



Universidad de Valparaíso

Tesina de Derecho

“Inteligencia artificial y uso de algoritmos en las decisiones judiciales ¿es un método exento de subjetividades?”

Autores:

Bruno Figueroa Vilos.

Camila Muñoz Rubilar.

Profesora guía:

Dra. Patricia Reyes Olmedo

Diciembre, 2021

Tabla de contenido:

Introducción	4
I. Una aproximación a la Inteligencia Artificial: avances, desarrollos y su impacto en el Derecho.	5
1. ¿Qué es la inteligencia artificial?.....	5
2. ¿Cómo opera la inteligencia artificial?.....	9
3. Impacto de la Inteligencia artificial en el Derecho: usos actuales.	12
II. Uso de algoritmos en sede judicial: ventajas y desventajas.....	16
1. Ventajas del uso de sistemas algorítmicos de IA	17
1.1. Eficiencia y descongestión.	17
1.2. Objetividad, certeza y despersonalización de la función jurisdiccional.	18
2. Desventajas o problemas del uso de algoritmos en sede judicial.	20
2.1. La trampa de la neutralidad: sesgos y discriminación algorítmica	20
2.2. Falta de transparencia y opacidad de los sistemas de aprendizaje profundo	22
3. Sesgos algorítmicos en sistemas de justicia criminal: casos emblemáticos para visibilizar la trampa de la neutralidad.	24
III. Sistemas, modelos y experiencias de las aplicaciones de IA en el mundo y en Chile	29
1. El caso de China	31
2. El caso de Estonia.....	31
3. El caso del Reino Unido	32
4. El caso de Emiratos Árabes Unidos	32
5. El caso de Canadá	32
6. El caso de Australia.....	33
7. Situación en Chile.....	33
IV. ¿Cómo minorizar los riesgos? La ética en el uso de la IA	36
1. ¿Qué han señalado los organismos internacionales?	36
2. ¿Qué se ha señalado desde la dogmática?	39
V. Algunas reflexiones sobre la implementación de la IA en el proceso chileno	45
1. Algoritmos en el proceso penal	47
2. Algoritmos en el proceso civil.....	51
3. Al menos Prometea.....	55
Conclusiones.....	58
Referencias bibliográficas:	60

Tabla de abreviaturas

1. IA	Inteligencia Artificial.
2. TIC	Tecnologías de la Información y la Comunicación.
3. ML	Machine Learning o Aprendizaje Automático.
4. DL	Deep Learning o Aprendizaje Profundo.
5. SEJs	Sistemas de Expertos Jurídicos.
6. CEPEJ	Comisión Europea para la Eficiencia de la Justicia.
7. OECD	Organización para la Cooperación y el Desarrollo Económicos.
8. CAPJ	Corporación Administrativa del Poder Judicial.
9. ODR	Online Dispute Resolution.
10. CRT	Civil Resolution Tribunal.

Resumen

El presente trabajo busca, a partir de la experiencia comparada, entregar algunas reflexiones respecto de un eventual uso de la inteligencia artificial en el sistema de justicia chileno, y sobre los potenciales riesgos y problemas que podría presentar. A partir de una explicación sistemática de los fenómenos tecnológicos involucrados, se aborda la manera en que estos se han ido aplicando en ordenamientos jurídicos occidentales en los últimos tiempos, los problemas tanto prácticos como teóricos que pueden advertirse en su implementación, para finalmente aterrizar en el ordenamiento chileno analizando las consecuencias de una posible implementación de estos sistemas comparados a la vista de las implicancias jurídicas -principalmente en sede civil y penal- y sociales que pudiere generar.

Palabras claves

Inteligencia artificial – Algoritmos – Derecho – Justicia – Sesgos algorítmicos

Introducción

La irrupción progresiva de la tecnología en los distintos aspectos de la vida cotidiana constituye un hecho inevitable en un mundo hiperconectado como en el que vivimos. Desde tiempos remotos, por medio de su intelecto, el ser humano se ha dotado de instrumentos serviciales a sus tareas, artificios que facilitan las funciones que son innatos a los fenómenos que genera la razón humana. Uno de esos fenómenos en que irradia la tecnología es el jurídico, y en efecto, el avance tecnológico-computacional de las últimas décadas cada vez tiene más que decir respecto de su rol a cumplir en ese orden normativo.

Indudablemente, uno de los protagonistas del nuevo mundo digital es la inteligencia artificial, noción de la informática cuyo horizonte es replicar los mecanismos de accionar del cerebro humano, y, en consecuencia, su incidencia tiene cabida en la labor intelectual de los operadores jurídicos. Por lo mismo, será necesario antes que todo delimitar conceptualmente los aspectos de esta tecnología disruptiva, para luego estacionarlo en el fenómeno jurídico.

En líneas generales, este trabajo abordará el impacto de la inteligencia artificial y la evolución de los llamados algoritmos, particularmente en lo que respecta a la administración de justicia estatal, en su dimensión civil y en su dimensión penal, estudiando sus ventajas y desventajas teóricas a partir de lo que han planteado juristas dedicados a este tema, así como reflexionar sobre los problemas que en la práctica se han podido avistar con la implementación reciente de sistemas de inteligencia artificial en distintos ordenamientos jurídicos, principalmente en Estados Unidos y Europa, y que dicen relación con sesgos de información traducidos, en lo particular, en discriminación; y en lo general, en vulneración de derechos fundamentales.

Servicial a lo anterior, en el camino se conocerán los principales sistemas de inteligencia artificial que se han aplicado en el espacio geográfico comentado y las experiencias más emblemáticas que se pueden relatar al respecto en vista de los posibles problemas que genera el uso de estos mecanismos en la administración de justicia. Todo esto nos llevará finalmente a la situación nacional, en el que analizaremos el uso de estas herramientas en nuestro ordenamiento jurídico, y con ello, las potenciales consecuencias positivas y negativas dado el contexto institucional y social del país.

I. Una aproximación a la Inteligencia Artificial: avances, desarrollos y su impacto en el Derecho.

El uso de las Tecnologías de la Información y la Comunicación (TIC), así como de la Inteligencia Artificial (IA), han permeado la generalidad de los ámbitos de la cotidianidad y áreas del saber humano, y el Derecho no ha estado exento de ello. La IA es una tecnología que está revolucionando¹ la manera en que la sociedad interactúa y se expresa, aunque el impacto que esta ha causado en el Derecho pareciera ser, a simple vista, menos perceptible si es que se la compara con otras disciplinas –por ejemplo, con la medicina–, debido a que enfrenta mayores obstáculos y reticencias estrechamente ligadas a las formas tradicionales de estudiar, investigar y ejercer el Derecho, así como el desconcierto que suele acompañar a los cambios tecnológicos.

Sin embargo, el hecho de que, en el ámbito jurídico, la adopción de soluciones y avances tecnológicos se produzca con mayor lentitud y despierte reticencias, no es razón suficiente para desconocer que las aplicaciones de la IA cuestionan los roles tradicionales e imponen desafíos a los operadores del derecho, quienes tienen, al menos, el deber de integrar estas tecnologías y la necesidad de responder, no simplemente de forma reactiva, sino también precautoria.

1. ¿Qué es la inteligencia artificial?

En lo que a la IA se refiere, aunque no exista un total consenso acerca de lo que la expresión «*Inteligencia Artificial*» significa, sino que, muy por el contrario, dependiendo de la perspectiva o enfoque, es una palabra que puede designar significados muy diferentes, sí que podría decirse, como noción previa, que se utiliza para referir a “sistemas que manifiestan un comportamiento “inteligente”. Que son capaces de analizar su entorno y, con cierto grado de autonomía, tomar decisiones para alcanzar ciertos objetivos específicos” (Borges Blázquez, 2020: p. 55), siendo una suerte de noción omnicomprendiva, como se demostrará más adelante.

¹ Reyes Olmedo (2020) indica que se la ha reconocido como la base de la “Cuarta Revolución Industrial”, continuadora de la denominada “Revolución Digital” y que estuvo basada en el uso intensivo de tecnologías de información y comunicaciones (TIC). En este nuevo contexto, la IA es señalada como elemento central de esta revolución, y con ella se relacionan los fenómenos de big data, uso de algoritmos y la interacción permanente y masiva de sistemas y dispositivos digitales.

Por lo demás, no existe pleno consenso en torno a esta expresión, pues, a lo largo de la historia se han desarrollado diversos enfoques de la IA. Para Russell y Norvig (2004), se pueden resumir en cuatro: el comportamiento humano, que es el enfoque del test de *Turing*, según el cual, un sistema de IA es inteligente en la medida en que para un observador, realiza conductas humanas; el pensar como humano, enfoque del modelo cognitivo, a partir del cual, un sistema de IA es inteligente en la medida que piense como humano; el pensamiento racional, óptica de las leyes del pensamiento, conforme al cual, un sistema es inteligente en la medida que resuelve problemas lógicos formales; finalmente, el actuar de forma racional. Un agente racional, según estos autores, es algo que razona, pero, además, debe tener otros atributos que lo distinguan de los programas convencionales: el estar dotado de controles autónomos, que perciba su entorno, que persista durante un período de tiempo prolongado, que se adapte a los cambios, etc.

Ahora bien, en lo que respecta a la IA como disciplina, Malpica Velasco la describe señalando que “es la rama de las Ciencias de la Computación que estudia el software y hardware necesarios para simular el comportamiento y comprensión humanos” (citado por Torres Manrique, 2017: p. 212). Esto de simular el comportamiento y comprensión humanos fue lo que, desde sus orígenes, ha distinguido a la IA de otras disciplinas, precisamente, como señala López Oneto:

“Desde sus inicios, pues, la IA se diferenció de otras disciplinas por su declarado objetivo de duplicar las facultades humanas tales como la creatividad, el autoaprendizaje y el uso del lenguaje, utilizando la informática y construyendo máquinas que funcionen en medios complejos y cambiantes, y que encarnen o posean esas características, permitiéndole decir a un observador que interactúa con ellas, que tiene frente a sí, a entes inteligentes, a sapiens” (2020: p. 41)

En el mismo sentido anterior, Nieva Fenoll da cuenta de que la IA “describe la posibilidad de que las máquinas, en alguna medida, piensen, o más bien imiten el pensamiento humano a base de aprender y utilizar las generalizaciones que las personas usamos para tomar nuestras decisiones habituales” (2018: p. 21), aunque en la práctica, este propósito de replicar la manera en que funciona el cerebro humano ha sido difícil de lograr, por la compleja característica de pensamiento abstracto propia de los humanos, limitándose sus avances y desarrollos a replicar el mismo resultado o uno similar en calidad a aquel que se obtiene cuando es un cerebro humano el que realiza los procesos, sin necesariamente replicar los procesos cognitivos cerebrales.

Lo anterior recuerda aquella clasificación propuesta, en los 80, por el filósofo norteamericano John Searle, en un artículo crítico con la IA, que distingue entre la IA fuerte y la IA débil y que constituye, al mismo tiempo, una forma de aproximarse conceptualmente a la IA. Una investigación que se hace cargo de la distinción da cuenta que “la IA fuerte implicaría que un ordenador convenientemente programado no simula una mente, sino que es una mente, por consiguiente, debería ser capaz de pensar igual que un ser humano” (Meseguer González y López de Mántaras, 2017: p. 10). La IA fuerte, alcanzaría tal grado de autonomía, que no solo es capaz de obtener un resultado y realizar una actividad con igual o mayor eficiencia con la que la que la ejecuta un ser humano, sino que, además, es capaz de replicar de forma exacta y superior, los procesos neuronales que se llevan adelante durante una actividad determinada en el cerebro humano, razonando como éste último, pero sin ser un humano.

La IA débil, en cambio, sería el tipo de inteligencia que se encuentra en la mayoría de los avances tecnológicos y “consistiría, según Searle, en construir programas que ayudan al ser humano en sus actividades mentales en lugar de duplicarlas” (Meseguer González y López de Mántaras, 2017: p. 10). Por esta razón, dependen de la guía de las personas que la programan, careciendo del dinamismo de la IA fuerte, respecto a lo que la resolución de problemas se refiere. Si bien, no logran razonar de un modo autónomo y abstracto, como lo haría un ser humano, pueden, sin embargo, ofrecer iguales o mejores resultados de forma rápida y eficiente.

Una buena parte de los autores cuyas ideas se sintetizan en López Oneto (2019), han afirmado la idea de que, hasta el momento, no se han construido IA fuertes, en el sentido de Searle o en cualquier sentido que pudiera implicar un nivel de simulación de la inteligencia humana en ambientes complejos. A diferencia de la IA débil, donde sí se han alcanzado niveles extraordinarios, a través de los sistemas expertos, que superarían la pericia humana y que mantendrían viva la ilusión de que, en un futuro próximo, será factible elaborar sistemas artificiales capaces de asimilar todos los procesos mentales humanos, o incluso generar otros.

Con todo, es posible analizar el asunto desde la perspectiva de Barona Vilar (2021). Para esta autora, la IA débil se caracteriza por la realización de una tarea concreta, por ejemplo, la función de asistencia en Siri de Apple. Se trata de sistemas que no pueden ir más allá de aquello para lo que originalmente fueron programados y que tienen una única misión, por ende, solo pueden

llegar hasta donde ha sido delimitada su función. En estos casos, la IA débil es capaz de guardar información, pero solo transitoriamente, desapareciendo una vez que se ha realizado su función o la utilidad de su tarea.

En cambio, según la misma autora, la IA será fuerte cuando puede desarrollar una multifuncionalidad, siendo capaz de aplicar la inteligencia a cualquier problema, más allá de un ámbito específico y predefinido, por ejemplo, un juez-robot capaz de admitir una demanda o una querrela, efectuar el seguimiento de los procesos, admitir o denegar pruebas, rechazar testigos, decidir el resultado del proceso, argumentar el sentido del fallo, adoptar medidas cautelares, dirigir la ejecución, etc. Esta IA es fruto de la evolución tecnológica, muy especialmente, de la posibilidad de crear aparatos inteligentes que tienen habilidad de la intercomunicación entre sí y es la que puede llegar a sustituir integralmente al ser humano.

En definitiva, cuando se habla de IA fuerte, se hace por referencia a una máquina con capacidad de aprendizaje que permitiría alcanzar soluciones autónomas a la diversidad de situaciones que la realidad puede plantearle, de manera más rápida, optimizada y fiable. Es dentro de este tipo de inteligencia, donde es posible situar los avances que se han generado a partir del aprendizaje automático y el aprendizaje profundo, cuya aplicación a diversos ámbitos del Derecho y de la justicia ya ha propiciado consecuencias, como iremos analizando.

Dejando de lado estas discusiones y cualquiera que sea el concepto de IA que se siga, como ya se señaló, se trata de un término omnicompreensivo, según Borges Blázquez (2020), utilizado a modo de cajón de sastre, para referirse a un conjunto de ciencias, teorías y técnicas dedicadas a mejorar la habilidad de las máquinas en la realización de tareas que requieren de inteligencia. En la IA se concentran componentes de distintas áreas y todas ellas pretenden arrojar un resultado de inteligencia, de modo que concurre el razonamiento automático, la demostración de teoremas, los sistemas de expertos, el procesamiento de lenguaje natural, la robótica, las redes neuronales, el aprendizaje, los lenguajes de IA, los algoritmos genéticos, entre otros.

Por lo demás, la IA alberga dentro de sí diversas técnicas avanzadas de procesamiento matemático de datos, entre ellas, Barona Vilar (2021) destaca los procesos de *Big Data*, que permiten una gestión eficiente de grandes cantidades de datos; las técnicas de *Datamining*, que

facilitan la toma de decisiones desde la creación patrones que se derivan del tratamiento de un importante volumen de datos; de *Machine Learning* (ML), que tienen por objeto el aprendizaje de las máquinas a medida que incorporan datos actualizados. En esa innovación de las máquinas, hay un área específica encargada de diseñar programas informáticos capaces de aprender por sí mismos y de realizar predicciones.

2. ¿Cómo opera la inteligencia artificial?

Dejando de lado las cuestiones relativas al concepto, y centrándonos ahora en la forma en que la IA opera, el funcionamiento de la IA tiene que ver fundamentalmente “con el procesamiento de la información para resolver problemas y tomar decisiones a partir de máquinas que operan a través de los llamados algoritmos inteligentes” (Corvalán, 2018: p. 299) que, entre muchos otros fines, se utilizan para identificar tendencias económicas, predecir delitos, diagnosticar enfermedades, predecir comportamientos digitales, etc. Para llevar a cabo este propósito de resolver problemas y tomar decisiones, Borges Blázquez (2020) da cuenta de que los sistemas de IA utilizan información que previamente ha sido proporcionada por humanos o por otras máquinas para, por medio de estos, percibir entornos reales o virtuales. Luego, abstraen esas percepciones y generan modelos. Finalmente, a partir de esos modelos, producen resultados, sea en forma de recomendaciones, predicciones o decisiones.

En otras palabras, “un sistema de IA necesita de una secuencia de instrucciones que especifique las diferentes acciones que debe ejecutar el computador para resolver un determinado problema. Esta secuencia de instrucciones es la estructura algorítmica que emplea el sistema de IA” (Navas Navarro, 2017: p. 24). De ahí que, pese a que bajo la IA se aglutinan una multiplicidad de técnicas diferentes de recolección, interpretación y procesamiento masivo de datos, todas encuentran sus bases en el uso de los algoritmos que establecen reglas lógicas para entregar una respuesta que emule a la de los seres humanos (Simón Castellano, 2021).

Respecto a cómo se ha entendido la expresión «algoritmos», diversas son las nociones que se han entregado. Así, se ha señalado que los algoritmos son “procedimientos informáticos que permiten transformar automáticamente datos en resultados apropiados para lograr un determinado objetivo” (Cerrillo i Martínez y Velasco Rico, 2019: p. 293); otra definición señala

que “es el procedimiento para encontrar la solución a un problema mediante la reducción del mismo a un conjunto de reglas” (Navas Navarro, 2017: p. 24); por su parte, la Agencia de los Derechos Fundamentales de la Unión Europea señala que “es aquella «secuencia de comandos que permite a una computadora tomar entradas y producir salidas» y añade que «usarlos puede acelerar los procesos y producir resultados más consistentes»” (Eguíluz Castañeira, 2020: p. 329).

