisión individual de cometer un delito da lugar a un proceso de reacción estatal, en una retribución legitimada. Por lo tanto, al menos en teoría, este ámbito del derecho debería ocuparse de todo lo que se descubra sobre el funcionamiento del cerebro.

Muchos autores de derecho visitan la neurociencia para averiguar si ya existe una respuesta tanto a la causa (el proceso de toma de decisiones del individuo) como a la eventual reacción del Estado. Si realmente se desentraña el funcionamiento del cerebro, las propuestas para curar al delincuente serán las más adecuadas al Estado de derecho. Para tal finalidad, sin embargo, cabría a las ciencias médicas explicar el cerebro con mucho más detalle, y no parece que sea el caso por el momento. Sin embargo, podemos importar un poco de lo que se ha discutido entre el proceso de toma de decisiones en neurociencia y el derecho penal para responder a nuestra primera pregunta: ¿hay algo en la toma de decisiones humana que

la diferencia en esencia de lo que puede imitar una máquina? Tal pregunta significa lo mismo que preguntar si la voluntad humana es *libre*. Si existe el libre albedrío, la voluntad de querer. Tal vez una de las preguntas más antiguas de la filosofía.

5. CEREBRO HUMANO Y NEUROCIENCIAS

Los médicos que estudian el cerebro han intentado, por supuesto, explicar el método humano de toma de decisiones. Aunque no del todo actualizados, los experimentos más referidos¹⁷ por los juristas son las desarrolladas por Benjamin Libet, que tuvo la idea de monitorear las ondas electromagnéticas del cerebro del individuo, en el momento en que debe tomar una decisión. *Grosso modo*, su experimento consistía en someter a los participantes en la investigación a un electroencefalograma, mientras observaban un cronómetro muy preciso¹⁸. Cuando se les estimula en momentos diferentes e inciertos, se proponen (en su controvertida ‘*urge to move*’) un movimiento de muñeca y anotar el momento exacto en que fueron conscientes de este impulso. A continuación, el propio sujeto de la investigación señalaba, y luego informaba, el instante en que había tomado conciencia de su decisión, la decisión de moverse.

Utilizando la encefalografía, el médico californiano demostró que el cerebro humano empieza a moverse una centésima de segundo antes de que el individuo se dé cuenta de que ha decidido moverse. Esto significa que el proceso consciente de elección es *posterior* a una decisión inconsciente previamente tomada; por tanto, la conciencia de la decisión es ilusoria. Uno de los primeros estudios de Libet publicados en este sentido, en 1983, concluye, tras analizar los datos de dos grupos de tres personas cada uno, que “la iniciativa cerebral de un

¹⁷ Consulte: JÄGER, Christian. Willensfreiheit, Kausalität und Determination: Stirbt das moderne Schuldstrafrecht durch die moderne Gehirnforschung? *Goldammer's Archiv für Strafrecht*. n.01, 2013, pp.8 e ss, (adiante comentados). También RUBIA, Francisco J. El controvertido tema de la libertad. *Revista de Occidente*. n.356, 2011, p.7.

¹⁸ El cronómetro, llamado "reloj osciloscopio", se reproduce de forma esquemática en un artículo más reciente de Libet. Se trata de un punto luminoso que se desplaza alrededor de un círculo, marcado con puntos del 00 al 60, como si fuera un reloj que marca los sesenta segundos de un minuto. Sin embargo, como describe la figura, el punto de luz rodea la periferia del reloj a los 2,56 segundos, en lugar de los 60 segundos del reloj, con lo que “each maked off ‘second’ (in the total of 60 markings), representes 43 msec of actual time here”. LIBET, Benjamin. Do we have free will? *Journal of Consciousness Studies*. vol.6, n. 8/9, 1999, p.48 (Figure 01).

acto libre y espontáneo puede comenzar de forma inconsciente, antes de cualquier toma de conciencia de que la decisión ya ha comenzado en el cerebro”¹⁹.

Estos estudios han reforzado la hipótesis de que el cerebro decide por sí mismo, independientemente de una voluntad exterior a él²⁰. Más de quince años después de este primer experimento —en 1999 para ser exactos— Libet lo retomó y volvió a realizarlo para revisar sus conclusiones. Ahora ha reproducido sus experimentos con técnicas más avanzadas, y confirma una vez más que las decisiones libres van precedidas de un cambio eléctrico específico, 550 microsegundos antes de que se produzca la acción²¹, y unos 200 microsegundos antes de darse cuenta de que se desea actuar. Tiempo más que suficiente, por tanto, para seguir afirmando que el proceso de decidir moverse es anterior a la conciencia. Década y media después, Libet repite el experimento, con mayor precisión, y retoma la posibilidad de que el cerebro actúe mecánicamente pero el individuo pueda vetar *libremente* la acción²². Para él, las consecuencias sociales de declarar que el hombre no es libre para decidir serían inaceptables²³, pero estas consideraciones iban más allá de lo que había observado.

¹⁹ “It is concluded that cerebral initiation of a spontaneous, freely voluntary act can begin unconsciously, that is, before there is any (at least recallable) subjective awareness that a ‘decision’ to act has already been initiated cerebrally. This introduces certain constraints on the potentiality for conscious initiation and control of voluntary acts.” LIBET, Benjamin; GLEASON, Curtis; WRIGHT, Elwood; PEARL, Dennis. Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): the unconscious initiation of a voluntary act. *Brain*. n.106, 1983, p. 623.

²⁰ Desde aquel experimento de 1983, pues, las conclusiones de Libet ya han sido controvertidas por los propios investigadores, lo que no quiere decir que el estudio sea menos relevante. Se asume, como se ha dicho, la conclusión a la que llegó Rubia, disertando sobre los experimentos de Libet: “Éstos son los experimentos que han llevado a pensar que la impresión subjetiva de la voluntad libre es una ficción. (...) La hipótesis construida sobre estos datos sí puede discutirse y de hecho así se ha hecho. Pero hoy por hoy estos datos apuntan a que la libertad, tal y como la entendemos, es decir de acción y de decisión, parece una ficción”. RUBIA, Francisco J. El controvertido tema de la libertad. cit. p.9.

²¹ “Freely voluntary acts are preceded by a specific electrical change in the brain (the ‘readiness potential’ RP) that begins 550 ms before the act. Human subjects became aware in intention to act 350-400ms after RP starts, but 200ms before the motor act.” LIBET, Benjamin. Do we have free will? *Op.cit.* p. 24.

²² “All of us, not just experimental subjects, have experienced our vetoing a spontaneous urge to perform some act. This often occurs when the urge to act involves some socially unacceptable consequence, like an urge to shout some obscenity at the professor.” LIBET, Benjamin. Do we have free will? *Op.cit.* p. 52.

²³ Esto se debe a que el texto muestra que el científico no rehace los experimentos ni niega rotundamente lo que ya había concluido. Se limita, como decíamos, a subrayar una hipótesis que ya se había planteado en el momento del estudio original (la posibilidad de veto). Y el texto termina sintomáticamente con la cita de una entrevista concedida por Isaac B. Singer, el escritor, afirmando que el libre albedrío es el gran don recibido por la humanidad. Con esto sólo mostramos el subjetivismo de la interpretación, que sólo resulta de la convicción personal de la libertad.

Se han desarrollado nuevas técnicas,²⁴, especialmente en aquellos investigadores que, a diferencia de Libet, no están tan dispuestos a revisar sus conclusiones para confirmar el “*free will*”, o libre albedrío. En 2008 Haynes reprodujo un experimento similar²⁵²⁶, de las nuevas tecnologías. Itzhak Fried y su equipo han remodelado²⁷ lo mismo, ahora con la introducción de electrodos para monitorizar las neuronas de los pacientes epilépticos, y concluyeron casi lo mismo²⁸: el cerebro se mueve antes de ser consciente de la decisión.

Los estudios neurocientíficos siempre seguirán la línea de que el sistema cerebral anticipa las decisiones. No habrá ningún lugar, como el llamado *teatro cartesiano*, en el cerebro donde las decisiones se produzcan libremente. Esto no significa, sin embargo, que somos el *automaton materiale*, una *materielle Maschine* a la que se refiere Kant. La filosofía, además, relativizará mucho estas cuestiones.

6. LIBERTAD DE DECISIÓN EN CLAVE JURÍDICA

Aunque los nieguen en algún momento, los neurocientíficos pretenden demostrar el determinismo, es decir, el hecho de que todos nuestros pensamientos y decisiones siguen una regla físicoquímica, según la cual funciona el cuerpo humano y todo el universo conocido.

²⁴ Así, Smith: “Libet’s result was controversial. Critics said that the clock was distracting, and the report of a conscious decision was too subjective.” SMITH, Kerri. Taking aim at free will. *Nature*. vol.477, 2011, p.24.

²⁵ “Second, the time delay between the onset of the readiness potential and the decision is only a few hundred milliseconds. It has been repeatedly argued that potential inaccuracies in the behavioral measurement of the decision time at such short delays could lead one to misjudge the relative timing of brain activity and intention.” HAYNES, John Dylan; SOON, Chun Siong; BRASS, Marcel; HEINZE, Hans Jochen. Unconscious determinants of free decision in the human brain. *Nature Neuroscience*. vol.11, n.5, 2008, p. 543.

²⁶ “To rule out the idea that any leading activity merely reflects unspecific preparatory activation, it is necessary to study free decisions between more than one behavioral options”. HAYNES, John Dylan; SOON, Chun Siong; BRASS, Marcel; HEINZE, Hans Jochen. Unconscious determinants of free decision in the human brain. *Nature Neuroscience*. vol.11, n.5, 2008, p. 543.

²⁷ “We adopted a paradigm originally described by Libet and colleagues (Libet et al, 1983). Subjects were presented with an analogue clock depicted in a laptop and were instructed to fixate at the center. A clock dial rotated on the screen with a period of 2,568 ms. Subjects were instructed to place their right index finger on a key on the laptop keyboard, to wait for at least one complete revolution of the dial, and then press the key whenever ‘they felt the urge to do so.’ (...) We note that this ‘urge to move’ can be interpreted as a decision for self-initiated movement.” FRIED, Itzhak; MUKAMEL, Roy; KREIMAN, Gabriel. Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron* vol.69, n.3, 2011, p. 551.