Como se puede apreciar de las definiciones, los algoritmos son procedimientos que se conforman por una secuencia de instrucciones o comandos, desarrollados por humanos, que se dirigen a la ejecución de un fin determinado, esto es, la realización de tareas y el encontrar soluciones de forma automatizada. Se utilizan para realizar cálculos, resolver problemas, tomar decisiones o cualquier otra actividad susceptible de descomponerse o visualizarse en procesos.

Ahora bien, resulta fundamental comprender la manera en cómo funciona un algoritmo y conocer la lógica que se esconde detrás de sí. La clave para ello se encuentra, según De Miguel Beriain y Pérez Estrada (2019), en interiorizar la idea de que un algoritmo no busca certificar la veracidad de una hipótesis, sino buscar correlaciones deterministas entre unos datos y otros. Para funcionar, “el algoritmo utiliza una gran base de datos ordenados de una manera comprensible (Smart data). Más tarde, un modelo matemático irá utilizando esos datos de manera aleatoria hasta establecer patrones de correlación entre ellos que le permitirá realizar apreciaciones con exactitud” (Borges Blázquez, 2020: p. 58).

En este proceso, la IA puede apoyarse en algoritmos más avanzados y complejos denominados de aprendizaje o *Machine Learning* (ML). Esta última es la expresión que se utiliza para aludir al aprendizaje automático, o lo que es lo mismo, a la capacidad que pueden llegar a tener las máquinas de aprender por sí mismas. Eguíluz Castañeira (2020) explica que la disciplina del ML diseña modelos predictivos que construyen por sí mismos la relación entre las variables mediante el análisis de un conjunto de datos, la identificación de patrones y el establecimiento de criterios de clasificación. Una vez fijados los criterios, al introducir un nuevo conjunto de datos, el componente IA es capaz de realizar una interferencia, permitiéndole a los sistemas aprender directamente de ejemplos y experiencias.

Pese a que estos algoritmos resultan eficaces para anticipar comportamientos, tienen tras de sí un problema técnico de difícil solución: su funcionamiento, llegado a un punto de inclusión de datos y variables, resulta incomprensible para la mente humana y “no somos capaces de reconocer los patrones que ha hecho la máquina para llegar a una conclusión. Y, de serlo, tardaríamos meses, semanas o incluso años en explicar algunas de sus decisiones” (Borges Blázquez, 2020: pp. 58-59).

Lo anterior se agrava si consideramos que el aprendizaje automático, en los últimos años, ha evolucionado de tal modo en su capacidad de procesamiento, que ha dado lugar a metodologías aún más avanzadas, basadas en esta misma idea de automatización. Fruto de ello es el surgimiento de la *Deep Learning* (DL) o aprendizaje profundo que, de acuerdo a Barona Vilar (2021), son algoritmos específicos de ML, conocidos como redes neuronales, esto es, modelos matemáticos que emulan el comportamiento estructural del cerebro humano, tomando como base la interacción de las neuronas, y que han ido generando una gran expectativa.

Sobre este punto, se introduce Solar Cayón (2021) para dar cuenta que los sistemas de IA basados en algoritmos de DL resultan indescifrables por su propia naturaleza, en cuanto a que las decisiones que se obtienen no son el resultado de la ejecución de una serie de reglas previamente programadas, sino que son arquitecturas de redes neuronales capaces de identificar patrones ocultos y crear correlaciones inesperadas e indetectables para una inteligencia humana.

Este asunto guarda relación con una de las clasificaciones que es posible trazar de los sistemas de IA, y que distingue entre: los sistemas deterministas o de expertos y los sistemas no deterministas. Como indica Cotino Hueso (2019), en los primeros sistemas las reglas a partir de las cuales se procesa la información de los algoritmos se encuentran definidas, de manera que la respuesta que entrega, queda totalmente predeterminada. En cambio, en los segundos, las reglas no están completamente marcadas y el algoritmo tiene autonomía para procesar la información y producir resultados, siendo capaz de dar respuestas diferentes e imprevisibles debido al aprendizaje, las circunstancias y del entorno cambiante.

La mayoría de la tecnología basada en IA pertenece al sistema de tipo determinista, al paso que el sistema no determinista se utiliza relativamente poco debido a que estos últimos “tienen

la desventaja de la baja interpretabilidad de los resultados al perder los diseñadores la capacidad de control sobre el sistema” (Borges Blázquez, 2020: p. 39). Asunto que algunos autores como Corvalán (2018) identifican como el «lado oscuro de la IA», resultando importante poner el acento en los riesgos que se pueden generar a partir de la posibilidad de que los seres humanos pierdan el control sobre los sistemas de IA.

Todo esto, es sin perder de vista que, como se demostrará, los algoritmos no son neutros y su uso tampoco lo es, es decir, pese a que existe la tendencia a creer que la IA, al sustentarse en algoritmos matemáticos, se caracteriza por ser infalible, racional y lógica, no se debe olvidar que son creaciones humanas que utilizan datos que, previamente, una persona los ha ofrecido como ciertos. Serían más bien, “una opinión incrustada en matemáticas que, a pesar de su reputación de imparcialidad, refleja objetivos e ideologías” (Eguíluz Castañeira, 2020: p. 330).

A este problema, en el que se profundizará en los apartados siguientes, Borges Blázquez (2020) denomina «la trampa de la neutralidad» y constituye un motivo de preocupación, ya que puede dar lugar a la discriminación algorítmica, esto es, “aquellos sesgos que apareciendo en el mundo real son reproducidos en entornos de los sistemas informáticos o aquélla que surge en estos últimos producto de los datos procesados” (Reyes Olmedo, 2020: p. 6).

3. Impacto de la Inteligencia artificial en el Derecho: usos actuales.

“El hecho de que la IA pase de ser un fenómeno tecnológico a ser un fenómeno social² hace que el vínculo que tiene con el Derecho sea inevitable” (Morales Cáceres, 2021: p. 42). En ese camino de vinculación, explica Martínez Bahena (2012), los avances y desarrollos de la IA aplicada al Derecho no datan únicamente de años recientes, como erróneamente podría pensarse, sino que, desde ya hace muchas décadas, expertos se han dedicado a reflexionar y producir sistemas basados en prototipos de modelamiento del razonamiento legal y de la toma de decisiones jurídicas con el objeto de resolver todo tipo de problemáticas de relevancia jurídica.

² Un fenómeno social, por un lado, porque la IA ha penetrado todo ámbito de la cotidianidad y, por el otro, porque el ser humano ha creado un vínculo cercano con la IA al estar en contacto permanente con ella.

Prueba de lo anterior, como demuestra Barona Vilar (2021), es el hecho de que la primera Conferencia Internacional sobre la Inteligencia Artificial y Derecho fue celebrada en mayo de 1987 en la ciudad de Boston, Estados Unidos, pudiendo entenderse como el punto de partida del nacimiento de la vinculación de la IA con el Derecho. De forma paralela, esta conferencia se ha celebrado en Europa desde 1988, creándose primero, la Asociación Internacional para la IA y el Derecho y luego, en 1992, la Revista que lleva por nombre “*Artificial Intelligence and Law*”.

Antes de que ello ocurriera, las primeras investigaciones sobre IA y Derecho se remontan a los años cincuenta, aunque como indica Nava González (2019), otros autores lo atribuyen a un estudio pionero acerca de la posible aplicación de la IA al campo del Derecho publicado en 1970 por Buchanan y Headrick, y que lleva por título «Algunas especulaciones sobre Inteligencia Artificial y el Razonamiento jurídico», en donde plantean la cuestión de si el razonamiento jurídico es o no computable y, entre otras cosas, estiman que ha llegado el momento en que abogados y científicos realicen un trabajo interdisciplinario serio para explotar el potencial informático del Derecho. Este artículo se considera un detonador, pues a partir de allí, las investigaciones aumentaron y empezaron a producirse los primeros sistemas automatizados.

Luego, “con el nacimiento de las grandes bases jurídicas de datos, privadas o públicas, en los años setenta y ochenta, el interés por los sistemas de clasificación en base a relaciones se incrementó” (Casanovas, 2015: p. 832). Es por ese motivo que uno de esos modelos computacionales que más éxito tuvo fue el de los sistemas jurimétricos o la disciplina conocida como Jurimetría³, esta última realiza “una suerte de análisis lógico del derecho a fin de aplicar a las normas jurídicas modelos derivados de la lógica, el uso del ordenador en la actividad jurídica y la previsión de sentencias futuras” (Barona Vilar, 2021: pp. 360-361).

Desde sus inicios, como llama la atención Casanovas (2015), la IA aplicada al Derecho o “*Artificial Legal Intelligence*”, surgió con la finalidad de solucionar un conjunto de problemas que se plantean en el ámbito jurídico, algunos que persisten hasta el día de hoy: la organización de grandes bases de datos, la clasificación y ordenación de textos jurídicos que varían con el tiempo, el interfaz con los distintos tipos de usuarios, el modelado de las operaciones realizadas por los

³ Con todo, hay quienes afirman que la “*Artificial Legal Intelligence*” apareció junto a la Jurimetría, es decir, la computarización de la ley, propuesta a finales de los años 40 y principios de los 50 del siglo pasado, debido a la acción de la Escuela americana de la Jurimetría.

agentes políticos y jurídicos, el modelado de determinados ámbitos del conocimiento jurídico y el modelado de la argumentación razonable en función de normas o de precedentes. Por tanto, ya en esos momentos es posible hablar de máquinas que ayudan y complementan habilidades del mundo jurídico a través de la automatización de los trabajos más básicos (Barona Vilar, 2021).

Más allá de la documentación y archivística propias de la jurimetría y modelos que aplican la lógica computacional, “la IA se centró desde sus inicios en los problemas planteados por el modelado del conocimiento, los conceptos y la argumentación y razonamiento jurídicos” (Casanovas, 2015: pp. 832-833). De ahí que la principal aplicación de la IA en el campo del derecho ha sido el desarrollo de los denominados Sistemas de Expertos Jurídicos (SEJs)⁴.

Los SEJs iniciales, según Casanovas (2015), se presentaban para agilizar la tarea de recopilación de información útil para el planteamiento del caso, amén de la elaboración de documentos, explotación de los existentes, análisis de lenguajes jurídicos, etc. Posteriormente, evolucionaron en SEJs programados para actuar como verdaderos asistentes inteligentes en la solución de un problema jurídico, respondiendo a las cuestiones que les plantean los usuarios y guiándoles hacia la solución de cuestiones jurídicas, utilizando para ello el conocimiento heurístico y formal de los expertos que les asisten en su creación, diseño y alimentación, ofreciendo, además, explicaciones de sus líneas de argumentación.

Pero el modus operandi de los SEJs sigue cambiando desde SEJs iniciales a softwares que cumplen roles de asistencia como la argumentación de escritos, solicitudes o resoluciones y cuya evolución ha sido resultado de un aprendizaje estadístico, a través de la denominada DL. En este sentido, Barona Vilar (2021) da cuenta que los avances de los programas computacionales, con incorporación y manejo de estructuras de datos y procedimientos de interpretación de lenguajes, han permitido emular el pensamiento humano, como sucede con las redes neuronales artificiales, favoreciendo la toma de decisiones, han trabajado con la resolución de problemas y el aprendizaje (razonamiento basado en casos), e incluso hay SEJs que ofrecen respuestas lógicas racionales propias del ser humano e incluso van más allá, permitiendo actuar como humanos, como sucede con el robot o con los agentes inteligentes.

⁴ Sistemas computacionales que codifican el conocimiento jurídico de los expertos (humanos) y que lo manipulan como aquellos para el desempeño de las actividades de su competencia

La evolución, manifestaciones y desarrollo de la tecnología de última generación ha aterrizado en el mundo jurídico, abarcando plurales y asimétricas tareas. Se ha señalado una serie de actividades que, en el ámbito jurídico, se realizan por medio de sistemas de IA. Así tenemos: a) Analizar textos, resoluciones, documentos, extrayendo información relevante; b) Generar documentos escritos, contratos e informes; c) Realizar tareas de planificación y organización que permita coordinar agendas de jueces, fiscales, policías, abogados, favoreciendo audiencias, citaciones, comparecencias; d) Realizar predicciones económico-financieras con incidencias en el mundo del derecho, que permitan asesoramiento en mercados, cambios de estrategia inversora, determinación de riesgos en inversiones internas o internacionales; e) Redacción de demandas tanto para procesos judiciales como para procesos arbitrales; f) Planteamiento de expedientes automatizados; g) Actividad de redacción, control y valoración legislativa; h) La automatización del proceso de negociación (Navas Navarro, 2017; Barona Vilar, 2021).

Sin embargo, como se verá, los avances de la IA y los algoritmos no se conformarían, sin más, con la tarea de complementar las tareas que desarrollan operadores jurídicos humanos, sino que cada vez se avanza en este proceso de sustitución, en donde la máquina inteligente comienza a sustituirlos, caminando hacia un posible reemplazo del hombre por la IA. En el caso de la justicia y de la función jurisdiccional, implicaría el tratar de emular no solo una de las posibles funciones que desempeña un juez, sino que la función judicial en su totalidad, en una suerte de incorporación de estructuras inteligentes diseñadas para sustituir a los jueces.

Nava González (citada por Rincón y Martínez, 2021) considera que, hasta la fecha, los desarrollos de la IA aplicada al Derecho se han orientado a los siguientes modelos: 1) Modelos para la argumentación y la toma de decisiones; 2) Clasificación y extracción de textos legales; 3) La extracción de información de textos legales, y 4) La creación y planificación de un sistema legislativo. De estos ámbitos, nos quedaremos con el primero, asistiendo al debate que, en el ámbito de la administración de justicia, ha surgido en torno a las posibilidades de aplicar sistemas de IA para simplificar y automatizar distintas fases del proceso judicial, para introducir en los tribunales procedimientos automatizados de negociación y mediación, para asistir a los jueces en la toma de decisiones o incluso tomar directamente esas decisiones en sustitución de ellos, emulando función judicial en su totalidad. Estos resultan ser asuntos sumamente delicados y complejos, con importantes implicancias teóricas y prácticas.

II. Uso de algoritmos en sede judicial: ventajas y desventajas.

Como explica Barona Vilar (2021), sucesivamente en la Justicia comienza a introducirse la idea de que los algoritmos asuman funciones que, hasta el momento, solo realizaban los operadores jurídicos humanos, de manera que puedan no solo complementar, como sucede en la actualidad con los múltiples software que facilitan, dan información y también fijan criterios a considerar en el ejercicio de la función jurisdiccional, sino también permitir la automatización judicial a través de sistemas computacionales, propiciando la robotización judicial.

Lo anterior, lejos de ser futurible, es una realidad ya actual, que emerge con fuerza e irradia diversas manifestaciones de la capacidad de discernir y de alcanzar soluciones a los conflictos jurídicos, ya sea por vía judicial, o por vías extrajurisdiccionales o extrajudiciales, como el arbitraje. Solar Cayón (2021), sobre este punto, manifiesta que, en los últimos años, diversos países han introducido en sus sistemas de justicia mecanismos para la resolución automática de conflictos, inspirándose en los sistemas de resolución automática de disputas en línea desarrolladas por compañías privadas en el ámbito del comercio electrónico. Algunos de los casos son los siguientes: Canadá y el ‘British Columbia Civil Resolution Tribunal’, el Reino Unido y su tribunal en línea ‘Online Solutions Court’, iniciativas similares en China, Francia y Estonia. Asimismo, sistemas para elaborar propuestas de resolución judicial como ‘Prometea’ en Argentina o para auxiliar a los tribunales en la toma de determinadas decisiones en Colombia.

Ahora bien, entre los motivos por los cuales se comienza a optar por este tipo de administración de justicia, Martínez Bahena (2012) señala, el poder proporcionar rendimientos más eficientes que aquellos producidos por el ser humano ya que las nuevas tecnologías disminuyen drásticamente las limitaciones por razones de espacio, tiempo y de saturación laboral, estando siempre en pleno rendimiento. Otra razón argüida en su favor, indicada por Batelli (2020), es la exactitud de una decisión objetiva, mucho más precisa, neutra y rápida que la que pueden proporcionar los seres humanos, siendo una exigencia compatible con la despersonalización de la figura del juez. Estos asuntos pasarán a analizarse a continuación.

1. Ventajas del uso de sistemas algorítmicos de IA

1.1. Eficiencia y descongestión.

Cotino Hueso (2019), al hacerse cargo de las bondades del uso de la IA, señala que en todos los ámbitos la IA puede mejorar, y mucho, la innovación, eficiencia, eficacia, competitividad o calidad de las decisiones. Así, por ejemplo, en el sector público, puede usarse para innovar y mejorar la eficiencia, la eficacia y los tiempos, así como especializar la calidad técnica y jurídica de las actuaciones administrativas y la prestación de servicios públicos. Es más, puede anticiparse a las necesidades de cada sujeto con el objeto de prestarle servicios precisos.

Por su parte, Rincón Cárdenas y Martínez Molano (2021), aludiendo ahora al ámbito de la justicia, indica que sabido es que muchos sistemas judiciales, especialmente en Latinoamérica, han tenido a lo largo del tiempo grandes problemas en relación con la baja producción, la congestión y la mora judicial por volúmenes de demanda superiores a la capacidad de respuesta y demora en la resolución, lo que afecta directamente el acceso a la justicia. Por ello, uno de los motivos que podría llevar a optar por una administración de justicia basada en algoritmos, sería el de proporcionar rendimientos más eficientes que aquellos producidos por el ser humano, ya que este tipo de IA se muestra como una herramienta capaz de procesar enormes cantidades de datos y conocimientos, logrando igualar o superar ampliamente ciertas capacidades cognitivas, a partir de procesar más eficientemente los datos (Batelli, 2020).

Una justicia que implementa sistemas de IA puede gestionar y dirigir claramente de forma automatizada un procedimiento, puede analizar datos, información y documentos y explotar la información ofreciendo una suerte de justificación que simula la decisión judicial. Además, permitiría, entre otras cosas, reemplazar la fuerza laboral, pudiendo, por ejemplo, desarrollar sus funciones las 24 horas del día, sin necesidad de vacaciones, remuneración, alimentación, no disminuyendo su rendimiento por factores externos. Esto causaría, como consecuencia directa, un aumento en la productividad laboral, lo significa una ventaja desde cualquier punto de vista.

Un ejemplo que da cuenta de la eficiencia que es posible alcanzar se puede observar en Prometea. En un trabajo reciente (Rincón Cárdenas y Martínez Molano, 2021), basado en un estudio publicado en el año 2018, se puso en evidencia que los seres humanos tardan

aproximadamente 172 días en resolver 1.000 casos judiciales de baja complejidad, al paso que Prometea puede dar respuesta a estos mismos 1.000 casos en 42 días.

Con todo, en esta reflexión sobre la eficiencia y eficacia del sistema judicial, habría que tener en cuenta si la descongestión que se produce a partir de ellas serían razones suficientes para justificar los costos e impactos en los derechos y garantías de los justiciables. Sobre el particular, Bellosó Martín expresa el punto de vista de la Comisión Europea para la Eficiencia de la Justicia (CEPEJ) del Consejo de Europa y señala que:

“La CEPEJ, tal como se establece en la Carta, parte de la asunción de que la aplicación de la IA en el ámbito de la justicia puede contribuir a mejorar su eficiencia y la calidad. Pero, a la vez, que debe aplicarse de manera responsable, respetando los derechos fundamentales garantizados por el ordenamiento y, en particular, por el Convenio Europeo de Derechos Humanos (CEDH) y por el Convenio del Consejo de Europa para la Protección de Datos Personales” (2019: pp. 11-12)

De ahí la importancia y preocupación por tener siempre un enfoque a partir de los riesgos que puede implicar el implementar estas tecnologías, pues “por muy automática que acabe siendo una parte sustancial de la labor del poder judicial, no se pueden olvidar jamás los derechos humanos” (Nieva Fenoll, 2018: p. 128).

En definitiva, es una realidad en la que se debe sopesar la opinión de quienes señalan que la IA aplicada a la justicia constituye una herramienta que ofrece grandes dosis de eficiencia procesal, y que se anuda a una respuesta que combina la calidad de la toma de decisiones con una reducción de la actividad judicial y la opinión de aquellos a los que, la aplicación de la IA, les genera numerosas dudas en torno a la independencia judicial, especialmente si se tiene en cuenta que las decisiones judiciales ya no serán el fruto del ejercicio humano de la función jurisdiccional, sino del ejercicio algorítmico, generándose una especie de determinismo jurisdiccional.

1.2. Objetividad, certeza y despersonalización de la función jurisdiccional.

En la toma de decisiones judiciales concurre un factor que ha ejercido un rol fundamental en la función jurisdiccional, esto es, la predisposición subjetiva del juez que comprende el

conocimiento de las normas, su capacitación y la adaptación y/o interpretación de las mismas al caso concreto, en atención a las circunstancias que concurren y a los sujetos afectados. Estos factores deben tomarse con precaución y con sentido crítico, precisamente, por la discrecionalidad judicial, que puede generar desigualdades, amén de dilatar los procesos y exteriorizar prejuicios.

No se debe perder de vista que, aunque no se quiera, en la toma de decisiones judiciales, juega un papel esencial, en cuanto al ser humano se trata, los afectos, las emociones, las sensibilidades, y en general el componente subjetivo, aun cuando se pretenda objetivarlo. Sobre el punto, Nieva Fenoll ha señalado que:

“Se ha dicho, no sin motivo, que los seres humanos no somos tan racionales como parecemos, en el sentido de que debiéramos ser siempre eficientes y consistentes, y sin embargo no lo somos, a diferencia de las máquinas, que ni se fatigan, ni se distraen o se aburren, ni tienen emociones. Por ello se ha llegado a afirmar, probablemente de forma un tanto pretenciosa, que en el futuro los jueces serán cada vez menos necesarios” (2018: p. 44).

Es ahí donde se muestra otra de las ventajas de la IA, ya que se estima que podría eliminar la discrecionalidad y, en consecuencia, las discriminaciones injustas del sistema judicial. Quedaría fuera de órbita la mirada subjetiva que tienen los jueces en cuanto humanos en vista de garantizar que los tribunales impartan justicia de manera consistente y objetiva, y con ello alcanzar mayor certeza jurídica. Como expresa Batelli “la justicia robótica parece, en el mejor de los casos, un "paradigma ideal-típico" que nace como reacción ante aquellas doctrinas de la interpretación que a menudo acabaron legitimando la arbitrariedad y el subjetivismo del juez” (2020: p. 54).

Aunque, igualmente, subsiste la duda de si realmente la toma de decisiones judiciales, por parte de máquinas, es realmente más idónea que la de un juez, un cuerpo colegiado de jueces o un jurado de ciudadanos, quienes tienen experiencias de vida, están envueltos de una cultura y procesos de resiliencia que ponen a su vista contextos ajenos a su propia realidad que, aunque le sean desconocidos, pueden provocar su sensibilización, teniendo la posibilidad de un moldeaje más justo que correcto (Vivar Vera, 2021) o si acaso la falta de sensibilidad y comprensión de

esos elementos sociales por parte de las máquinas generaría, más que una objetivación óptima, una comprensión distorsionada de la realidad que la alejaría de la noción de justicia.

2. Desventajas o problemas del uso de algoritmos en sede judicial.

Como se ha señalado, resulta evidente que el empleo de la IA en el ámbito judicial importa una serie de aspectos positivos y un motor de avance, pero al mismo tiempo, se configura como un elemento distorsionador. Claro está, no se trata de poner trabas al progreso a través de la tecnología, sino que el Derecho y sus operadores sean conscientes de los riesgos y peligros que comportan para las garantías y los derechos fundamentales, estableciéndose estos como un verdadero contrapeso, para evitar, de esa forma, que se transformen en armas antojadizas.

Un ejemplo en el que esto queda en evidencia es el que tiene lugar en los sistemas de justicia penal que utilizan la puntuación de riesgo automatizada, donde su empleo puede, al mismo tiempo, reducir el número de individuos pertenecientes a grupos mayoritarios que son innecesariamente encarcelados, pero también las deficiencias del sistema pueden aumentar la tasa de encarcelamientos erróneos de individuos que pertenecen a grupos marginados (Castellanos Claramunt y Montero Caro, 2020). Algunas de estas desventajas se analizan a continuación.

2.1. La trampa de la neutralidad: sesgos y discriminación algorítmica

La primera de estas distorsiones, cuestión central de este trabajo, ya fue adelantada y refiere a que la toma de decisiones, a partir de sistemas de IA basados en algoritmos predictivos, se suele presentar como un cuerpo preciso e infalible, que promete un mayor grado de rigor y consistencia que el que puede ofrecer un juez humano. Sin embargo, concebir a la IA despojada de cualquier sesgo es, al menos, imprudente ya que su uso puede producir fácilmente resultados discriminatorios.

La concurrencia de sesgos y de discriminación algorítmica es palmaria, y para entender estos fenómenos resulta preciso destacar lo señalado por Reyes Olmedo, para quien:

“La discriminación es un trato diferente y perjudicial que se da a una persona debido a categorizaciones arbitrarias o irrelevantes y que en virtud del derecho humano fundamental de igualdad ante la ley se encuentra prohibida. En el caso, la calificamos de “algorítmica” porque aparece a partir del uso de algoritmos utilizados por el aprendizaje automatizado y de la inteligencia artificial” (2020: p. 6).

En estos sistemas de IA, “los sesgos algorítmicos pueden devenir de los propios patrones no aleatorios en los que se plasma la realidad en cuestiones como la edad, el género, o la etnia, mientras que otros pueden producirse por defectos analíticos por falta de datos” (Miró Llinares, 2018: p. 121). En otras palabras, se produciría, o bien, porque los datos reflejan ciertos sesgos personales de quienes recogieron los datos o programaron los algoritmos, o bien, por adolecer de algún defecto que devenga en su imperfección, ya sea por ser insuficientes, excesivos, desactualizados o erróneos.

Según el mismo autor, existen dos tipos de sesgos en los algoritmos: sesgos en los datos de entrenamiento y sesgos por una distribución desigual real de las variables. Para entender los primeros, se debe tener en cuenta que las IA actuales se basan en modelos de aprendizaje automático, es decir, aprenden con nuevos ejemplos⁵. Por ello, se ha hablado de un triple riesgo de sesgo: 1) que la máquina haya sido alimentada con datos erróneamente calificados por el programador; 2) que el set de datos inicial sea el resultado de un muestreo no representativo y 3) que los parámetros del aprendizaje se realicen sobre intervalos temporales fuertemente acotados⁶. En lo que se refiere al segundo grupo de sesgos, corresponden a variables que efectivamente se distribuyen de manera desigual en la realidad, factores aparentemente neutrales recogidos por la IA, pero que no se refieren al actuar de la persona en concreto, sino al actuar pasado de sus iguales, con los que comparte, por ejemplo, género, etnia, edad, entre otros.

Este último tipo de sesgos se observan en ciertos sistemas de IA predictiva que se utilizan en Estados Unidos –como COMPAS– que se basan en un código de fuente que toma en consideración aspectos de raza y género, como ocurrió en el caso *State v. Loomis*, que se analiza

⁵ En contraposición con aquellos algoritmos basados en reglas precodificadas y completamente especificadas.

⁶ La inobservancia de este tipo de limitaciones acerca de los datos utilizados en estos algoritmos predictivos perpetúa, en efecto, la discriminación algorítmica hacia los grupos infrarrepresentados, pudiendo lesionar o poner en riesgo sus derechos fundamentales.

más adelante, y que constituye uno de los casos más notables en la materia (Corvalán, 2018). En ese contexto, resulta trascendental asegurar los principios de igualdad y de no discriminación de los individuos y conciliarlos con el uso de estas tecnologías, de manera que la IA que se implemente sea más justo que COMPAS.

Con todo, hay quienes advierten, que estos sesgos “podrían no ser peores que los que se producen cuando el análisis para la toma de decisión lo llevan a cabo personas sin el uso de sistemas informáticos” (Miró Llinares, 2018: p. 121). Además, cuando la IA se canaliza a través de sistemas más sofisticados, la injerencia de sus programadores disminuye, reduciéndose sus potenciales sesgos, pero da paso al siguiente inconveniente: la imposibilidad de explicar por qué se llega al resultado emitido, tratándose de técnicas cargadas de secretos y falta de transparencia.

2.2. Falta de transparencia y opacidad de los sistemas de aprendizaje profundo

La segunda distorsión dice relación con ciertos aspectos que pueden dificultar la transparencia requerida para la utilización de este tipo de herramientas y que se refieren, según Solar Cayón (2021), a dos aspectos: por un lado, a factores jurídicos, relativos al régimen de protección de los algoritmos, propiedad en su mayoría de compañías privadas, quienes no revelan el código de fuente –por el secreto empresarial– y, por otro lado, ciertos factores técnicos que tienen que ver con la opacidad de los algoritmos de DL.

A este último problema se le denomina *black box* o caja negra, que, en términos sencillos, se refiere a una “herramienta computacional en la que uno entiende los datos ingresados y los resultados, pero no comprende el procedimiento subyacente” (Corvalán, 2018: p. 305). Esto es así, en opinión de Boix Palop (2020), porque los sistemas de IA basados en algoritmos de DL funcionan de manera tal, que dependen no solo de operaciones e inferencias probabilísticas muy complejas (lo que, hasta cierto punto, y aun con costos enormes, podría ser calculado por la inteligencia humanas) sino, también, de un tipo de programación que permite una paulatina y sucesiva evolución del propio programa y de las soluciones a las que va decantando.

En otras palabras, su funcionamiento resulta indescifrable, lo que implica que ni el sistema, ni sus diseñadores pueden explicar el porqué de sus resultados y predicciones, llegando a esas decisiones sin ayuda, es decir, de una manera que no está especificada de antemano. Lo anterior,

se tornaría particularmente grave si es que este sistema de IA se aplica para decidir sobre la condena o absolución de una persona, pero no deja de tener relevancia en otros asuntos sometidos a conocimiento de un tribunal, como cuestiones de derecho de familia como sería el decidir sobre cuidado personal de los hijos, o incluso en asuntos patrimoniales.

Por otro lado, esta distorsión no sólo apunta a lo anterior, pues, como indica Hernández y Polanco la idea de caja negra “hace referencia no sólo a la opacidad de la información, sino también a su legibilidad, esto es, a la capacidad de que la información y los resultados puedan ser interpretados y aprehendidos por los usuarios afectados” (2020: p. 93).

Es a propósito de la opacidad y falta de transparencia que se genera uno de los debates más importantes en torno al empleo de la IA en la justicia entre quienes opinan, como advierte Barona Vilar (2021), que no es posible tolerar de ningún modo que un software o máquina tome decisiones tan importantes sin que sepamos por qué lo ha hecho, de manera que hasta que no se tenga una auténtica IA explicada, no se podría utilizar redes neuronales para impartir justicia, manteniéndose serias dudas acerca de si efectivamente un “juez robot” puede sustituir al juez humano sin vulnerar el modelo judicial de garantías.

Sin embargo, otros si bien se inclinan y adhieren a la demanda de mayor transparencia, igualmente dudan acerca de si el tipo de transparencia que se esconde detrás de muchas de estas demandas sea siempre útil. En ese sentido se han decantado algunos autores (Zerilli et al., 2019), para quienes ese paradigma de la transparencia pueda resultar igualmente ilusorio en relación a los procesos de decisiones humanas. Argumentan que es cierto que los agentes humanos pueden dar razones de sus decisiones, pero eso no es lo mismo que iluminar los procesos cognitivos que conducen a una conclusión. Estos movimientos mentales que subyacen a las elecciones humanas, especialmente en áreas en las que un elemento crucial de la intuición y la impresión personal están impulsando la deliberación, están, de hecho, lejos de ser transparentes. Por lo que sería un ideal que se basaría en una estimación completamente irrealista del grado de transparencia que es posible obtener en las decisiones humanas y la complejidad de sus procesos mentales, exigiéndosele a las máquinas lo que no se les exige a los humanos.

3. Sesgos algorítmicos en sistemas de justicia criminal: casos emblemáticos para visibilizar la trampa de la neutralidad.

La IA se ha utilizado en los sistemas jurídicos con diversos fines, entre estos, facilitar el acceso a la justicia, por ejemplo, mediante los *chatbots* o a través de medidas alternativas de solución de conflictos automatizada. Pero más allá de ello, “un bloque más concreto de usos son los relativos a la sustitución o apoyo por parte de los sistemas de IA aplicados al proceso de determinación judicial de la responsabilidad por la perpetración de un delito” (Simón Castellano, 2021: p. 7), asunto que despierta un importante interés dogmático por las consecuencias que puede traer a los derechos fundamentales y por la duda de si la IA va a reforzar el proceso penal garantista o ayudarnos a liquidarlo.

Gran parte de los trabajos sobre la materia, cuando abordan esta concreta cuestión, hacen referencia al trascendente caso *State v. Loomis* y la sentencia que a su respecto dictó la Corte Suprema de Wisconsin, el año 2016. Una investigación (Borges Blázquez, 2020) que se hace cargo de este caso, relata que el ciudadano Eric Loomis fue detenido por la policía de Wisconsin el año 2013, imputándosele una serie de delitos ligados a un tiroteo desde un automóvil: posesión de armas de fuego, intento de evasión, puesta en peligro de la seguridad pública y conducción de vehículo ajeno sin autorización. Loomis negó haber participado del tiroteo, pero admitió haber conducido un vehículo no autorizado, logrando acordar con la fiscalía la concurrencia de ese delito y del de puesta en peligro a la seguridad pública, quedando libre de los demás cargos.

El juez de primera instancia aceptó el acuerdo pero al momento de conmensurar la pena recurrió a COMPAS, un sistema de evaluación de riesgos, el que determinó que Loomis presentaba un riesgo extremadamente alto de reincidencia. En consecuencia, en vista del informe algorítmico, la gravedad del crimen y su historial en materia de supervisiones judiciales, se denegó la suspensión del procedimiento y se fijó una pena de seis años de presidio y cinco de supervisión postpenitenciaria (Martínez Garay, 2018).

Loomis solicitó al tribunal la revisión de la pena, argumentado una vulneración al debido proceso (*Due process*), lo que fue negado, puesto que, según el juez, COMPAS sólo venía a corroborar las estimaciones realizadas jurisdiccionalmente, por lo que la pena hubiese sido la

misma, independientemente de su uso. Frente a esto, nace el caso *State v. Loomis*, al recurrir Loomis al Tribunal Supremo de Wisconsin (Browning y Shahrestani, 2019).

Eric Loomis señaló como razones que: 1) Se habría infringido el derecho a ser juzgado por medio de datos fiables y precisos. Esto se debe a que COMPAS es una tecnología desarrollada por la compañía privada Northpoint y se rige por el secreto comercial, de manera que el procesamiento de información que realiza resulta inaccesible para el condenado, afectando su derecho de defensa; 2) No se le habría dictado una sentencia plenamente individualizada ya que el sistema se basa en datos de casos pretéritos que agrupa matemáticamente y; 3) Su género se tuvo en cuenta de manera indebida para determinar la pena (Freeman, 2016).

Ninguno de los motivos fue admitido por la corte, la cual ratificó la sentencia y desentrañó cada punto respectivo de la siguiente manera: 1) El derecho a la información no fue vulnerado porque el secreto comercial se estimó esencial para el negocio y porque sí se accedió a los datos objetivos empleados por la IA, que básicamente eran su historial penal; 2) No se vulnera el derecho a sentencia individualizada, pues, el fallo no se basó únicamente en aquella tecnología, sino que fue sólo un complemento; 3) No habría una discriminación de género indebida ya que la ciencia ha verificado una mayor reincidencia y criminalidad violenta de los varones por sobre las mujeres. La omisión de este elemento aumentaría el margen de error del sistema, lo que afectaría tanto al acusado como a la administración de justicia (Martínez Garay, 2018).