²⁸ Por razones éticas, pues, todos los sujetos del experimento tenían algún tipo de enfermedad cerebral ya diagnosticada, pero esto no parece ser tan relevante para el método en sí mismo.

Este determinismo biológico, por muy avanzada que sea la tecnología, será objeto de varias críticas a lo largo de la filosofía, porque el libre albedrío es lo más fundamental a considerar para la autoconciencia del individuo, como responsable de sus actos.

Entre estas miles de críticas a las conclusiones deterministas precipitadas, ofrecemos humildemente la nuestra, que resumimos aquí para los fines de este capítulo²⁹. La primera de ellas (1) es que, tratándose de experimentos tan simples como los de Libet, no alcanzan, ni siquiera en hipótesis remotas, el verdadero proceso humano de toma de decisiones. Se refieren a elecciones sobre apretar botones, que es muy distinto de emitir un juicio. Meditar, por ejemplo, sobre el hecho de que una persona sea acusada de difamación en plena campaña electoral, lo que demanda reflexionar sobre los límites de la libertad de expresión, el derecho a la información en tiempos de competición, las exigencias lingüísticas, la legitimidad y las consecuencias de la sanción, todo al mismo tiempo, no es lo mismo que pulsar un botón. Así pues, la existencia de un movimiento mental previo a la conciencia no anula la libertad de querer —dondequiera que se oculte—.

Como consecuencia de (2), la combinación de factores que intervienen en la decisión es mucho mayor de lo que puede controlarse, por lo que adoptamos un punto de vista algo más categórico. Significa que identificar grandes zonas del cerebro que se mueven para inducir comportamientos desviados no significa desvelar cómo funciona este órgano en todas sus resoluciones. Estamos condicionados por elementos fisicoquímicos, pero no solo: nuestra memoria, nuestra educación (los textos que conocemos y reproducimos), nuestro propio carácter, todo ello aún no ha sido desentrañado.

Es interesante el contraargumento de que la distancia entre estos experimentos de laboratorio, que identifican grandes zonas del cerebro, y una comprensión completa de las decisiones humanas es una simple cuestión cuantitativa: si se procesaran más datos, incluso casi infinitamente, conoceríamos todos los elementos que se combinan para determinar el cerebro y, consecuentemente, todo el pensamiento humano. Esto permitiría, para nuestros fines, recrear en una máquina la forma individual de tomar decisiones. Lo que sería un juez electrónico perfecto.

²⁹ Ver RODRÍGUEZ, Víctor Gabriel. Libre albedrío y derecho penal. São Paulo: Marcial Pons, 2019.

El argumento no es aceptable, y resumimos dos razones para nuestra negativa. En primer lugar, el número de combinaciones cerebrales es muy superior al que cualquier máquina puede controlar. No son solo vibraciones, sino vibraciones influidas por una cadena de elementos, que se extiende a todo el universo: las cadenas químicas del cuerpo estarán sin duda influidas por todo el entorno, incluidos los planetas³⁰. Si la posición de la luna determina las cosechas y las mareas, por lo tanto, tendrá una reacción en nuestras decisiones. Miles de millones de neuronas, en reacciones multidimensionales, con los demás factores de todo el universo, todo esto hace que ningún ordenador lo imite jamás.

Si hacemos el paralelismo, tan de nuestro gusto, con una partida de ajedrez, la cuestión se aclara. Por muy avanzados que sean los ordenadores actuales, que hacen sucumbir a cualquier jugador humano en una disputa sobre el tablero, ninguna de estas máquinas ha conseguido “cerrar” la partida (*solve the game*), es decir, calcular todas sus posibilidades. Esto significa, para los informáticos, que un ordenador siempre puede ser superado por otro más potente, pero para nosotros tiene una implicación adicional: si es imposible calcular todo el funcionamiento de un simple juego bidimensional, con solo 32 piedras moviéndose por solo 64 posiciones, qué decir de una combinación de unos cuantos miles de millones de neuronas con todos los demás elementos del universo³¹. Cuando mucho vamos a construir ordenadores que superan matemáticamente los sistemas conocidos de criptografía, utilizados, por ejemplo, para sistemas de mensajes en los móviles, o seguridad bancaria. Pero el universo está criptografado de modo mucho algo más complejo que la contraseña del banco.

El tercer punto (3) es que la física cuántica ha demostrado que las posibilidades binarias de sí y no se superan cuando llegamos al universo subatómico. Esto significa que allí, aunque no lo deseemos, existe, si no hay una situación de caos, una situación imposible de prever, porque cualquier observación de la misma, por teoría, la cambia sustancialmente. En otras palabras, si el cerebro, en algún momento de sus operaciones, invade el “reino” de

³⁰ La física cuántica ha confirmado hasta ahora la comunicación (interacción) de partículas a una velocidad superior a la de la luz, en la paradoja Einstein-Podolski-Rosen.

³¹ De hecho, las posibilidades del ajedrez, que, por lo que se sabe, llegan a 10^{120} , fueron calculadas y publicadas en 1950 por Shannon, en un artículo en el que ya se cuestionaba que el ordenador pudiera “pensar”. En 1950. El artículo original está disponible en los portales de investigación científica: “Electronic computers can be set up to play a fairly strong game, raising questions of whether they can think” SHANNON, Claude E. A Chess-Playing Machine. Scientific American. vol.182, n. 2, 1950, pp.48-51. Como hemos dicho anteriormente, el ajedrez es una base cognitiva muy importante para comparar razonamiento humano y del ordenador.

la mecánica cuántica en lugar de la física newtoniana, habrá un momento en que sus reacciones sean científicamente inobservables, como en el dilema del gato de Schrödinger, o dependientes de un entrelazamiento que involucra, en acción y reacción, a partículas (sin exagerar) de toda la galaxia, como en la teoría del entrelazamiento cuántico. Ésta ya se ha demostrado. Como mínimo, se muestra que, en aquel conjunto de factores de influencia para el determinismo cerebral, tendríamos que incluir la posibilidad de que partículas subatómicas a años luz de nuestro cuerpo ejerzan su influencia, de un modo que siquiera podemos observar.

A estos tres puntos debo añadir solo uno más (4), tan rudimentario como difícil de responder. La afirmación circular de Cohen, según la cual: “si nuestros cerebros fueran lo suficientemente simples como para ser comprendidos, no seríamos lo suficientemente inteligentes como para comprenderlos”³². Comprender nuestro pensamiento significa, decimos, crear otro pensamiento, lo que requeriría una nueva comprensión, y así *ad infinitum*. Es una paradoja que no se solventa.

En resumen, sin menoscabar en absoluto la relevancia de la neurociencia, esta no logrará comprender el proceso de toma de decisiones del individuo en su conjunto, que es diferente del proceso del cerebro como órgano formado por miles de millones de neuronas en constante interacción. La toma de decisiones está vinculada a muchos procesos concomitantes que, por fuerzas metafísicas o caos combinatorio, dan lugar al libre albedrío. Esto, aunque ilusorio en cierto plano dimensional, está mucho más allá de lo que cualquier ordenador, por potente que sea, podrá procesar.

7. SI HAY UN DERECHO FUNDAMENTAL A UNA DECISIÓN JURÍDICA HUMANA

No es difícil conciliar dos posturas que hemos adoptado aquí, aparentemente contradictorias. Por un lado, hemos dicho que los ordenadores ya aprenden, que son capaces de crear (para lo que nos importa) textos muy específicos. Quizá mejor que los seres

³² Jack Cohen e Ian Stewart “*If our brains were simple enough for us to understand them, we'd be so simple that we couldn't.*” (COHEN, Jack; STEWART, Ian. *The collapse of chaos: discovering simplicity in a complex world*. London: Penguin Paperbacks, 1995, p. 8). La afirmación nos fue presentada por la tesis de máster de Thales Coelho, supervisada por nosotros, USP, 2018.

humanos³³. Por otra, damos a entender que no imitan la capacidad humana para tomar una decisión justa, que no es lo mismo que escribir una frase formalmente perfecta.

Un ordenador podría crear juicios basados en tantos criterios hasta reproducir juicios más perfectamente que los realizados por jueces humanos. Serían mejores en la forma, estadísticamente más equitativas, incluso con variaciones estéticas, lo que no implicaría necesariamente justicia. Entre los muchos argumentos que pueden esgrimirse para razonar esta falta de justicia, uno nos parece de lo más elemental: el ordenador puede utilizar infinitos criterios para tomar una decisión, pero nunca puede elegir *cuál* de estos parámetros utilizar. Y si desarrolla una forma de hacerlo, esta elección no será humana, lo que le llevará a cuestionarse qué criterios ha puesto en marcha *para* elegir sus criterios, en una espiral de interrogantes que, una vez más, tenderá al infinito. La decisión humana tampoco puede ofrecer estos criterios, porque nuestro cerebro puede *decidir*, como diría Frankfurt³⁴, lo que quiere, pero no *querer lo que quiere*. Así pues, las motivaciones de *fondo* de cualquier decisión humana jamás podrán exhibirse, porque el individuo no será consciente de ellas. La nota distintiva, sin embargo, será que se trata de una decisión humana, idéntica a la de los argumentadores y las partes en el procedimiento.

En otras palabras, un ordenador puede imitar la inteligencia humana, incluso mejor que el propio ser humano, pero nunca la igualará. Cómo un falsificador pinta un Van Gogh más identificable como Van Gogh que uno auténtico³⁵, pero nunca sería un cuadro original. Si un coleccionista decide que quiere un Van Gogh, pintado por el propio holandés, una imitación más que perfecta carece de valor. La razón por la que un coleccionista quiere un cuadro *original* es aparentemente un axioma, es decir, algo tan evidente que no debería cuestionarse. Sin embargo, solo el pensamiento mágico lo explica: desear la obra original de

³³ Dentro de unos años, no me cabe duda, por ejemplo, de que un ordenador será capaz, procesando sin cesar las grandes piezas musicales de la historia, de componer una sinfonía extremadamente compleja y, además, muy agradable al gusto humano. Sin embargo, el ordenador no podrá emocionarse al escucharla. Aunque, si está programado para ello, puede identificar los elementos que provocan emociones en los humanos y, a partir de ahí, reproducir las reacciones estándar ante ellos.

³⁴ “A person who is free to do what he wants to do may yet not be in a position to have the will he wants.” FRANKFURT, Harry G. Freedom of the will and the concept of a person. *The Journal of Philosophy*. vol.68, n.1, 1971, p.17.