Con todo, la importancia del caso reside en que el Tribunal Supremo estableció limitaciones sobre cómo y para qué pueden ser usadas las evaluaciones de riesgo. Estos sistemas no pueden emplearse para decidir la privación de libertad ni para determinar el grado de la pena, debiendo limitarse a las siguientes hipótesis: 1) Proposición de medidas alternativas a delincuentes de bajo riesgo; 2) Señalamiento de la posibilidad de supervisión segura y efectiva por la comunidad; 3) Sugerencia de plazos, condiciones de suspensión de la pena, libertad condicional, y consecuencias de su infracción (Browning y Shahrestani, 2019).

Este caso marcó la historia procesal de Estados Unidos, pues, consolidó la admisibilidad de los algoritmos y diversos tribunales han ido expresando su voluntad de hacer uso de las herramientas de IA para mejorar el funcionamiento de la justicia, aunque también instaló una

serie de dudas relativas a los derechos fundamentales, ya que en definitiva se corre el riesgo de ratificar y extender los sesgos emanados de la discrecionalidad humana (Freeman, 2016). Esto último no es una mera especulación, sino que una realidad ya palpable en lugares como Estados Unidos, dónde reportajes y artículos han advertido situaciones de sesgo algorítmico ligados a temas de género –como se vio en *State v. Loomis*– así como de raza, temas que vienen a demostrar que los algoritmos no hacen más que reflejar los prejuicios del ser humano como especie productora de estos instrumentos (Browning y Shahrestani, 2019).

Bien es sabida la compleja situación estadounidense en relación con su composición étnica, la violencia motivada por ella y estigmatizaciones históricas que perviven hasta hoy. En efecto, se ha acusado que los sistemas de evaluación de riesgos tienden a agravar la situación procesal de unas personas por sobre otras. A modo ejemplar, en una investigación desarrollada por ProPública en el año 2016, se llegó a la conclusión que COMPAS “exageraba el riesgo de reincidencia en personas afroamericanas como más propensas a volver a cometer crímenes, mientras que la raza blanca no era considerada un factor de riesgo” (Belloso Martín, 2019: p. 19).

No es que la máquina se configure deliberadamente para generar discriminaciones por racismo, sino que, al recurrir estos sistemas a archivos del pasado, traen de antemano la excesiva representación histórica que tienen las minorías en los historiales criminales en el sentido de ser proporcionalmente más condenadas y pasar más tiempo en prisión, y estos antecedentes acaban por perpetuar los sesgos que la institucionalidad ha mantenido por largo tiempo (Tonry, 2014). En informática se ha empleado al respecto la expresión “*Bias In, Bias Out*”, y la consecuencia que genera es que, al recurrir a una visión fragmentada de la realidad, termina por predecir el futuro mirando al pasado, esto es, determinando la concurrencia de delito, su gravedad y consecuencias jurídicas en base a la historia colectiva más que al caso concreto (Mayson, 2019).

Otro caso relevante, en ese sentido, es el de Brisha Borden, una joven afroamericana de Florida que, junto con su amiga, fueron arrestadas y acusadas de robo y hurto de un scooter y una bicicleta por un monto de 80 dólares. Teniendo en cuenta un historial de delitos menores perpetrados antes de cumplir la mayoría de edad, el sistema de evaluación de riesgo predijo un alto riesgo de reincidencia. Un año antes, un hombre blanco llamado Vernon Prater fue detenido por robo de tienda en un valor de 86,35 dólares, cuantía similar al otro caso. Sin embargo, siendo

su historial delictual mucho más lesivo —abarcando robos con armas—, ese mismo sistema que evaluó a Brisha, lo calificó con un bajo riesgo de reincidencia. Lo cierto es que la máquina, además de dar una mayor puntuación de riesgo a personas negras que a personas blancas, no predijo correctamente, pues Brisha Borden no reincidió en absoluto, mientras que Vernon Prater sí lo hizo, pasando 8 años en prisión por un robo posterior (Borges Blázquez, 2021b)⁷.

Ahora, así como en Estados Unidos se emplea COMPAS, “con propósitos similares pero un alcance geográficamente más acotado, el equivalente europeo podría ser HART (*Harm Assessment Risk Tool*) utilizada en Durham para informar sobre el riesgo de reincidencia del delincuente” (Miró Llinares, 2018: p. 110). Al igual que COMPAS, HART se sostiene mediante diversos factores de caracterización de las personas y de su historial criminal, y académicos se han mostrado igualmente críticos ante posibles sesgos como los ya mencionados, pero también sesgos de clase: el uso de esta información llevaría a perpetuar mayores custodias para personas de menos recursos. Ello se suma a la cantidad ingente de información que maneja HART, sumamente interconectada, lo que dificulta su transparencia (Burgess, 2018; Oswald et al., 2018; Simón Castellano, 2021).

El problema de los algoritmos en relación con aspectos como el género ha supuesto incluso reprimendas jurídicas a instituciones, como ocurrió en España, en la localidad de Sanlúcar la Mayor, Andalucía en el año 2016, en el que Ángela González solicitó ante la Guardia Civil una orden de protección alegando ser víctima de violencia de género (Casanueva, 2020). El departamento policial en cuestión consta para el efecto de un cuestionario de cribado, en virtud del cual, la IA evalúa el riesgo en que se encuentra la persona, y el resultado fue de riesgo bajo. En base a esta información brindada por la tecnología, y sin ahondar más, la Guardia Civil calificó al caso con un riesgo no percibido, valoración que fue determinante para que el juzgado rechazara de plano la medida de protección (Borges Blázquez, 2020).

Al cabo de un mes, Ángela fue asesinada por su marido, algo que la autoridad administrativa —que terminó por sancionar al Ministerio del Interior cuatro años después— calificó como un

⁷ Como advierte Barona Vilar (2021), casos como los señalados demuestran que los sistemas de predicción, al incluir grandes dosis de sesgos, pueden favorecer el derecho penal de autor. Se abre la puerta a que la gravedad de la condena no se base ya en los hechos cometidos, sino que incluya circunstancias personales y elementos de la personalidad.

hecho evitable, pues la Guardia Civil se quedó sólo con los aspectos formales y burocráticos y no ahondó en una relación directa y protectora con la víctima, siendo “la falta de concienciación social e institucional sobre la violencia de género unida a la confianza ciega que muchas veces depositamos en los sistemas de IA la que conllevó al fallecimiento de la solicitante de protección” (Borges Blázquez, 2020: p. 66). Esa confianza ciega en la máquina genera un anclaje inconsciente a lo que ésta dice, sólo que esta vez es por parte de las policías, quienes también forman parte importante del engranaje de la justicia estatal (De Miguel Beriain y Pérez Estrada, 2019).

Esta clase de sesgos, se han ido repitiendo en el historial jurisprudencial y policial recientes, y demuestran una imperfección práctica de estos sistemas en su origen y configuración. Estos problemas opacan la eficiencia e imparcialidad que los entusiastas de estas tecnologías proponen, al perpetuar la discrecionalidad jurisprudencial y los prejuicios históricos de una comunidad determinada en un ciclo de no acabar (Martínez Garay, 2018). Aunque, como bien advierte Borges Blázquez (2021), lejos de quedarnos en la lectura pesimista, hay que aprender del pasado y ser capaces de corregirlo. Sin intervención, el futuro repetirá los datos ya que el análisis predictivo es un espejo del pasado que se proyecta en el futuro. Predecir el futuro bajo las condiciones *status quo* es simplemente proyectar las desigualdades.

Por tanto, la IA lejos de ser objetiva, está impregnada de situaciones humanas, contextos comunitarios e individuales desiguales que son base para la decisión judicial. Por ello, “si queremos emplear la IA en el sistema penal, tenemos que ser conscientes de que ninguna herramienta de gestión del riesgo va a ser capaz de dar marcha atrás y cambiar siglos de injusticia racial y desigualdad de género” (Borges Blázquez, 2021: p. 163).

Ha quedado expuesto qué pros y contras conviven en el paisaje algorítmico, no sólo desde un punto de vista teórico, sino que en la existencia de casos de la vida real. Por tanto, pese a que uno de los argumentos que se esgrimen a favor de la decisión judicial algoritmizada es la neutralidad, entendida como ausencia de subjetividad en la decisión judicial, la concurrencia de sesgos niega tal argumento ¿Por qué no podrían replicarse esos sesgos en la administración de justicia en nuestro país, que se caracteriza por la desigualdad de su población, por el encarcelamiento masivo y la criminalización de la pobreza? ¿Qué podemos hacer para evitarlos?

III. Sistemas, modelos y experiencias de las aplicaciones de IA en el mundo y en Chile

En un apartado precedente, se dio cuenta la forma en que la IA ha impactado el ámbito del Derecho y cómo su relación ha ido evolucionando en el tiempo. El paso siguiente será el de configurar las funciones de los algoritmos en el ámbito de la justicia para, posteriormente, evidenciar diversas experiencias que existen en el mundo.

La aplicación de la IA al campo de la justicia, como indica Miró Llinares (2018), suele denominarse “Inteligencia Artificial Judicial”, así, uno de los principales intereses se encuentra en los sistemas de IA que informa a los tribunales para la toma de decisiones en el ámbito penal.

Respecto al primer cometido, Barona Vilar distingue las siguientes funciones de la IA en el ámbito de la Justicia: “a) Asistencial (como complemento o facilitador); b) En la prevención (a través, aunque no solo, de la irrupción de la Justicia predictiva); c) En la investigación penal; d) En la toma de decisiones judiciales (complemento o sustitución del juez por el juez-robot); y e) En la ejecución de sentencias” (2021: pp. 363-364).

Es importante tener en cuenta que los niveles de inteligencia, en cada una de estas funciones, serán asimétricos y heterogéneos, de modo que, en ciertos casos, van a asistir y complementar las habilidades humanas, mientras que en otros se plantea el reemplazo del humano por la máquina o la integración máquina-humano, lo que se denomina “*cyborg*” o “híbrido”, siendo la IA la que cubre las falencias de actuación o decisión del ser humano.

Tratándose de la primera de las funciones, en el mercado judicial son variados los modelos con la experticia de un asesor jurídico que permiten, por ejemplo, el diagnóstico legal, programados para actuar como asistentes inteligentes en la solución de problemas jurídicos. Se puede citar como ejemplos: ARAUCARIA, creado en la Universidad Dundee de Escocia, “un software que apoya el análisis y la reconstrucción de textos argumentativos mediante la diagramación” (Mora Pedreros, 2017: p. 60); DART que es un soporte para argumentación jurídica, o RESONABLE o CONVINCER ME, que dan soporte a los argumentos; CASEMAP, una base de datos relacional diseñada específicamente para aplicaciones relacionadas con los juicios, y que permitiría preparar la defensa de los clientes (Barona Vilar, 2021).

Otros sistemas que prestan un servicio de asistencia son aquellos que crean textos jurídicos (contratos, informes, documentos, dictámenes judiciales, entre otros) y que ayudan a la realización de tareas organizativas de los operadores jurídicos. Un ejemplo cercano de un sistema de IA que permite crear dictámenes judiciales es PROMETEA, el “primer sistema de inteligencia artificial jurídico en Latinoamérica que trabaja con un asistente de voz (tal como lo hace Siri de Apple) y permite realizar un dictamen jurídico de manera íntegra” (Corvalán, 2018: p. 303), y cuyo estudio se difiere al último capítulo de este trabajo, por las ventajas que su uso puede ofrecer al sistema judicial chileno.

La IA y los algoritmos se ha adentrado también en el ámbito de la predicción, prevención e investigación criminales. Por ejemplo, PROTOBADI, una aplicación utilizada en Bangladesh que avisa a sus usuarias de la posible cercanía de personas con tendencias delictivas sexuales, alertando a posibles víctimas de posibles ataques o abusos sexuales; asimismo, se puede indicar la aparición de las herramientas estructuradas de valoración de riesgo como el empleo de Big Data en la ciudad de Chicago y COMPAS (Borges Blázquez, 2021a).

En la ejecución de las sentencias, Barona Villar (2021) señala que países más avanzados en materia de IA, como China, han comenzado a incorporar como piloto, los modelos de prisiones SMART PRISON, en los que la IA se despliega en la construcción de los edificios y en su funcionamiento cotidiano, en los sistemas de vigilancia con robots y con modelos computacionales de control de los presos, generando sistemas de contacto de éstos con el sistema de justicia a través de procedimientos electrónicos. En el caso de Suecia, en un estudio realizado por Kaun y Stiernstedt (2020), se señala un número creciente de iniciativas de prisiones inteligentes, por ejemplo, IA para analizar imágenes del circuito cerrado o para evitar peleas e intentos de fugas, pero no sólo se trata de tecnología inteligente para fines de vigilancia y fines administrativos, sino que también se incluye la IA en el trabajo de rehabilitación de las personas encarceladas, como por ejemplo, UTSIKT, una aplicación cuyo objeto principal es mejorar la asistencia a las reuniones de libertad condicional y proporcionar a sus usuarios directrices que les brinden apoyo en episodios críticos, con el objeto de prevenir la reincidencia.

Ahora bien, en lo que se refiere a la toma de decisiones judiciales, como ya ha sido señalado en reiteradas ocasiones, uno de los debates más intensos que se generan en torno a la justicia es

el del uso instrumental de los algoritmos y la IA, como “Justicia-máquina perfecta”, tanto como “complemento” de la decisión judicial, a través de la incorporación de criterios que favorecen, mediante un sistema de algoritmos, la decisión de los jueces, como respecto de la posibilidad de que sea la misma máquina la que tome la decisión judicial. Son dos modalidades de aplicación de la IA al proceso: empleándose el sistema de expertos o la robotización judicial.

Por tanto, el paso siguiente, en el desarrollo de esta tecnología, sería ya no trabajar integrando máquina-persona en un sistema de expertos, sino que, ante la acción de la IA de última generación, el operador jurídico se tornaría innecesario.

A continuación se describe el estado de desarrollo en algunos países analizados:

1. El caso de China

En China, a partir del año 2017, se han creado los TRIBUNALES DE INTERNET. El primero de ellos, como indica un artículo en la materia (Bryan Lynn, 2019), se creó en la ciudad de Hangzhou, luego en Beijing y Guangzhou, siendo competentes para conocer disputas por contratos que involucran compras, servicios y pequeños préstamos en línea o derechos de autor.

Son sistemas de litigación online en el que una IA, dotada de una interfaz audiovisual, se comunica con los litigantes, resuelve las controversias asignadas y consultas relativas a los procedimientos de los que conoce, funcionando las veinticuatro horas del día. Además, incluye el empleo de jueces robot, programados con IA, empleando *blockchain*, que permite reducir la carga de papeles y crear registros transparentes del proceso judicial.

2. El caso de Estonia

Estonia es uno de los países europeos más digitalizados y que ha diseñado todo un modelo de sociedad digital. En lo judicial, ha sido interesante la experiencia piloto de jueces robot dirigidos a la resolución de conflictos de pequeña cuantía (hasta 7.000 euros) y que sólo requieren para su solución la aplicación de normas o de cláusulas contractuales.

En cuanto a su procedimiento, “prevé una tramitación digitalizada y telemática por medio de una plataforma virtual judicial que pretende eliminar la retardación de la resolución de esta modalidad de conflictos (small claims), amén de simplificar la tarea –reduciéndola– de jueces y personal judicial” (Barona Vilar, 2021: p. 395) y en el que solo se les exigiría a los litigantes el subir a la plataforma la documentación e información necesarias. En base a ello, el sistema computacional tomaría la decisión, que luego puede ser apelada ante un juez.

3. El caso del Reino Unido

En el Reino Unido, Otero (2018) explica que, en el año 2015, se creó el “robot abogado”, DONOTPAY, un portal web creado para reclamar multas de estacionamiento en juicios de pequeña cuantía y que, en el 2016, evolucionó hasta convertirse en un *chatbot*, un programa capaz de mantener una comunicación con usuarios, generando recursos y respondiendo a las consultas de sus clientes. Con el tiempo se ha ido extendiendo a causas en las que se demandan cantidades pequeñas en litigios de poca cuantía. Para realizar todas esas operaciones, el robot emplea el sistema de comparación de textos a partir de la concurrencia de palabras claves.

4. El caso de Emiratos Árabes Unidos

En este país se emplea un sistema que se denomina REEM, creado por la empresa española Pal Robotics. Se trata de un robot humanoide que ofrece una suerte de sede electrónica judicial, a la que cualquier ciudadano puede acudir para encontrar asistencia judicial, por ejemplo, para pagar una multa o para denunciar discusiones, accidentes de tráfico, etc. (Silva Barona, 2021)

5. El caso de Canadá

En el marco de funciones de negociación o mediación, en Canadá existe la herramienta SMARTSETTLE ONE creada por iCan Systems Inc. y cuyas virtudes se pueden observar en el hecho de que consigue resolver en una hora un conflicto civil que llevaba más de tres meses tratando de resolverse. Es un software diseñado como una especie de “adjudicador” de controversias, basado en una negociación bilateral electrónica y mediante algoritmos y que “funciona como un procedimiento electrónico, cooperativo, online y que requiere del acuerdo

final de todos los implicados sobre la base de las opciones que va generando la interfaz interactiva” (Martín Díaz, 2020: p. 67).

6. El caso de Australia

En Australia existe el sistema SPLIT-UP, un sistema de expertos basados en reglas y redes neuronales “cuyo dominio es el derecho familiar y subdominio es la partición de bienes” (Martínez Bahena, 2012: p. 839). Su función es prestar un servicio de asesoramiento jurídico sobre la forma en que deben distribuirse los bienes una vez disuelto un vínculo matrimonial, así como de las decisiones que deben tomarse sobre la guarda y custodia de los menores.

7. Situación en Chile

En Chile, en el ámbito de la justicia estatal, pese a que el debate académico se encuentra instalado, los avances y desarrollos de la IA se hallan al debe, encontrándose lejos de alcanzar el nivel de desarrollo de los países expuestos. Esta situación ha quedado en evidencia a propósito de la pandemia mundial provocada por el coronavirus que, en principio, puso en jaque al sistema judicial, impidiendo el normal funcionamiento de los tribunales, demostrando su retraso en lo que al uso de las nuevas tecnologías se refiere y profundizando las brechas de acceso a la justicia.