³⁵ Puede ocurrir, por ejemplo, que un cuadro original de Van Gogh, pintado al principio de su carrera, no contenga los rasgos que los caracterizan. Sobre la falsificación de obras de arte y su curioso método, léase VAN DEN BERGHE, Rene Alphonse. Por amor al arte: memorias del ladrón más Famoso del mundo. Barcelona: Planeta, 2012.

Van Gogh, y no otra más perfecta que la suya, implica relacionarse con el hecho de que cada movimiento del pintor queda marcado en el cuadro y, de algún modo, se transfiere a quien lo observa o, más aún, a quien lo adquiere. Incluso menos perfecto técnicamente que su réplica, el original es objeto de deseo.

Nuestro paralelismo ya está establecido: el individuo tiene derecho a que una disposición judicial sea dictada por seres humanos, aunque técnicamente sea menos refinada que la del ordenador. De forma tácita, asumimos que cualquier imperfección judicial es la regla del juego, y debe ser corregida, pero la cuestión filosóficamente va más allá: el ser humano es el fin último de toda norma y, aunque lo consideremos una colectividad, ésta no incluye a las personas electrónicas. El individuo, como centro del ordenamiento jurídico, tiene derecho a ser juzgado por un homólogo. Eso es un querer humano.

Y la base de esta exigencia es, sin exageración alguna, todo lo que se expone en esta obra: el ciudadano-individuo necesita a un conciudadano como juez porque es con él con quien sabe dialogar, es a él a quien puede dirigir sus argumentos, de modo que una defensa planteada ante la máquina, por perfecta que sea, no es una defensa, es una guerra de capacidad de cálculo entre *softwares*, como ocurre cuando dos megaordenadores se enfrentan en programas de ajedrez: bonito de ver como enfrentamiento tecnológico o incluso como lección para futuras técnicas a adoptar por los humanos, pero totalmente inválido en los parámetros del juego como deporte. Y, si se nos permite estirar un poco más el paralelismo, este mismo deporte, aun siendo uno de los grandes temas para los retos de cálculo informático, sólo tiene funcionalidad como entretenimiento de gran valor histórico, no como cálculo en sí mismo. Es decir, sale de los tableros de ajedrez y se inserta en las máquinas como una forma de ilustración para el humano; pero dentro del ordenador, no será más que un conjunto de algoritmos programados y aplicados, porque la máquina nunca entenderá su analogía con la realidad, con la guerra entre reinos, el poder bélico de la dama, el movimiento sesgado de los alfiles, la imprevisibilidad de los caballos o la función de retaguardia de las torres. También la literatura: el ordenador de hoy puede construir un poema o un ensayo con sentido del estilo, pero no será literario porque le faltará la interacción, el sentido de la observación que lleva a

la creación³⁶. Del mismo modo, el derecho (como dogma) puede insertarse en la máquina, sí, como un conjunto de reglas, pero jamás existirá para un procesador electrónico como sentido de la justicia. Solo devolverá lo que se le haya programado, o lo que haya aprendido que sería correcto por ese mismo reglamento; el hecho de que, en el aprendizaje automático, la computadora aprenda elementos de decisión, y que tales elementos sean, de tan complejos, inaccesibles para el ser humano, no significa que sean justos. La complejidad del cálculo, que despierta nuestra admiración, no importa en la distribución de la justicia.

Para hacernos una idea más práctica, los estudiosos de la ingeniería lingüística más avanzada en la actualidad señalan tres debilidades principales para el lenguaje artificial: (a) en textos largos, en algún momento tendrán lagunas de coherencia; (b) pueden estar llenos de prejuicios, y no habrá forma de filtrarlos; (c) pueden aportar información falsa, ya que sus fuentes serán desconocidas³⁷. Uno puede, utilizando el ejemplo de los autores, pedir que se construya un texto sobre la supremacía de un determinado grupo étnico sobre otro, y el ordenador lo construirá. Por supuesto, se basará en mentiras, pero será un argumento formalmente perfecto. No hay que ir mucho más lejos para decir que la ley no debería admitirlo³⁸.

³⁶ La cuestión es más sensible de lo que puede aparentar. Si asumimos que el ordenador construye una obra de arte sin la capacidad de asociación entre lo que allá está se representa y su propia vida (porque no la tiene), entonces hay que asumir que cualquier obra, como un romance, es una copia de estilos que no fueron comprendidos, sino sencillamente combinados. Eso también pasa, hoy en día, muy claramente en la música. Con lo cual, se podrá construir un buen romance o una excelente música sólo por *machine learning*, pero la cuestión es si ella puede ser aceptada por la sociedad. Sería, en nuestra opinión, siempre como una copia impresa de un cuadro de Van Gogh: la misma imagen, sin ningún valor. De ahí nace otro problema, que son los individuos que sencillamente prestan su nombre a creaciones del ordenador, como si fuesen los verdaderos creadores de las obras hechas mecánicamente. Eso ya pasa en la música, en los trabajos escolares, en alguna sentencia judicial. El ordenador puede ser un excelente auxiliar para un fallo, pero, como se dirá, no puede hacerlo por sí mismo, al faltarle la experiencia de vida necesaria a la decisión.

³⁷ “Its outputs may lack semantic coherence, resulting in text that is gibberish and increasingly nonsensical as the output grows longer. Its outputs embody all the biases that might be found in its training data: if you want white supremacist manifestos, GPT-3 can be coaxed to produce them endlessly. Its outputs may correspond to assertions that are not consonant with the truth.” DALE, Robert. *GPT-3: What’s it good for? Natural Language Engineering*. vol.27, 2021, pp. 116 e ss. Veja-se también: LIU, Xiao, *et al.* GPT understands, too. *arXiv preprint arXiv:2103.10385*, 2021, e MCGUFFIE, Kris; NEWHOUSE, Alex. The radicalization risks of GPT-3 and advanced neural language models. *arXiv preprint arXiv:2009.06807*, 2020.

³⁸ Ese ejemplo es extremo, pero ya hay muchas injusticias cometidas por la inteligencia artificial debidamente comprobadas: usar el promedio de criminalidad del barrio en que vive el reo para determinar su culpabilidad, usar género y color de piel para seleccionar plazas para empleo, sin que los propios empleadores supieran que esta tecnología lo hacía. Fue el caso de la IA de Amazon, véase: HAMILTON, Isabel Asher, Amazon built an AI tool to hire people but had to shut it down because it was discriminating against women, Oct 10, 2018 in: *Business Insider*, NY October, 2018

Podríamos ir más lejos, a cuestiones más filosóficas, sobre todo al decir que la inteligencia artificial, aunque capaz de un aprendizaje mucho mayor que el humano, nunca podrá hacerse *preguntas*, saber qué debe *resolver*, porque no está afligida por los dramas y problemas que rigen la creación, incluso en el derecho. También se cuestionará hasta qué punto se admitirá que el ordenador intervenga por sí mismo en la decisión, ya que, como sabemos, todo conocimiento social es participativo, es decir, el propio observador forma parte de él. Más aún, todo conocimiento jurídico es proposicional (propone interpretaciones y cambios), de ahí la cuestión de qué interés protegería un ordenador-decisor³⁹.

En este punto, sin embargo, hay una gran paradoja, casi oximorónica, que nosotros hemos detectado y a la que nos gustaría encontrar propuestas para una relativa solución. En definitiva, la justicia solo la hará la Inteligencia Artificial a través de un algoritmo cuyo procesamiento nuestro cerebro no será capaz de entender debido a su complejidad. Ahí radica la primera parte de la paradoja, típica de la informática: el ser humano es apto a desarrollar una máquina precisamente para que procese muchos más datos de los que la inteligencia del creador puede alcanzar, como en el ejemplo de los programas de ajedrez, pero esto ya está trasnochado. El núcleo de nuestro oxímoron es la transparencia⁴⁰: a diferencia de una decisión judicial humana, cuyo sentido de la justicia, por muy racional que sea, tiene un núcleo impenetrable (el libre albedrío), la decisión informática tiene un algoritmo programado y cuyo despliegue, en conjunción con los datos que recoge de los hechos del caso, la información que tiene de los individuos en su base de datos, además del ordenamiento jurídico completo, hasta llegar a la decisión judicial final, quedará todo registrado, paso a paso. Sin embargo, a pesar de todas estas combinaciones y acoplamientos, este verdadero *rastró digital*, por complejo que sea, no será comprendido por el ser humano. Por lo tanto, la *motivación* de la decisión, derecho fundamental del tribunal, **está tan minuciosamente registrada que no existe**, porque nuestra razón no llega a ella. Si se pide a un ordenador que desvele sus razones para decidir, o bien (a) presentará su propio algoritmo, indescifrable para los operadores del derecho, o bien (b) creará, mediante otro algoritmo de inteligencia

³⁹ Más sobre este tema se encontrará en nuestro RODRÍGUEZ, Víctor Gabriel, *Argumentación Jurídica: Texto, Razonamiento y Lógica Informal*, Valencia: Tirant lo Blanch, 2024

⁴⁰ Oxímoron, como en el caso, con alguna licencia poética, podríamos decir que es una *transparencia opaca*, o una desinformación total, por el exceso informativo. Algo que, en derecho tecnológico, ya no es novedad.

artificial de texto, un escrito que fundamente la decisión en lenguaje humano. El caso (a) implica la práctica inexistencia de razonamiento, ya que la explicación de las razones para decidir es el propio algoritmo. Por tanto, el fundamento es el propio ordenador, como si un juez humano nos dijera, al estilo de Luis XIV: “las razones de la sentencia son mi cerebro y mi alma”; (b) en el segundo caso, la razón para decidir sería un texto creado *ex post*, en el peor sentido del realismo jurídico, en una farsa preprogramada, porque no corresponde al algoritmo, sino a un constructo desvinculado de la realidad digital, pero que agrada a las partes jurisdiccionalizadas, al imitar la sentencia judicial humana. En otras palabras, se trataría de una estafa programada⁴¹. Una solución no imposible sería que el poder judicial se viera obligado a someter cada proceso a una decisión por IA, y luego mostrar esa sentencia artificial a las partes de la demanda. El juez no estaría obligado a seguir la decisión, pero la enseñaría. La IA funcionaría entonces como un elemento para investigar y aconsejar, proporcionando al juez información y razonamientos que pueden haberse escapado. Así, se anhela la seguridad jurídica, la homogeneidad de las decisiones, que es tan necesaria. Por supuesto, esto requeriría la aceptación del mismo sistema de IA para toda la judicatura, además de otras cuestiones administrativas.