Con todo, existen varias e importantes iniciativas que pueden considerarse un paso previo a la incorporación de la IA en el ámbito de la justicia, pero que, hasta el momento, solo configuran manifestaciones de procesos de automatización que, a diferencia de los sistemas de IA, no están diseñados para simular el pensamiento humano, y, por consiguiente, no son capaces de realizar tareas de manera “inteligente”, tomar decisiones o realizar acciones sin que hayan sido previa y específicamente programados para ello.

Estas iniciativas no solo encuentran su asidero en años recientes, sino que son el resultado de un proceso de transformación digital que ha experimentado el Poder Judicial hace más de 30 años. Algunos de estas iniciativas son las siguientes: la creación, en la década de los 90, de la Corporación Administrativa del Poder Judicial (CAPJ), que vino a reemplazar a la antigua Oficina de Presupuestos de la Corte Suprema, y, cuya misión, es administrar los recursos humanos, físicos, financieros, en especial, impulsar el desarrollo tecnológico del Poder Judicial.

Luego, en el 2000, se inicia un proceso de modernización, en primer término, con la puesta en marcha de la Reforma Procesal Penal que buscó la incorporación de nuevas herramientas tecnológicas, interconectando a los distintos intervinientes del proceso penal. Luego, lo suyo ocurrió con las reformas que, en materia de familia y en materia laboral, se introdujeron en los años 2004 y 2008, respectivamente, y que, entre muchos otros fines, buscaron la utilización de las capacidades tecnológicas disponibles en beneficio del trabajo jurisdiccional.

También cabe mencionar la creación de la primera intranet y la página web del Poder Judicial, que permitió a los usuarios consultar sus causas; el Sistema Informático de Apoyo a la Gestión Judicial (SIAGJ) que hizo posible la operatividad y conexión de los nuevos tribunales reformados; y la equiparación, a través del Acta 91-2007 de la Corte Suprema, del uso del expediente digital con el de papel.

En los últimos años, un importante capítulo, en este proceso de transformación tecnológica de la justicia, lo constituyó la promulgación de la Ley N°20.886 sobre Tramitación Electrónica, que modifica el Código de Procedimiento Civil y establece la tramitación digital de los procedimientos judiciales. Asimismo, la habilitación de la Oficina Judicial Virtual⁸ y la implementación de la aplicación móvil “Receptores Judiciales”⁹.

Ahora bien, la emergencia sanitaria y el estado de excepción constitucional de catástrofe por calamidad pública obligó, por un lado, a dictar la Ley N°21.226 que estableció un régimen jurídico de excepción para los procesos judiciales, y, por el otro, a la dictación, por parte de la Corte Suprema, de las Actas 41-2020 y 53-2020, que permitieron utilizar todos los medios tecnológicos para mantener el servicio de justicia, entre ellos, el teletrabajo y la realización de las audiencias a través de plataformas virtuales. Además, llevó al desarrollo de programas innovadores como “CONNECTA”, una plataforma que tiene por objetivo entregar a la ciudadanía una “ventanilla” única de atención virtual en donde pueden realizar todo tipo de consultas mediante chat o videollamada.

⁸ Plataforma digital que permite que los usuarios se vinculen con los tribunales de justicia, permitiendo ingresar demandas y escritos, por vía electrónica, desde cualquier lugar y horario.

⁹ Aplicación que permite a los receptores judiciales tramitar y georreferenciar sus actuaciones.

El próximo desafío sería la incorporación de mecanismos de IA en ciertos proyectos ligados a la justicia, por ejemplo, en la presentación de escritos para facilitar y hacer más eficiente el trabajo de los tribunales, lo que brindaría al juez la posibilidad de automatizar tareas para, de esa manera, concentrar más su tiempo en la toma de una mejor y más informada decisión.

Sin embargo, el uso de la IA y los algoritmos en la Justicia todavía resulta ilusorio, aunque sí se debe destacar una manifestación incipiente de cómo la IA comienza a penetrar en el sistema judicial chileno, esto es, el proyecto piloto de IA desarrollado por la Defensoría Nacional junto a la consultora PriceWaterhouse y la Defensoría de Puente Alto. Se trata, como señala la Defensoría Penal Pública (2020), de una aplicación cuya finalidad es entregar información básica a los defensores que se encuentren en audiencias de control de detención, con antecedentes de la causa y orientaciones para exponer en audiencias, de conformidad al delito imputado. Estos insumos básicos incluirían antecedentes de detenciones anteriores, causas vigentes y cerradas, así como un cruce estadístico que indique las posibilidades de que se declare ilegal una detención, o la posibilidad que se decrete una prisión preventiva, tomando como base causas similares, antecedentes del imputado, resoluciones históricas, entre otros antecedentes.

Como ha sido posible observar, la progresiva introducción de las herramientas de IA en la toma de decisiones judiciales suscita serias preocupaciones. En los siguientes apartados profundizaremos en ellas, poniendo énfasis en los desafíos ético-jurídicos¹⁰ y en las herramientas para combatir la discriminación y la opacidad algorítmicas, siempre desde la perspectiva de que el Derecho debe incorporarse a esta propulsión veloz a la que nos están sometiendo la tecnología, pero mantenerse, al mismo tiempo, como garante de que la irrupción de la IA no se convierta en una aniquilación de garantías y derechos, teniendo presente que los sistemas de IA tienen que ser robustos, seguros, estar al servicio de las personas y utilizarse con respeto de los derechos humanos, valores democráticos y la diversidad, con equidad y transparencia.

¹⁰ Por estos motivos, la ética tiene y tendrá cada vez más, un rol decisivo para el desarrollo de sistemas inclusivos y justos, poniendo límite a los sesgos que puedan encontrarse tanto en los datos como en la construcción de los modelos.

IV. ¿Cómo minorizar los riesgos? La ética en el uso de la IA

Reflexionar sobre los problemas y riesgos que suponen la interacción de la IA y la administración de justicia lleva a considerar, como eje central, el análisis de los desafíos ético-jurídicos derivados de su implementación; los daños “accidentales” que se pueden causar a las garantías de los justiciables, principalmente, como consecuencia de los sesgos y discriminaciones que pueden ocasionar el uso de algoritmos entrenados en sistemas de ML y DL; y la necesaria transparencia y rendición de cuentas asociadas a la toma de decisiones automatizada.

Estos dilemas éticos deben ser resueltos a la postre de su correcta aplicación, lo que importa especialmente si se considera el protagonismo que ha ido teniendo esta tecnología, ocupando roles sociales que antes sólo eran colmados por seres humanos y considerando su injerencia en la esfera íntima de las personas. Por ello, lo que se busca es imponer directrices y regulaciones técnico-legales para limitar usos poco éticos, especialmente aquellos vinculados con la no discriminación y la privacidad, así como fortalecer el control y garantías de los individuos (Cotino Hueso, 2017), aunque bien podría discutirse cuál sería la tipología de fuentes apropiada para su regulación (autorregulación, heterorregulación, códigos, normas de conductas, etc.) que, entre otros factores, puede depender de si los algoritmos se incorporan al sector público o privado.

1. ¿Qué han señalado los organismos internacionales?

Amén de la preocupación por los riesgos que puede significar el empleo de la IA y, al mismo tiempo, conscientes de la utilidad que puede generar el impulsar regulaciones internacionales en la materia, organismos han propuesto principios que rijan su funcionamiento, en donde la cooperación supra e internacional para la elaboración de Códigos Éticos de IA se erige como el antídoto que permita amortiguar los sesgos que se extrapolen de la mente humana a la máquina inteligente o bien sea la máquina inteligente que los genere (sesgo del lenguaje).

Entre las diversas propuestas, la Organización para la Cooperación y el Desarrollo Económicos (2019) (OECD) ha señalado cinco principios complementarios para la administración responsable de la IA, centrados en el ser humano: 1) La IA no es un fin en sí mismo, sino un medio que debe servir a la comunidad con el objetivo de aumentar su bienestar; 2) Los sistemas de IA deben diseñarse de manera que respeten el estado de derecho, los derechos

humanos, los valores democráticos y la diversidad, e incorporar mecanismos conducentes a una sociedad justa y equitativa; 3) Debe existir transparencia y divulgación responsable en torno a los sistemas, por ende, los actores de IA deben proporcionar información significativa, adecuada al contexto y coherente; 4) Los sistemas de IA deben evaluar potenciales riesgos de seguridad que pongan en peligro a los usuarios y; 5) La OECD insta a que las organizaciones y las personas que desarrollen, desplieguen o gestionen sistemas de IA, respondan por su correcto funcionamiento en consonancia con los principios precedentes.

Por su parte, la CEPEJ, en diciembre del año 2018, adoptó el primer texto europeo que establece principios éticos relativos al uso de la IA en sistemas judiciales: *“La Carta Ética Europea sobre el uso de la Inteligencia Artificial en los sistemas judiciales y su entorno”*. Su finalidad es “proporcionar un marco de principios que pueden guiar a los responsables políticos, legisladores y profesionales de la justicia ante el rápido desarrollo de la IA en los procesos judiciales nacionales” (Belloso Martín, 2019: p. 11). La CEPEJ (2018) parte de la base de que la IA puede contribuir a mejorar la eficiencia y calidad de la administración de justicia, pero a la vez, debe aplicarse de manera responsable, respetando los derechos fundamentales.

Este instrumento reconoce: el *principio de respeto de los derechos fundamentales*, que implica que el uso de la IA en la justicia no puede poner en entredicho a las garantías derivadas del derecho a la tutela judicial efectiva y las garantías del proceso; el *principio de no discriminación*, esto es, que las decisiones judiciales no contengan sesgos que generen discriminaciones; el *principio de calidad y seguridad*, que llama a no vulnerar la protección de datos personales ni la privacidad; el *principio de transparencia* que pone el énfasis en la accesibilidad y comprensión de las técnicas de procesamiento de datos, así como en la posibilidad de que autoridades y expertos independientes realicen auditorías externas; el *principio «bajo el control del usuario»* que significa que ante el desconocimiento que puede producir en los usuarios, la complejidad de la IA, deben adoptarse medidas de formación, asesoramiento e información, asimismo, articularse recursos que permitan revisar las decisiones judiciales adoptadas en uso de los algoritmos de IA (Cerrillo i Martínez y Velasco Rico, 2019).

Con posterioridad, en febrero de 2020, la CEPEJ ha publicado *“El Libro Blanco de la Inteligencia Artificial, un enfoque europeo orientado a la excelencia y la confianza”* que busca apoyar un enfoque

regulatorio y orientado a la inversión con el doble objetivo de promover la adopción de IA, pero al mismo tiempo, abordar los riesgos asociados con ciertos usos de esta nueva tecnología (Barona Vilar, 2021), tales como la opacidad en la toma de decisiones, las discriminaciones, la intromisión en la privacidad y su empleo con fines delictivos. Al reflejar estos factores en la justicia, la CEPEJ (2020) termina por calificarla como un “sector de elevado riesgo para el uso de la IA”, al ser un ámbito en el que los derechos pueden verse afectados más directamente, haciéndose “preciso arbitrar una acción política y normativa que minimice en todo lo posible estos riesgos, de tal manera que se pueda generar un ecosistema de confianza con un enfoque antropocéntrico” (De Hoyos Sancho, 2021: p. 131).

Para la generación y mantenimiento de ese ecosistema de confianza, uno de los pilares del Libro Blanco de la IA, la CEPEJ (2020) señala que el establecimiento de un marco regulador juega un papel fundamental, debiendo reunir una serie de características:

En lo que a los *datos de entrenamiento* se refiere, deben adoptarse las medidas necesarias que garanticen que los datos con los que se entrena el sistema de IA respetan los valores y normas de la Unión Europea en relación con la seguridad y protección de derechos fundamentales;

En la *conservación de registros y datos*, se deben poder mantener los registros sobre el método de programación de los algoritmos, así como conocer el origen de los datos utilizados. Lo anterior permitirá hacer seguimiento y supervisión de las decisiones adoptadas;

Tratándose del *suministro de información*, los sistemas IA deben facilitar información adecuada y de manera proactiva acerca de cómo deben ser utilizados aquellos que se califican “de alto riesgo”. En particular, tal información se referirá a las capacidades y limitaciones del sistema, a los objetivos a los que se destina, a las condiciones en que se espera funcione, así como al nivel de exactitud previsible en la consecución de sus objetivos;

Todos los sistemas IA deben ser sólidos y exactos para que las haga altamente fiables, sobre todo en las aplicaciones de la IA de “riesgo elevado”. En su desarrollo y configuración se habrán valorado de forma previa y adecuada todos los riesgos conocidos que su uso puede conllevar; además, se habrán adoptado todas las medidas idóneas para reducir al mínimo el peligro de que su utilización, en todas las fases de su ciclo de vida prevista, pueda producir daños;

Finalmente, en la operatividad de la IA, *siempre debe existir la supervisión humana final*, lo que ha de evitar que la IA provoque efectos no deseados o perversos, especialmente cuando opera en sectores de alto riesgo. Los resultados de la IA no serán efectivos hasta que un humano los revise y valide. Además, deberá ser posible realizar un seguimiento del sistema IA mientras funciona, intervenir en tiempo real y, en su caso, proceder a su desactivación.

Los lineamientos de la OECD, la Carta Ética Europea sobre el uso de la IA y el Libro Blanco de la Inteligencia Artificial de la CEPEJ son enfoques que deben ser observados, tanto por las autoridades chilenas, como por las de toda la región, como un modelo mínimo a seguir y como instrumentos necesarios para la incorporación de la IA en la cotidianidad, en especial, en los sistemas judiciales, como una oportunidad de eficiencia, pero al mismo tiempo velando por el respeto de las garantías y de los derechos fundamentales.

2. ¿Qué se ha señalado desde la dogmática?

Más allá de los instrumentos internacionales y las respectivas adaptaciones normativas que, a propósito de estos, se han generado en países europeos, juristas han analizado, desde el prisma ético, las decisiones basadas en algoritmos que están impactando la vida en sociedad. De ahí el surgimiento de subdisciplinas de la ética, que vienen a reflexionar y hacerse cargo de los impactos que el progreso de la tecnología significa para las personas. Al respecto, Reyes Olmedo señala:

“Floridi y Taddeo, han desarrollado una nueva subdisciplina bautizada como *Data Ethics*, entendida como una nueva rama de la ética que estudia y evalúa los problemas morales relacionados con los datos (generación, grabación, almacenaje, procesamiento, difusión y uso de los datos), los algoritmos (IA, agentes artificiales, aprendizaje automático y robots) y prácticas conexas (innovación responsable, programación y diseño de IA, hacking y códigos profesionales)” (2020: p. 7)

Sin embargo, como enfatiza Barona Villar (2021), el interés por estos debates éticos también se ha visto en pensadores, quienes denunciando las consecuencias de la algoritmización de la sociedad, han propuesto principios que permitan a las máquinas inteligentes actuar de forma adecuada en el desarrollo del aprendizaje autónomo, esto es, una suerte de protocolo ético que debiera respetarse desde la intervención humana y de la intervención de la máquina inteligente.

Antes de entrar en los principios y las propuestas de la dogmática, cabe destacar distintas posiciones doctrinales que han confluído para afrontar los avances tecnológicos, desde un punto de vista ético. En una síntesis de las ideas formuladas por Pascual (2017), en primer lugar, se aprecia al *bioconservadurismo* que, defendiendo la idea de mantener la identidad y dignidad humanas, considera que el avance de la tecnología entraña peligros para la humanidad y su supervivencia, por lo que su eventual auge podría hacer perder la noción de la ética. En segundo lugar, el *transhumanismo* pretende aprovechar la tecnología para hacer más próspera la humanidad, construyendo humanos con mayor capacidad. Con todo, su vertiente más dura pretende alcanzar una fusión entre máquina y ser humano y cae en el riesgo de aumentar la desigualdad social. Su posición ética defiende que el valor moral del ser humano está contenido en lo que hace y no en los rasgos propios de su especie. Finalmente, el *humanismo abierto desde los derechos humanos* que adopta lo mejor de las concepciones anteriores, es decir, pone al ser humano en el centro de trascendencia de todo suceso empírico, teniendo como horizonte una vida digna para las personas, pero complementándola con la idea de progreso del transhumanismo, dándole protagonismo pero siempre en el marco de los derechos humanos.

Quedándonos con la última postura, parece preciso entonces, que el desarrollo de la IA sea compatible con un “modelo de derechos humanos”, y para ello se debe impulsar una regulación que incorpore una serie de principios que derivan de la dignidad de las personas en el mundo digital y que se vinculan con la identidad y vulnerabilidad algorítmicas (Corvalán, 2018), una ética centrada en la garantía de derechos fundamentales. Este enfoque permite establecer una primera barrera jurídica al desarrollo tecnológico, funcional al modelo constitucional y democrático, situando la dignidad del ser humano en el centro, configurándose como “una IA que no responda a criterios de mera eficiencia económica, sino que se centre en la función social de la inteligencia artificial y en el empleo de los datos para el bien común” (Martínez, 2019: p. 73).

En esencia, se debe robustecer el sistema de protección incorporando al bloque de juridicidad principios tales como la prevención/precaución, la autodeterminación y transparencia algorítmicas, la imparcialidad del validador, la trazabilidad de la IA, el máximo acceso y la no discriminación (Corvalán, 2018). Advirtamos, en todo caso, que, sin perjuicio de la importancia de establecer estos sólidos presupuestos éticos, jurídicamente el punto de partida

son los principios comunes del Derecho Constitucional, especialmente los derechos y libertades de sociedades democráticas (Cotino Hueso, 2017), aunque puedan requerir readecuación.

De los principios antes señalados interesa destacar, para los fines de este trabajo, el de *transparencia y de no discriminación algorítmicas*. Como destaca Corvalán (2018), de lo primero se sigue que el diseño, desarrollo y uso de la IA debe garantizar que no se configuren cajas negras, siendo la IA transparente en sus decisiones, pudiendo inferir o deducir una explicación entendible acerca de los criterios en que se basa para una conclusión, sugerencia o resultado. Mientras que, para este mismo autor, el segundo principio busca que el diseño y/o implementación de los algoritmos evite procesar la información o los datos bajo sesgos o distinciones, por motivos de etnia, sexo, religión, ideología, nacionalidad, estatus económico o cualquier otra condición social.