El derecho a ser juzgado por un ser humano no implica no creer en la posible perfección del juicio de un ordenador: menos sujeto a errores, desigualdades, quizá incluso prejuicios. Sin embargo, toda condena tiene un reproche moral, y será difícil aceptar que un ordenador tenga autoridad moral sobre los seres humanos, por muy imperfectos que seamos.

⁴¹ Tener un humano como juez es, por tanto, una forma de entender la sociedad, inscrita como derecho fundamental desde el deber de prestación jurisdiccional. En lugar de demostrarlo con otros argumentos que nos distraerían del tema, nada mejor que una ilustración tan paradigmática como la de Orwell en 1984, cuando nos habla del mundo en el que las máquinas componen poemas. “Cuando la boca de la mujer no estaba ocupada por broches y ropa, cantaba con una poderosa voz de contralto: ‘Fue sólo una ilusión sin esperanza/que pasó como un día de abril; pero esa mirada, esa palabra/y los sueños que despertó/se apoderaron de mi corazón’ Aquella canción había sido la moda en Londres durante varias semanas. Era una de las producciones de la subsección del Departamento de Música, destinada a los proles. Las letras de estas canciones se componían sin intervención humana, mediante un aparato llamado “versificador”. Pero la mujer las cantaba con tan buen gusto que el horrible ritmo se convertía en sonidos casi agradables” ORWELL, G. 1984, p. 141 (traducción libre). El “versificador” en la obra de Orwell produjo una canción exitosa. Poética, políticamente correcta y agradable para el ser humano hasta el punto de ser un “hit”. En términos tecnológicos, ya no es una utopía, una ficción, pues ya existen programas capaces de construir textos estéticamente agradables. Sólo su uso nos parece distópico, pero aquí ya hemos cerrado nuestro paralelismo.

Resumiendo: (a) el centro del orden jurídico es el ser humano, individual o colectivamente pensado; (b) como tal, tiene derecho a ser juzgado por otro individuo, porque sólo puede dialogar con su contraparte - y este diálogo es el que ejerce la defensa; (c) siempre quedará la cuestión de que cualquier consideración jurídica, aunque sea formal, parte de un punto de vista del observador, del sujeto cognoscente. Un juicio artificial otorga al ordenador el derecho de ser el conocedor y el decisor de la sociedad, incluso en la construcción del texto más formal; (d) un ordenador aporta en su procesamiento todas las razones para decidir, pero no puede transmitirlos al ser humano, lo que imposibilita cualquier tipo de contradicción; (e) es a través de los derechos fundamentales como se impedirá que los ordenadores, aunque sean capaces de construir textos técnicamente perfectos, decidan judicialmente o escriban el razonamiento de las sentencias; (f) algo totalmente distinto es que los jueces humanos utilicen la IA para orientar sus decisiones, incluso enseñando a las partes cuál sería la forma que un ordenador interpreta las pruebas y, especialmente, cómo las compara con la jurisprudencia, las decisiones anteriores del mismo tribunal.

8. DIALÉTICA Y ORDENADOR: UNA ÚTIL CONJUNCIÓN

Dejar bien fijado que existe un derecho fundamental de recibir la decisión judicial de un ordenador humano es condición para enmarcar las ventajas que la IA puede traer para el todo el Poder Judicial. Nuestra concepción es de algún modo sencilla: así que la IA evolucione un poco más en su capacidad de análisis de contenido, ella puede ser dirigida a *opinar* sobre cada proceso judicial, especialmente en las Altas Cortes.

Eso significaría que la máquina sería consultada para, sin cualquier posibilidad de interferencia judicial previa, establecer una *decisión* para cada caso. Esa decisión tomaría en cuenta las pruebas del proceso y la argumentación de las partes y, luego, traería una *propuesta* de solución. Tal propuesta, aunque con las desventajas de la ausencia de un pensamiento abductivo, conllevaría un potencial que el ser humano no tendría: la de comparar, sin sesgos, la jurisprudencia del tribunal e indicar, directamente, cuáles los fallos antecedentes de la Corte, que serían aplicables al caso concreto.

Esa decisión previa de la Inteligencia Artificial, por más que la máquina aprenda sola a manejar el estilo y la lógica judicial, tendría potenciales injusticias, con lo cual no sería una decisión final. Si, empero, ella es, sin intervención judicial, sometida directamente a las partes para respuesta, ellas se manifiestan sobre puntos fuertes y debilidades de la decisión robótica. Luego de ese ejercicio dialectico vendría la decisión del magistrado humano, pero ya con un elemento procesal que no se puede borrar: la indicación, por la máquina, de que hay que mantener una coherencia con decisiones anteriores. Si el magistrado pretende cambiar la decisión del ordenador tendrá toda la autoridad para hacerlo, y con un elemento facilitador: su razonamiento será también orientado para asentir o discordar de la sentencia provisional de IA generativa.