Por otra parte, ante el hecho indesmentible de que el avance de la tecnología no espera, autores han propuesto la necesidad de adoptar decisiones estratégicas que se encuentren guiadas por un conjunto de principios, entre ellos, Martínez (2019) señala: asegurar un desarrollo tecnológico que tenga en cuenta los derechos fundamentales como límite y objetivo; garantizar que los procesos de ese desarrollo consideren el ordenamiento jurídico preexistente, siendo capaces de identificar los principios, los valores y las reglas que deban ser aplicadas; asegurar que el abordaje jurídico de la IA logre una aproximación equilibrada a la tecnología, en lugar de enfoques reactivos centrados en la prohibición y la limitación; y ser capaces de entender que dicha aproximación jurídica exige un enfoque coherente con los propios procesos de diseño.

Otros autores han centrado sus esfuerzos en realizar propuestas para adoptar precauciones en la aplicación de los sistemas de IA en el mundo jurídico. En este sentido, Bostrom y Yudkowsky (2014) han propuesto las siguientes: 1) El control e inspección de los sistemas expertos humanos a través de auditorías; 2) Inspecciones periódicas de los algoritmos; 3) Conocimiento de los sistemas que permitan la predictibilidad como garantía de la ciudadanía, en donde la predictibilidad y la previsibilidad se convierten en una suerte de principio de legalidad y se presentarían como una vía necesaria para alcanzar la seguridad jurídica, es decir, que todos los ciudadanos conozcan los criterios mediante los cuales estos sistemas actúan y toman decisiones; 4) Trabajar la incorruptibilidad de los algoritmos, para evitar injerencias externas con fines perversos y sean lo más seguros posibles; 5) Garantizar la responsabilidad en caso de fallos

del sistema, pues si el sistema inteligente comete un error que afecta a las personas y sus derechos, deben existir parámetros para determinar con claridad quién asume dicha responsabilidad: el diseñador, el fabricante, el encargado de su ensamblaje, el comerciante, los usuarios, etc.

Respecto a la responsabilidad frente a fallos del sistema, “se debe tener en cuenta como lo ha establecido la Montreal Declaration for a Responsible Development of Artificial Intelligence (2018), que la inteligencia artificial no debe contribuir a que se busque disminuir la responsabilidad de las personas en la toma de decisiones” (Rincón Cárdenas y Martínez Molano, 2021: p. 16), ya que sólo los seres humanos pueden ser responsables de las decisiones jurídicas formuladas por herramientas de IA y de las acciones originadas de ellas.

Ahora, más allá del nivel de responsabilidad de quienes participen en el proceso, lo relevante es que los sistemas de IA deben diseñarse de manera transparente y de una forma compatible con los valores morales, sociales y culturales que rijan en una sociedad, en un lugar y tiempo dados, con seguridad, sostenibilidad, democracia, participación, transparencia, etc. Al cumplir con estos parámetros se constituiría en una verdadera IA responsable, entendiendo por ella, una tecnología de confianza basada en valores éticos y sociales y con un enfoque antropocéntrico.

Como se ha observado en los instrumentos internacionales y en las pautas doctrinarias, el respeto a los derechos fundamentales y del principio de transparencia es clave para construir una IA responsable y basada en la confianza. De ahí que parezca al menos discutible, desde un punto de vista ético, la posibilidad de implementar sistemas no deterministas que vengán a sustituir el rol del juez, con completa autonomía, precisamente, por la desconfianza que produce la baja interpretabilidad de los resultados obtenidos por estos, al perder los diseñadores la capacidad de control sobre el mismo (Borges Blázquez, 2020), en cuyo caso, las personas sobre las pesan las resoluciones, se colocan en una inevitable situación de indefensión, al enfrentarse a un sistema en el que no se puede comprobar la veracidad de lo que ha llevado a la toma de decisiones.

Por esa baja interpretabilidad, hay quienes califican a estos sistemas de IA como “oráculos artificiales” y entregan el siguiente ejemplo: dos personas discuten acerca de si una de ellas le debe o no una suma de dinero y a propósito de aquello acuden a la justicia. Ahora, el juez en su sentencia resuelve que no se le debe nada, pero no les dice cómo los evalúo. En este caso, se

estaría frente a un típico caso de sentencia arbitraria (Corvalán, 2018). Si a lo anterior se le añade que las decisiones adoptadas pueden verse afectadas por sesgos algorítmicos, y se extrapola el ejemplo al sistema de justicia criminal en donde se decide acerca de la responsabilidad penal de un sujeto y de si es procedente o no aplicar la pena más gravosa del ordenamiento jurídico, ahí se estaría frente a uno de los debates éticos más complejos e importantes a los que se va a enfrentar la justicia penal en las próximas décadas (Borges Blázquez, 2020).

Por tanto, el reto ético-jurídico consiste, por un lado, en poder utilizar los algoritmos, el aprendizaje automático y la IA evitando la discriminación y sesgos, y por el otro, en la necesidad de superar la opacidad algorítmica (Reyes Olmedo, 2020). Mientras ello no ocurra, no se debería autorizar el empleo de herramientas que pudieran ser determinantes del sentido de una resolución judicial, si las partes no pueden conocer los elementos que integrarán los algoritmos y el propio funcionamiento del sistema de IA (De Hoyos Sancho, 2021).

Finalmente, aunque en la actualidad, desde un punto de vista ético, los usos de esa IA fuerte estén más bien limitados, y según la CEPEJ (2018), el estado de desarrollo de las técnicas de aprendizaje automático no permite hoy alcanzar resultados confiables con respecto a la predicción de decisiones judiciales, si su avance llegara a tal punto de implementar la justicia predictiva con total prescindencia del juez humano, por razones de eficacia, eficiencia y economía procesal, esos peligros son los que se pretenden paliar con el mantenimiento de una buena estructura de principios y garantías. Barona Vilar (2021), reflexiona sobre este punto y señala que habría que pensar que, del mismo modo que los jueces ingresan en la carrera judicial por determinados sistemas y ascienden a través de medios legalmente establecidos, la operatividad de máquinas que pudieren llegar a sustituir al juez (juez-robot totalmente), deberían basarse en unas normas-protocolos que garantizaran el tipo de máquina que se va a utilizar, quién la diseñó, cuál es su contenido y quién puede llegar a controlarla. Obviamente, que existan mecanismos de control a posteriori, un sistema de fiscalización/validación del funcionamiento del sistema computacional judicial, o si se quiere, una auditoría del funcionamiento.

Desde luego que la lista de criterios, principios y medidas enunciados en este apartado no es de ningún modo exhaustiva, pero puede ser útil como una pequeña muestra de lo que debería estar pensando una sociedad cada vez más informatizada. Así, cada vez que se confíe a un

algoritmo la toma de una decisión, será necesario establecer de antemano las reglas y criterios que le serán aplicables.

Sin embargo, no basta únicamente una declaración de criterios, principios y medidas, sino que, al mismo tiempo, se debe formar ética y jurídicamente a los diseñadores, lo que significa que el equipo debe contar con una base de formación que incorpore los valores éticos humanistas y de garantía a los derechos fundamentales. Desde luego que “no se trata de convertir a los desarrolladores, ni a los gestores en expertos juristas. Sencillamente obliga a que en el conjunto de la organización exista una cultura de garantía y respeto del derecho, que oriente las decisiones” (Martínez, 2019: p. 75). Ello comporta un proyecto colaborativo y multidisciplinario entre expertos informáticos y expertos del derecho.

V. Algunas reflexiones sobre la implementación de la IA en el proceso chileno

Muchas veces se suele hablar de un desarrollo asimétrico que es inherente a los países menos aventajados, lo que en el ámbito de la implementación y desarrollo de la IA se ha manifestado con mayor fuerza, dando cuenta de un fenómeno de diversas asimetrías. Así, hasta el momento, las tecnologías de IA han sido desarrolladas y desplegadas más rápidamente en el norte global y la preparación de los Estados latinoamericanos para adoptarlas ha sido prácticamente mínima.

Esto sucede, pues, a medida que los sistemas de IA se integran progresivamente en la cultura cotidiana, en las más diversas dimensiones sociales, su desarrollo impacta las diferentes brechas digitales. En este sentido, se ha señalado que “dada la dependencia de la IA en Big Data (datos masivos) y el poder computacional, es muy probable que la difusión de estos sistemas amplifique desigualdades existentes, incrementando las disparidades en el acceso a la tecnología” (Lombana, 2018: pp. 21-22). Esta perspectiva de que la IA crea una nueva brecha digital supone que el acceso a las tecnologías de IA ofrece una ventaja a países que puedan desarrollarlas, implementarlas y apropiárselas. Los que “tienen” IA pueden optimizar sus operaciones, automatizar procesos e innovar, mientras que, los que no, se quedan atrás en la evolución digital.

Ahora, como se ya se advirtió, la situación de Chile no ha sido muy auspiciosa en lo que a los avances y desarrollos de la IA respecta, y sólo hace unos años atrás ha comenzado a gestarse una “Política Nacional de Inteligencia Artificial”, más allá de ciertos proyectos que en el sector privado han sido implementados.

En la administración de justicia no ha sido diferente, y sólo es posible observar manifestaciones incipientes, por ejemplo, la aplicación ya comentada, que busca apoyar la tarea de los defensores públicos en distintas instancias del proceso penal, así como también algunos proyectos de innovación del Poder Judicial, en un nivel muy básico. Dentro de estos últimos, destacan el sistema para la tramitación masiva de juicios ejecutivos y para la tramitación masiva de recursos de protección contra ISAPRES que ha sido implementado en las Cortes de Apelaciones, consistiendo en un mecanismo que hace un reconocimiento a partir de los datos ingresados por las partes, entregando una sentencia judicial a partir de los datos que se extraen de la información entregada por los usuarios; el sistema de reconocimiento de texto con el que

trabaja las Cortes de Apelaciones que se utiliza a propósito de órdenes de no innovar; finalmente, el empleo de herramientas de IA en la gestión judicial en el Juzgado de Familia de Coronel. Aunque cabe destacar que estas manifestaciones no pasan de ser programas piloto, y no una política que trascienda el Poder Judicial y la administración de justicia.

Sin embargo, la contingencia que ha transitado Chile desde octubre del año 2019, a propósito de la crisis político-social y, a partir del año 2020, con ocasión de la pandemia mundial producida por el COVID-19, dio cuenta de la importancia estratégica de la IA para los gobiernos de todo el mundo, lo que ha obligado al Poder Judicial y a la CAPJ, institución encargada de adquirir los recursos tecnológicos y los sistemas inteligentes para la justicia, a generar un proceso de transformación digital.

A propósito de esa importancia estratégica, el legislador chileno ha tomado cautelas en el avance de las nuevas tecnologías. Así, el 25 de octubre de 2021 se dictó la reforma constitucional de la Ley N°21.383, que modifica el artículo 19 N°1 de la actual Constitución Política de la República, y que busca garantizar que el desarrollo tecnológico y científico se encuentre al servicio de las personas y que debe llevarse a cabo respetando el derecho a la vida y la integridad física y psíquica.

Ahora, otros países de América Latina han ido más allá, impulsando algunas iniciativas para introducir sistemas digitales en la administración documental o de decisión autónoma en sus procesos judiciales, con el objeto de solucionar “cuellos de botellas” que, en ocasiones, los caracterizan. Esto problemas afectan directamente el derecho de acceso a la administración de justicia y redundan en la congestión judicial, una de las principales problemáticas que puede afectar a la justicia, haciéndola más lenta, menos accesible, más costosa y generando un impacto negativo en la sociedad (Rincón Cárdenas y Martínez Molano, 2021). Un fenómeno similar tiene lugar en la prestación de justicia en Chile, en que diversos motivos, como la falta de acceso a atención profesional eficiente, los altos costos, el formalismo judicial en cuanto al lenguaje, formas y rituales, han generado un déficit en cuanto a acceso a la justicia para el ciudadano común (Lillo, 2020).

Esto nos lleva a reflexionar acerca de la posible implementación de la IA y algoritmos predictivos en el proceso chileno, compatibilizando las garantías y principios que lo gobiernan con la celeridad y oportunidad en la decisión. Con todo, por la extensión de este trabajo, las reflexiones tendrán como objeto de análisis el proceso penal y civil, analizando su conveniencia –o no– y las formas de implementar la IA en esos ámbitos de la justicia.

1. Algoritmos en el proceso penal

En lo referido al proceso penal, una primera cuestión que se debe tener en cuenta es la reflexión que realiza Beltrán Román y Preminger Samet, quienes señalan que:

“Independientemente del ordenamiento jurídico que se trate, la combinación entre el uso de la tecnología y la intervención humana, en el marco de la justicia penal, puede presentar problemas. Tanto durante el proceso de implementación, como de utilización de la inteligencia artificial, garantías y derechos de los ciudadanos en el marco de un proceso penal y principios constitucionales íntimamente vinculados a estos podrían verse menoscabos si no es aplicada de forma correcta” (2020: p. 10)

En el mismo sentido, Nieva Fenoll (2018) ha manifestado que su automatización es más compleja y peligrosa pues, al fin y al cabo, un delito no es algo tan frecuente como un conflicto civil y suele estar lleno de circunstancias personales, del reo y de la víctima, que deben considerarse con cuidado. Es por ello que, en su opinión, quizá lo único ciertamente automatizable serían las conformidades en los delitos menos graves que puedan juzgarse solamente con las informaciones policiales. No obstante, para este autor, tampoco puede parecer muy conveniente, pues los bienes jurídicos exigen por parte de los jueces una comprobación de la actividad policial en cuanto a su corrección, coherencia y ausencia de lesión de derechos fundamentales y puede resultar complejo exigirle esa fineza a una máquina.

Teniendo en cuenta las precauciones que anteceden, se podría plantear una posible incorporación de la IA, empero con las más importantes reservas, a efectos de evitar esas situaciones de menoscabo e injerencias en derechos, y en todo caso, solo cuando se garantice efectivamente un proceso justo.

Pillado González (2021) señala que es evidente la utilidad que, dentro del proceso penal, podrían tener ciertas herramientas de IA como pueden ser los buscadores inteligentes de jurisprudencia, los que facilitan la redacción de escritos, los que ordenan los argumentos a favor y en contra de una posición, o aquellos que permiten automatizar actuaciones procedimentales, en cambio, cuestiona y critica el empleo de herramientas de IA programadas para valorar la probabilidad de una conducta futura, que facilitan la toma de decisiones sobre la base de una predictibilidad de la actuación del sujeto pasivo del proceso y que pueden servir al juez a la hora de decidir sobre la determinación de la pena o su extensión, la suspensión de la pena privativa de libertad impuesta o la aplicación del tratamiento penitenciario al condenado.

Este cuestionamiento se realiza, por un lado, porque si el significado de una sentencia penal fuera reducido a la decisión judicial condenatoria o absolutoria sobre la responsabilidad de una persona por la comisión de un delito, probablemente se resolvería que la asistencia o incluso el reemplazo del humano por un procesador tecnológico sería de suma utilidad. Sin embargo, como advierte Vivar Vera (2021) una sentencia penal es más que eso, tiene implicaciones de forma y de fondo, tiene significado simbólico y trasfondo teórico jurídico, político, social y económico. Es una construcción compleja que decide el futuro de las personas, la decisión última del proceso penal y ultimadora de un evento doloroso: decisión violenta para un acto violento.

Y, por otro lado, porque ya conocidos son los problemas que envuelve el uso de algoritmos: la opacidad y falta de transparencia. Asuntos graves, ya que las posibilidades de alegar, de probar o de impugnar, y sus resultados, dependen de que se puedan conocer con suficiente detalle los datos y razonamientos conforme a los cuales resolverá el órgano decisor, y/o sobre los cuales basa su actuación la contraparte o parte acusadora. Si esto no ocurre, se vulnera el derecho de defensa del imputado y no tanto porque no pueda ejercerlo, pues la contradicción estará garantizada formalmente, sino que el problema está en comprender cómo el algoritmo que se ha usado ha conseguido un determinado resultado y poder así rebatirlo (De Miguel Beriain y Pérez Estrada, 2019). Si se quiere hacer uso de algoritmos en el sistema de justicia penal, al menos, debemos preguntarnos qué nivel de transparencia tiene el resultado del algoritmo, esto es, cómo ha manejado los datos y qué valor se le ha entregado a cada variable (Borges Blázquez, 2020).

El otro problema es el de los sesgos que, además de afectar el principio de igualdad, si el algoritmo los presenta, el resultado de su análisis afectará a la imparcialidad del juzgador al momento de dictar la sentencia porque su convicción judicial estará condicionada al sesgo que presente la herramienta de IA. Sobre el particular, Vivar Vera (2021) recoge la opinión de autores que creen que la simple sospecha de un sesgo discriminatorio crea un clima de desconfianza de la justicia y de la decisión, por ende, para revertir esa apreciación, el estándar de duda razonable conllevaría a la absolución, por aplicación del principio de presunción de inocencia y la falta de convicción de la existencia del hecho delictivo y de la responsabilidad penal del acusado.

En este punto, conviene recordar el sistema COMPAS que, como se señaló, permitió sentar las bases para una eventual uso de IA en materia penal, permitiéndola para: 1) imponer medidas alternativas, 2) valorar si un delincuente puede ser supervisado de manera efectiva y segura por la comunidad y 3) determinar las condiciones de la suspensión de la pena y libertad condicional. Si esto se tiene en cuenta, para una futura implementación de la IA en el proceso penal chileno, podrían pensarse, en son de la política legislativa, su procedencia en la investigación penal, en las salidas alternativas, en los delitos de acción penal privada, en los delitos que, por su gravedad, tengan aparejadas penas más bajas, como ocurre con las faltas, etc. En efecto, debido a la complejidad normativa y ética que supone el uso de estas tecnologías en juicios criminales, parámetros como el de la corte de Wisconsin sirven para determinar *de lege ferenda* una aplicación adecuada de tecnologías algorítmicas al proceso penal chileno.

Así las cosas, estos sistemas claramente quedarían descartados a la hora de determinar la prisión preventiva de un imputado en cuanto a medidas cautelares se refiere, así como tampoco sería plausible una sentencia definitiva que concluya con la pena privativa de libertad. Ambas afirmaciones se deben a que, ante los peligros ya comentados de las IA, que pueden suponer afectaciones a derechos de las personas, se busca evitar que irradie a la esfera más agresiva del poder coactivo del aparato estatal: la prisión.

Aplicando la sentencia de *State v. Loomis* a un eventual escenario legislativo chileno, el primer punto sobre lo que la IA podría hacer tener cabida en el ordenamiento interno es a propósito de las salidas alternativas, institución que permite encontrar soluciones integrales respecto de delitos de baja gravedad o respecto de los cuáles es posible trazar acuerdos con el Ministerio Público

y/o la víctima de la causa. Así, un sistema de evaluación de riesgos en Chile estaría apto para participar en la decisión de una suspensión condicional de procedimiento o un acuerdo reparatorio una vez verificadas las respectivas hipótesis normativas.

El segundo punto, por su parte, resulta más complejo en el escenario nacional pues parece exceder lo orgánico-procesal y se tendría que recurrir a otros ámbitos como la reinserción o los efectos de una salida alternativa, asunto complejo teniendo en cuenta que lo que prima en el derecho penal chileno es la prevención negativa (neutralización del agente) por sobre la positiva (educación). El tercer punto, sin embargo, resulta mucho más claro, y es factible que una máquina pueda participar en la toma de decisiones favorables a los derechos civiles, como la libertad condicional, siempre contando con los fundamentos adecuados al caso concreto; ciertamente, hay mucho que la tecnología puede hacer a propósito de la post-sentencia.

Descartando la presencia de sistemas algorítmicos en consecuencias jurídicas de alta intensidad, como bien se ha señalado, estos mecanismos sí podrían tener cabida a propósito de ilícitos penales de baja intensidad como las faltas, institución para la cual el ordenamiento contempla penas más benignas, distintas a la privación de libertad. Así también se podría argumentar la plausibilidad de una mayor presencia tecnológica respecto de los delitos de acción privada como lo son la injuria y la calumnia, al tratarse de conflictos que más bien escapan al interés público, y en dónde la voluntad de los contendientes tiene un mayor rol.

Cabe concluir, en todo caso, que, si nos ceñimos a los lineamientos jurisprudenciales del Tribunal Supremo de Wisconsin, en ningún caso se deberá tener la IA como única juzgadora a la hora de determinar si el imputado elude la pena privativa de libertad.

En suma, la participación de la tecnología en el proceso penal merece analizarse, criticarse y construirse, teniendo en cuenta que la decisión judicial de condena o absolución es resultado de un proceso racional del contexto fáctico, legal y probatorio que impacta en la vida de los justiciables. No es si no por ello que el uso de algoritmos en materia penal ha sido calificado por la CEPEJ (2020) como “usos a tener en cuenta con las reservas más extremas” porque las experiencias que existen en el mundo sobre la materia –COMPAS en Estados Unidos y HART en Reino Unido–, debido a la metodología que han utilizado, han llevado a resultados

incorrectos, por ejemplo, el hallazgo de que algunas personas afroamericanas están más involucradas en actos criminales ha llevado a un factor de riesgo más alto para toda la población afroamericana, aumentando injustamente la cuantía de sus sentencias.

Si reproducimos estos problemas en la justicia penal de nuestro país, redundarían en la quiebra de ciertas garantías que integran el derecho al proceso debido, muy especialmente, el principio de igualdad procesal, el derecho de defensa y la presunción de inocencia, que constituyen la base del proceso penal garantista chileno. Por tanto, el enfoque que adopte la IA tiene que ser uno respetuoso de esas garantías y debe ofrecer al individuo la posibilidad de rehabilitación y de reintegración.

En este sentido, se aboga por el gran potencial de la IA para dar una mejor respuesta del sistema procesal penal ante la criminalidad, pero no ya con el objeto de sustituir a los jueces u otros operadores jurídicos en el desempeño de sus funciones, sino utilizando esa tecnología como herramienta de apoyo o ayuda en cuanto supone la aportación de toda una información sobre el investigado o condenado que va a permitir ajustar mejor la respuesta judicial, ya sea en la imposición de una medida cautelar o a la hora de determinar la pena a imponer o en la toma de decisión sobre la suspensión de la ejecución de la pena privativa de libertad impuesta en la sentencia o sobre el tratamiento penitenciario más adecuado al reo para su rehabilitación. Se piensa entonces, en el empleo de algoritmos predictivos como asistentes de los jueces humanos, y no como se presenta en otras disciplinas jurídicas, como decisores autónomos.

2. Algoritmos en el proceso civil.

Para considerar una posible implementación de los algoritmos en el proceso civil resulta preciso partir de la base de que una gran parte de la labor que desempeñan los jueces y tribunales es mecánica. Esto ha traído como consecuencia que buena parte de los funcionarios judiciales invierten su tiempo utilizando modelos de resoluciones y modificando los datos identificatorios del proceso, en una especie de “copia y pega”, lo que redundaría en que pocas resoluciones y sentencias se redactan completamente *ex novo*, aunque claro está, esto depende del caso concreto y de la tendencia de cada juez (Nieva Fenoll, 2018). Desde luego que no debiera existir mayor

inconveniente en aquello si la motivación “copiada” es adecuada para el caso concreto, pero al margen de ello, esa realidad debiera reconducir a una reflexión.

Esa reflexión viene propiciada por la utilidad que la IA le puede entregar a esa labor mecánica porque no solo es capaz de obtener una mayor producción y copia de documentos, sino que también consigue una mayor capacidad de análisis. Combinadas esas funcionalidades resulta que una aplicación de IA podría ser increíblemente más rápida que un juez en la resolución de procedimientos previsibles, sobre todo en el análisis de la documentación, que en ocasiones un juez o el personal del juzgado solo puede revisar de manera superficial. Por este motivo, la irrupción de la tecnología, y más concretamente, de las diversas variables de la IA, puede llevar, como señala Barona Vilar (2021), a mejorar el desarrollo del procedimiento, dado que cabría trabajar con un aporte mayor de documentos creativos que se utilizan en estos actos procesales, además de permitir un análisis más profundo de los mismos.

A este respecto, la CEPEJ (2018) ha señalado que el procesamiento de decisiones judiciales por algoritmos de IA, según sus desarrolladores, es probable, en materia civil (al igual que en materia comercial y administrativa), para ayudar a mejorar la previsibilidad de la aplicación de la ley y la coherencia de las decisiones judiciales. Claro está, ello sujeto al cumplimiento de ciertos principios, entre ellos, el derecho de acceso a un tribunal, principio adversarial, igualdad de las partes en el proceso, imparcialidad e independencia de los jueces y derecho a un abogado. Al mismo tiempo, la previsibilidad de un juicio civil puede producir un efecto virtuoso respecto de la demanda de justicia, pudiendo reducir gradualmente los casos no fundados (Battelli, 2020).

De esta manera, un criterio que advierte Barona Vilar (2021) y que se ha venido argumentando, en relación con la posibilidad de la tramitación automatizada, es precisamente que allá donde el modelo procedimental sea reiterativo y no concurra oposición, o aun habiéndola ésta sea restringida, y se precise soporte documental, se facilita notablemente este modelo de justicia algorítmica. Entrega como ejemplo, en el Derecho Civil español, los procesos de desahucio u otros posesorios en los que el ilegítimo poseedor simplemente carece de motivo de oposición, o bien el proceso monitorio sin oposición. En general, en aquellos procedimientos muy sencillos en los que la intervención humana simplemente no es necesaria.

Ahora, si se piensa en las materias conocidas por los órganos jurisdiccionales chilenos, podríamos señalar aquellas respecto de las cuales se ha recomendado su desjudicialización. Actualmente, es posible apreciar que gran parte de los ingresos de causas civiles corresponden a procesos ejecutivos, gestiones preparatorias para la vía ejecutiva y de cobranza que tienen como patrón común el estar sometidas a ciertos trámites no decisionales, más o menos idénticos (Lillo, 2020).

Sobre estos procesos, existe un consenso respecto de la necesidad de dotar de mayor eficiencia la ejecución de los créditos por las múltiples deficiencias que el procedimiento ejecutivo presenta para la concreción de las decisiones judiciales dentro de un plazo razonable: la lentitud, la ritualidad, el exceso de excepciones que puede oponer el ejecutado y la saturación de los tribunales civiles, hacen que su desjudicialización sea altamente factible (Vargas Pávez, 2013). Este tipo de asuntos, al contener trámites en los cuales el juez realiza funciones de tipo administrativo o notarial, pueden realizarse digital y automáticamente, disminuyéndose costos y tiempo de tramitación, en el entendido que “la duración y dilación innecesarias llegarían al punto de constituir una verdadera negación de justicia” (Lillo, 2020: p. 121).

Otro grupo de materias corresponde a procedimientos esencialmente no contenciosos, en los que existe ausencia de controversias, por ejemplo, la apertura, protocolización y publicación del testamento, la posesión efectiva, la declaración de herencia yacente, el divorcio de común acuerdo, en que, por no existir intereses contrapuestos entre las partes para su realización, basta con el sólo cumplimiento de ciertos requisitos al modo de *checklist*.

Asimismo, es posible mencionar ciertos asuntos de naturaleza contenciosa que se caracterizan por ser de baja cuantía, o por la necesidad de requerir de soluciones rápidas, o por ser de una baja complejidad y cuyos procedimientos son relativamente sencillos. Así, las partes pueden resolver disputas respecto a deudas, daños a la propiedad, servicios no otorgados, o disputas de dominio. Las denominadas “pequeñas causas” que se enfocan en conflictos en las comunidades, cuestiones relacionadas a ruidos molestos, tenencia responsable de animales y/o mascotas, infracciones a la ley de urbanismo, infracciones a la ley de copropiedad inmobiliaria, entre otros, cuyas resoluciones están entregadas a los Juzgados de Policía Local.

Estos últimos asuntos podrían ser conocidos y resueltos mediante los mecanismos de resolución de conflictos online denominados *Online Dispute Resolution* (ODR), estructuras que permiten resolver las controversias a través de medios electrónicos, si no en todas las fases del procedimiento, si en gran parte del mismo. En derecho foráneo, la experiencia del *Civil Resolution Tribunal* (CRT) ha demostrado su eficiencia, configurándose como el primer tribunal en línea en Canadá y el primero de su tipo en el mundo. El CRT es competente para resolver reclamos menores y disputas sobre la propiedad, centrándose en la resolución temprana de disputas en línea y permitir que las personas se conviertan en participantes activos en la resolución de conflictos, empleado principios de diseño ágiles y centrados en el usuario para promover un cambio transformador en el sistema de justicia civil (Salter y Thompson, 2017).

En nuestro país un ODR puede liberar recursos judiciales y de los tribunales para tratar asuntos complejos y graves, puede agilizar los procesos judiciales y ampliar los métodos de acceso, reduciendo la necesidad de una amplia infraestructura física de los tribunales. Lo importantes es que cuando los litigantes acceden a una plataforma de resolución de disputas en línea, deben ser informados claramente si el procesamiento de su disputa se realiza de manera totalmente automatizada o con la participación de un mediador o un árbitro (CEPEJ, 2018).

Además, permitiría tramitar aquellas disputas que comprenden un amplio abanico de cuestiones cotidianas que hoy son conocidos por los Juzgados de Policía Local. Piénsese, por ejemplo, las infracciones a la Ley N°19.496 sobre protección a los Derechos de los consumidores o infracciones a la Ley N°18.700 sobre votaciones populares y escrutinios. En todo caso, ya existen algunas manifestaciones en esta línea: el ‘Sistema de Resolución de Controversias’ de NIC Chile para resolver conflictos por inscripción de nombres de dominio, a través de un arbitraje en línea, tramitadas mediante un sitio web y bajo un procedimiento estandarizado y el proyecto digital ‘Resolución en Línea’, un ODR del Centro de Arbitraje y Mediación de la Cámara de Comercio de Santiago que permite la resolución online de conflictos, con apoyo algorítmico, entre consumidores y empresas cuando no han encontrado una solución en sus canales de atención, destacando su utilidad en el *e-commerce*.

Se pueden observar algunas de las bondades que ofrecería la implementación de la IA tanto en tareas simples y repetitivas, como en aquellas materias de baja cuantía y en las que se ha

recomendado su desjudicialización. En este contexto, la IA resulta compatible y hasta necesaria para una Reforma Procesal Civil “2.0” donde los tribunales y Cortes dicten más y mejores fallos, lo que provocaría sustanciales externalidades positivas: aliviar la carga de jueces, ministros y relatores, dictar fallos más expeditos, fortaleciendo el acceso a la justicia, incrementar la calidad de las sentencias, incluyendo jurisprudencia y doctrina relevante, disminuir las contradicciones que se generan entre las salas de una Corte y a veces en una misma sala. Es en estos sentidos que la IA no hará la administración de justicia más artificial, sino que la hará todavía más humana.

3. Al menos Prometea

Es preciso destacar que las reflexiones que se han realizado en este apartado se hacen sobre la base de que la transformación digital es muy desigual en los sistemas judiciales de los distintos países, dado que mientras algunos han desarrollado un uso de esta tecnología y soportes muy avanzados, otros, como el sistema judicial chileno, apenas lo presentan como emergente. Asumiendo esta idea, es que puede resultar más apropiada la introducción de los algoritmos y sistemas de IA a la administración de justicia de una forma que resulte progresiva en el tiempo, teniendo como primera meta, el incorporar la IA principalmente a la gestión judicial.

Por lo tanto, ya no hablar de una IA que venga a reemplazar al juez y su razonamiento para fallar, sino que de una IA que se sustente sobre la base de los siguientes dos conceptos: el concepto de transformación digital y el concepto de asistencia. Es a propósito de estas nociones que interesa destacar las virtudes que Prometea, la primera IA al servicio de la justicia de Argentina, y la primera en Latinoamérica, puede ofrecer a la justicia de nuestro país.

Prometea es uno de los proyectos más desarrollado en el ámbito de la IA aplicada a la administración de justicia, más allá de Estados Unidos y de los países europeos. Pertenece a la denominada “inteligencia en la interfaz” que implica que el usuario simplemente interactúa y la tecnología resuelve los problemas mediante conexiones con diferentes sistemas que pueden responder a las necesidades del usuario y a partir del aprendizaje (Corvalán, 2018). Se trata de una herramienta implementada por el Ministerio Público de Buenos Aires, consistente en un asistente de búsqueda de documentos legales, que se puede activar por voz, y que permite reconstruir expedientes judiciales de forma sencilla (Cerrillo i Martínez y Velasco Rico, 2019).

Belloso Martín (2019) señala que el procedimiento es el siguiente: el usuario abre el expediente, y una vez que está en condiciones de proyectar un modelo, activa a Prometea por comando de voz o mediante un chat. Se carga el número de expediente y el sistema busca la carátula en la página del Tribunal Superior de Justicia. Después, lee las sentencias de primera y segunda instancia, realiza algunas preguntas al operador para completar datos y emite el dictamen. Su tarea se realiza sobre la base expedientes donde la Justicia ya tiene algo dicho, permitiendo que el operador jurídico pueda destinar su tiempo a los casos complejos.

Todo el proceso, de manera íntegra, se realiza a través de la IA. Desde un “Hola”, pasando por varias preguntas y respuestas entre Prometea y el usuario, lo que incluye que busque y “traiga” al dictamen leyes y decretos, hasta llegar al punto en el que Prometea diga “dictamen completado”. Luego, se le puede ordenar que “imprima”, “descargue” o que “envíe”, por mail o a una red interna, el proyecto borrador para que sea corregido. Así puede entregar, por ejemplo, un modelo citación, un modelo vivienda, un modelo autosuficiencia, etc. También puede notificar si el modelo de dictamen no es aplicable, ya sea porque están vencidos los plazos, o bien, por la falta de algún requisito formal (Corvalán, 2018).

Como se puede observar, lo que hace Prometea es proponer textos de resoluciones judiciales en base a la gestión documental. Identifica procesos similares y en base a una serie de requerimientos, va construyendo y proponiendo modelos que luego son revisados por un funcionario judicial. Por este motivo, es que no pretende reemplazar a los funcionarios judiciales y abogados, pues detrás de ella siempre debe haber un ser humano que, con su inteligencia natural, y no artificial, defina si la propuesta del sistema es adecuada o no.

Desde el año 2019, a partir de un trabajo colaborativo entre la Corte Constitucional de Colombia, el Laboratorio de Innovación e Inteligencia Artificial de la Universidad de Buenos Aires y la Universidad del Rosario, Prometea también comenzó a ser utilizado en Colombia, con el propósito de ayudar en el proceso de revisión de fallos de tutela que llegan a la Corte, aunque solamente aquellas que traten sobre la vulneración del derecho a la salud, siendo su principal objetivo “la selección y preselección de las tutelas que serán conocidas por la Corte Constitucional, mediante criterios generados por el propio juez” (Rincón Cárdenas y Martínez Molano, 2021: p. 3).

Este sistema ha comenzado a extenderse, pues sus ventajas resultan evidentes. Así, se ha estimado que este tipo de inteligencia puede impactar de manera decisiva para garantizar ciertos derechos de acceso e implica un salto cualitativo en relación con la velocidad y la precisión para prestar un mejor servicio de justicia (Corvalán, 2018). Una evidencia práctica de su capacidad y eficiencia se encuentra en el estudio publicado por el Banco Interamericano de Desarrollo, en el 2018, en el que consta que los humanos tardan aproximadamente 172 días en resolver 1.000 casos judiciales de baja complejidad; empero, Prometea puede dar respuesta a esos 1.000 casos en 42 días sin que sea necesaria la revisión gramatical y ortográfica posterior, teniendo importantes porcentajes de eficiencia que sobrepasan en la mayoría de los casos el 300% (Rincón Cárdenas y Martínez Molano, 2021). Prometea es capaz de adoptar resoluciones en 15 segundos con un 98% de acierto, pudiendo realizar análisis estadístico y generar certificaciones con la *tecnología blockchain* (Simón Castellano, 2021). Otro de los aspectos más importante se vincula con su extensión a otras áreas, por ejemplo, simplificar la interacción con un fiscal, organizar procesos judiciales internos, optimizar las relaciones ciudadanía-Estado y, sobre todo, enfocar su uso en los sectores vulnerables y en las personas con discapacidad (Corvalán, 2018).