09. CONTRAPUNTO Y CONCLUSIÓN

No se nos escapa que, caso se implemente esa propuesta, haya el riesgo de que el juez delegue al ordenador ese trabajo irrenunciable de redactar la decisión humana, sencillamente declarando una concordancia genérica con la sentencia de IA, o, lo que sería peor, copiando sin más críticas esa misma sentencia automática. Sin embargo, ese es un problema que, si bien observado, nuestra propuesta no haría más grave: ya existe IA aplicada en los tribunales, con el propósito único de filtrar argumentaciones y elegir modelos de decisión *pretensamente* tomada por seres humanos. Si bien analizada, nuestra propuesta sería funcionalmente buena, porque obligaría al juez a considerar la decisión artificial y su dialéctica como elementos añadidos al razonamiento de su sentencia final. Y, si quisiera cambiar la jurisprudencia de su Corte para agrandar a los intereses no-judiciales, tan comunes en los Altos Tribunales latinos, estaría obligado a enfrentar las razones homogéneas, estadísticas y lógico-deductivas del ordenador.

Si un día consideramos que la máquina tiene potencia suficiente como para interpretar el lenguaje natural y pronunciarse en la complejidad de un proceso, no hay motivo para que no se la introduzca como elemento paralelo, a fin de reforzar la ecuanimidad de las decisiones y la igualdad de los ciudadanos frente a la ley. Siempre y cuando se garantice que la toma

final de la sentencia es del juez o magistrado, observador externo al proceso de Inteligencia Artificial Generativa.

10. REFERENCIAS

ARRIETA, Alejandro Barredo et al. *Explainable artificial intelligence (xai): concepts, taxonomies, opportunities and challenges toward responsible ai*. Information Fusion, [s.l.], v. 58, p. 82-115, 2020. Disponible en: <https://doi.org/10.1016/j.inffus.2019.12.012>.

BEDOYA, Andrés Felipe Roncancio; RESTREPO MARULANDA, Mariana; AGUILERA IZAGUIRRE, Gustavo. La intervención estatal y el estado social de derecho: análisis de como las ramas del poder público materializan los derechos fundamentales. **Nuevo Derecho**, [s.l.], v. 28, n. 17, p. 1-14, 2021. Disponible en: <https://doi.org/10.25057/2500672x.1395>.

CLARKE, Arthur C. *2001: a space Odyssey*. New York: ROC Penguin Books, 1993.

DALE, Robert. *GPT-3: What's it good for? Natural Language Engineering*. vol.27, 2021, pp. 116 e ss. Veja-se também: LIU, Xiao, et al. GPT understands, too. *arXiv preprint arXiv:2103.10385*, 2021

DAVID, Omid E.; NETANYAHU, Nathan S.; WOLF, Lior. *Deepchess: End-to-end deep neural network for automatic learning in chess*. In: Artificial Neural Networks and Machine Learning–ICANN 2016: 25th International Conference on Artificial Neural Networks, Barcelona, Spain, September 6-9, 2016, Proceedings, Part II, v. 25. Springer International Publishing, 2016. p. 88-96.

FRANKFURT, Harry G. Freedom of the will and the concept of a person. *The Journal of Philosophy*. vol.68, n.1, 1971

HAMILTON, Isabel Asher, Amazon built an AI tool to hire people but had to shut it down because it was discriminating against women, Oct 10, 2018 in: Business Insider, NY October, 2018

HAYNES, John Dylan; SOON, Chun Siong; BRASS, Marcel; HEINZE, Hans Jochen. Unconscious determinants of free decision in the human brain. *Nature Neuroscience*. vol.11, n.5, 2008

JÄGER, Christian. Willensfreiheit, Kausalität und Determination: Stirbt das moderne Schuldstrafrecht durch die moderne Gehirnforschung? *Goltdammer's Archiv für Strafrecht*. n.01, 2013

KÖBIS, Nils; MOSSINK, Luca D. Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry. *Computers in human behavior*, 2021, vol. 114, p. 106553

LIBET, Benjamin; GLEASON, Curtis; WRIGHT, Elwood; PEARL, Dennis. Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): the unconscious initiation of a voluntary act. *Brain*. n.106, 1983

MCGUFFIE, Kris; NEWHOUSE, Alex. The radicalization risks of GPT-3 and advanced neural language models. *arXiv preprint arXiv:2009.06807*, 2020

MINSKY, Marvin. *Society of mind*. Amsterdam: Simon and Schuster, 1988

PASTOR, Daniel, ¿Sueña el sistema penal con jueces electrónicos?, En: *Derecho Penal y Comportamiento Humano*, Eduardo Demetrio Crespo (Director), Valencia: Tirant lo Blanch, 2022

RODRÍGUEZ, Víctor Gabriel, *Argumentación Jurídica: Texto, Razonamiento y Lógica Informal*, Valencia: Tirant lo Blanch, 2024

RODRÍGUEZ, Víctor Gabriel. *Livre Arbitrio e Direito Penal*. São Paulo: Marcial Pons, 2019

RUBIA, Francisco J. El controvertido tema de la libertad. *Revista de Occidente*. n.356, 2011, p.7

SHANNON, Claude E. A Chess-Playing Machine. *Scientific American*. vol.182, n. 2, 1950

SMITH, Kerri. Taking aim at free will. *Nature*. vol. 477, 2011

Submetido em 30.03.2024.

Aceito em 18.06.2024.