Finalmente, ante la interrogante respecto de qué es lo que puede hacer el Poder Judicial chileno para incluir a la IA y continuar con el proceso de transformación digital, se puede concebir su incorporación en los siguientes ámbitos: por ejemplo, en la atención a público, es decir, los procesos de consulta a los usuarios podrían ser automatizados a través de un proceso y de un asistente que lidere los recursos de los tribunales; en las etapas de conciliación de los diversos procedimientos; asimismo, ampliando el ámbito de las sumas y presumas que encabezan los escritos judiciales se lograría obtener los datos suficientes para lograr la estandarización de ciertos procedimientos, como en los recursos deducidos contra ISAPRES o en el proceso ejecutivo; en la gestión documental a través de un sistema que lea y procese documentos; en el apoyo en la dictación de resoluciones y sentencias definitivas, etc. Es en este sentido que un sistema como Prometea, al abrir una interconexión colaborativa entre humanos y máquinas, se erige como una herramienta útil a ser considerada por el Poder Judicial y por la CAPJ, una oportunidad para regular el flujo de las causas y reducir costos operativos judiciales.

Conclusiones

El objetivo de este trabajo ha sido mostrar, desde una perspectiva doctrinaria y comparada, cómo la IA y los algoritmos predictivos han impactado en el ámbito del Derecho, en especial, en los sistemas de justicia de diversos países, con miras a una futura, pero inevitable, incorporación de esta tecnología al sistema judicial chileno.

Es una realidad indesmentible que la tecnología avanza a gran velocidad y sus posibles usos, en los sistemas de justicia civil y criminal, como en las demás áreas del derecho, son prácticamente infinitos. Chile, por su parte, debiese estar preparado para estos cambios disruptivos y abrir el debate a fin de establecer principios y lineamientos para una eventual aplicación de estas tecnologías, de modo de hacer más efectiva y eficiente el acceso a la justicia, pero al mismo tiempo velar por el pleno respeto por los derechos y garantías fundamentales de los ciudadanos, en especial, de aquellos que pertenecen a grupos en situación de discriminación.

Los sistemas de IA, en la medida de que procesan datos, pueden cometer discriminaciones y las decisiones que se adopten a partir de estos sistemas, se pueden ver teñidas por los mismos sesgos y prejuicios que afectan a las personas que los construyeron. Por este motivo, a lo largo de este trabajo se ha cuestionado su supuesta objetividad, a fin de dar cuenta que los algoritmos pueden discriminar arbitrariamente, y ya lo han hecho. En este punto se expresa también la paradoja de la transparencia, por la que mientras que con los datos masivos se accede invasivamente a información privada y sensible, los resultados de estos tratamientos están casi completamente rodeados de secreto legal e industrial.

Frente a estas situaciones, el Derecho no ha de plantearse en un estatus de negación de los dilemas referidos, sino, anticiparse a fin de evitar mayores problemas a futuro. Si lo consideramos desde una perspectiva constitucional (especialmente a partir de la redacción de Constitución Política desde cero, a propósito de nuestro actual contexto social), estamos ante una oportunidad de innovar en cuanto a la asimilación de estas nociones en nuestra Constitución, siendo de suma relevancia el correcto diseño, desarrollo, aplicación y auditoría de los algoritmos, de forma ética, socialmente justa y responsable, y con una perspectiva antropocéntrica.

Por tanto, el protagonista en la utilización de los sistemas de IA y los algoritmos será el factor humano, que continuará siendo clave en todas las etapas del proceso. Factor humano que debe ser consciente de que la IA sólo nos beneficiará en la medida en que su diseño y su uso no perpetúen ni amplifiquen injusticias sociales. Para ello, debe generar herramientas que contribuyan al desmantelamiento de desigualdades estructurales, donde todas las personas puedan aprovechar las ventajas de la IA, sin dejar a nadie atrás.

En ámbito de la justicia civil, consideramos que la función asistencial guarda una mayor consonancia con las posibilidades reales, y legales, para su utilización en procesos que se caracterizan por la repetición de tareas. También puede jugar rol importante la implementación de ODR para aquellas materias de baja cuantía y en las que se recomienda su desjudicialización.

En el ámbito de la justicia criminal, resulta pertinente destacar dos asuntos. Por un lado, llama la atención la nula implementación de sistemas de IA que tengan por meta, ya no predecir el riesgo de criminal de los individuos, sino que tengan por objeto alcanzar fines constructivos, como podría ser la rehabilitación y la reinserción social de los reos. Por el otro, en este ámbito el uso de sistemas de IA despierta mayores temores y se le califica como una “actividad de elevado riesgo”, pues implica decidir sobre la condena y absolución de una persona a partir de sistemas de evaluación de riesgo que pueden contener datos discriminatorios, lo que puede resultar incompatible con nuestro sistema penal garantista.

Referencias bibliográficas:

— Amunátegui Perelló, C. (2020): *Arcana Technicae. El derecho y la Inteligencia Artificial*. [e-book] Valencia: Tirant lo Blanch. Disponible a través de: Tirant Online <https://latam-tirantonline-com.pucv.idm.oclc.org/cloudLibrary/ebook/show/9788413550213?showPage=0> [Consultado 15-11-2021].

— Barona Vilar, S. (2021): *Algoritmización del Derecho y de la Justicia*. [e-book] Valencia: Tirant Lo Blanch. Disponible a través de: Tirant Online <https://latam-tirantonline-com.pucv.idm.oclc.org/cloudLibrary/ebook/show/9788413786667?showPage=0> [Consultado 15-06-2021]

— Battelli, E. (2020): “La decisión robótica: algoritmos, interpretación y justicia predictiva”, *Revista de Derecho Privado*, (40), pp. 45-86. DOI: [10.18601/01234366.n40.03](https://doi.org/10.18601/01234366.n40.03)

— Belloso Martín, N. (2019): “Algoritmos predictivos al servicio de la justicia: ¿una nueva forma de minimizar el riesgo y la incertidumbre?”, *Revista de Faculdade Mineira de Direito*, 22(43), pp. 1-31. Disponible en: <http://periodicos.pucminas.br/index.php/Direito/article/view/20780/16029>. [Consultado 14-11-2021]

— Boix Palop, A. (2020): “Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la administración para la adopción de decisiones”, *Revista de Derecho Público: Teoría y Método*, 1, pp. 223-270. DOI: [10.37417/RPD/vol_1_2020_33](https://doi.org/10.37417/RPD/vol_1_2020_33).

— Borges Blázquez, R. (2020): “El sesgo de la máquina en la toma de decisiones en el proceso penal”, *Ius et Scientia*, 6(2), pp. 54-71. DOI: [10.12795/IETSCIENTIA.2020.i02.05](https://doi.org/10.12795/IETSCIENTIA.2020.i02.05).

— Borges Blázquez, R. (2021a): *Inteligencia Artificial y proceso penal*. Navarra: Thomson Reuters Aranzadi.

— Borges Blázquez, R. (2021b): “La inteligencia artificial en el proceso penal y el ¿regreso? De Lombroso”. En: Barona Vilar, S. (ed.). *Justicia algorítmica y neuroderecho*. [e-book] Valencia: Tirant lo Blanch. Disponible a través de: Tirant Online <https://latam-tirantonline-com.pucv.idm.oclc.org/cloudLibrary/ebook/show/9788413972022?showPage=0> [Consultado 24-11-2021].

— Bostrom, N. y Yudkowsky, E. (2014): “The ethics of artificial intelligence,” En: Frankish, K., Ramsey, W. M. (eds.) *The Cambridge Handbook of Artificial Intelligence*. Cambridge: Cambridge University Press, pp. 316–334. DOI: 10.1017/CBO9781139046855.020.

— Browning, J. G y Shahrestani, A. (2019): “Ghosts in the Machine: Algorithmic Bias and the Courts”, *The Computer & Internet Lawyer*, 36(12), pp. 1-7. Disponible en: EBSCOhost <https://search-ebSCOhost-com.bibliotecadigital.uv.cl/login.aspx?direct=true&db=bsu&AN=139883696&lang=es&site=ehost-live> [Consultado 27-11-2021].

— Burgess, M., (2018): “UK police are using AI to inform custodial decisions – but it could be discriminating against the poor”, *Wired*, 01 de marzo. Disponible en: <https://www.wired.co.uk/article/police-ai-uk-durham-hart-checkpoint-algorithm-edit> [Consultado 10-09-2021]

— Casanovas, P. (2015): “Derecho, tecnología, inteligencia artificial y web semántica. Un mundo para todos y cada uno”. En: Fabra Zamora J.L. y Núñez Vaquero, A. eds. *Enciclopedia de filosofía y teoría del derecho*. México: UNAM, Instituto de Investigaciones Jurídicas, pp. 825-887. Disponible a través de: Biblioteca Jurídica UNAM <https://archivos.juridicas.unam.mx/www/bjv/libros/8/3875/26.pdf> [Consultado 15-11-2021]

— Casanueva, I. (2020): “La Audiencia Nacional condena a Interior por la inadecuada protección a una denunciante de violencia de género que fue asesinada”, *Confilegal*, 14 de octubre, Disponible en: <https://confilegal.com/20201014-la-audiencia-nacional-condena-a-interior-por-la-inadecuada-proteccion-a-una-denunciante-de-violencia-de-genero-que-fue-asesinada/> [Consultado 10-09-2021]

— Castellanos Claramunt, J. y Montero Caro, M. D. (2020): “Perspectiva constitucional de las garantías de aplicación de la inteligencia artificial: la ineludible protección de los derechos fundamentales”, *Ius et Scientia*, 6(2), pp. 72-82. DOI: [10.12795/IETSCIENTIA.2020.i02.06](https://doi.org/10.12795/IETSCIENTIA.2020.i02.06).

— Cerrillo i Martínez , A. y Velasco Rico, C. (2019): “Jurisdicción, algoritmos e inteligencia artificial”. En: López Ramón, F. y Valero Torrijos, J. cords., *20 años de la Ley de lo contencioso-administrativo*. Madrid: Instituto Nacional de Administración Pública. Disponible a través de: Dialnet <https://dialnet.unirioja.es/servlet/articulo?codigo=7316169> [Consultado 19-11-2021].

— Corvalán, J. G., (2018): “Inteligencia artificial: retos, desafíos y oportunidades – Prometea: la primera inteligencia artificial de Latinoamérica al servicio de la Justicia”, *Revista de Investigações Constitucionais*, 5(1), pp. 295-316. DOI: [10.5380/rinc.v5i1.55334](https://doi.org/10.5380/rinc.v5i1.55334).

___ Comisión Europea para la Eficiencia de la Justicia (2018): *La Carta Ética Europea sobre el uso de la Inteligencia Artificial en los sistemas judiciales y su entorno*. Estrasburgo: Consejo de Europa.

___ Comisión Europea para la Eficiencia de la Justicia (2020): *El Libro Blanco de la Inteligencia Artificial, un enfoque europeo orientado a la excelencia y la confianza*. Bruselas: Consejo de Europa.

___ Cotino Hueso, L. (2017): “Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales”, *Dilemata*, (24), pp. 131-150. Disponible en: <https://www.dilemata.net/revista/index.php/dilemata/article/view/412000104/494>

[Consultado 14-09-2021]

___ Cotino Hueso, L., (2019): “Derecho y garantías ante el uso público y privado de inteligencia artificial, robótica y big data”. En: Bauzá Reilly, M. ed. *El Derecho de las TIC de Iberoamérica*. [e-book] Montevideo: La ley Uruguay. Disponible a través de: Researchgate <https://www.researchgate.net/publication/349494500>

___ Defensoría Penal Pública (2020): “Se inició período de prueba de inédito piloto de inteligencia artificial”, 11 de noviembre 2021. Disponible en: http://www.dpp.cl/sala_prensa/noticias_detalle/10903/se-inicio-periodo-de-prueba-de-inedito-piloto-de-inteligencia-artificial [Consultado 30-08-2021].

___ De Hoyos Sancho, M. (2021): “Premisas y finalidades del Libro Blanco sobre Inteligencia Artificial de la Comisión Europea: perspectiva procesal del nuevo marco regulador”. En: Barona Vilar, S. (ed.) *Justicia algorítmica y neuroderecho*. [e-book] Valencia: Tirant lo Blanch. Disponible a través de: Tirant Online <https://latam-tirantonline-com.pucv.idm.oclc.org/cloudLibrary/ebook/show/9788413972022?showPage=129>

[Consultado 24-11-2021]

___ De Miguel Beriain, I. y Pérez Estrada, M. J. (2019): “La inteligencia artificial en el proceso penal español: un análisis de su admisibilidad sobre la base de los derechos fundamentales implicados”, *Revista de Derecho UNED*, (25), pp. 531-561. DOI: 10.5944/rduned.25.2019.27013.

___ Eguíluz Castañeira, J. A. (2020): “Desafíos y retos que plantean las decisiones automatizadas y los perfilados para los derechos fundamentales”, *Estudios de Deusto*, 68(2), pp. 325-367. DOI: 10.18543/ed-68(2)-2020pp325-367.

___ Freeman, K. (2016): “Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in *State v. Loomis*”, *North Carolina Journal of Law & Technology*, 18(5), pp. 75-106. Disponible en: <https://scholarship.law.unc.edu/ncjolt/vol18/iss5/3/> [Consultado 14-09-2021].

- Hernández, J. y Polanco Medina, J. (2020): “Problemas ético-jurídicos de las decisiones algorítmicas y el big data”, *Revista de Humanidades Cuadernos del Marqués de San Adrián*, (12), pp. 75-108. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=7877699> [Consultado 15-11-2021]
- Kaun, A. y Stiernstedt, F. (2020): “Doing time, the smart way? Temporalities of the smart Prison”, *New Media and Society*, 22(9), pp. 1580–1599. DOI: [10.1177/1461444820914865](https://doi.org/10.1177/1461444820914865).
- Lillo, R. (2020): “La justicia civil en crisis. Estudio empírico en la ciudad de Santiago para aportar a una reforma judicial orientada hacia el acceso a la justicia (formal)”, *Revista Chilena Derecho*, 47(1), pp. 119-157. Disponible en: <http://ojs.uc.cl/index.php/Rchd/article/view/14896/12242> [Consultado 25-11-2021]
- Lombana Bermúdez, A. (2018): “La evolución de las brechas digitales y el auge de la Inteligencia Artificial (IA)”, *Revista mexicana de bachillerato a distancia*, (20), pp. 17-25. DOI: 10.22201/cuaed.20074751e.2018.20.65884.
- López Oneto, M. (2020): *Fundamentos para un derecho de la inteligencia artificial*. [e-book] Valencia: Tirant lo Blanch. Disponible a través de: Tirant Online <https://latam-tirantonline-com.pucv.idm.oclc.org/cloudLibrary/ebook/show/9788413368849?showPage=0> [Consultado 30-10-2021]
- Lynn, B. (2019): “Robot Justice: The Rise of China’s ‘Internet Courts’”, *Voa News*, 11 de diciembre. Disponible en: <https://learningenglish.voanews.com/a/robot-justice-the-rise-of-china-s-internet-courts-/5201677.html> [Consultado 20-11-2021].
- Martínez Bahena, G. C. (2012): “La inteligencia artificial y su aplicación al campo del Derecho”, *Alegatos*, (82), pp. 827-846. Disponible en: <https://biblioteca.corteidh.or.cr/documento/65450> [Consultado 15-11-2021].
- Martín Diz, F. (2020): “Justicia digital post-covid19: el desafío de las soluciones extrajudiciales electrónicas de litigios y la inteligencia artificial”, *Revista de Estudios Jurídicos y Criminológicos*, (2), pp. 41-74. DOI: [10.25267/REJUCRIM.2020.i2.03](https://doi.org/10.25267/REJUCRIM.2020.i2.03).
- Martínez Garay, L. (2018): “Peligrosidad, algoritmos y due process: El caso State v Loomis”, *Revista de Derecho Penal y Criminología*, (20), pp. 485-502. DOI: [10.5944/rdpc.20.2018.26484](https://doi.org/10.5944/rdpc.20.2018.26484).
- Martínez Martínez, R. (2019): “Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo”, *Revista Catalana de Dret Públic*, (58), pp. 64-81. DOI: 2436/rcdp.i58.2019.3317.

- Mayson, S. G. (2019): “Bias in, bias out”, *The Yale Law Journal*, 128(8), pp. 2218–2300. Disponible en: <https://www.yalelawjournal.org/article/bias-in-bias-out> [Consultado 24-11-2021].
- Meseguer González, P. y López de Mántaras Badia, R. (2017): *Inteligencia artificial*. [e-book] Madrid: Editorial CSIC Consejo Superior de Investigaciones Científicas. Disponible a través de: E-Libro <https://elibro.net/es/lc/uvalparaiso/titulos/42319> [Consultado 09-06-2021]
- Miró Llinares, F. (2018): “Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots”, *Revista de Derecho penal y Criminología*, (20), pp. 87-130. DOI: [10.5944/rdpc.20.2018.26446](https://doi.org/10.5944/rdpc.20.2018.26446).
- Morales Cáceres, A. (2021): “El impacto de la inteligencia artificial en el Derecho”. *Advocatus*, (39), pp. 39-71. DOI: [10.26439/advocatus2021.n39.5117](https://doi.org/10.26439/advocatus2021.n39.5117).
- Mora Pedreros, P. A. (2017): “Software para argumentación, exploración del uso de las tecnologías en ejercicios argumentales”, *Libros Universidad Nacional Abierta y a Distancia*, pp. 54-65. Disponible en: <https://hemeroteca.unad.edu.co/index.php/book/article/view/2464> [Consultado 15-11-2021]
- Nava González, W., (2019): “Inteligencia artificial para la solución de controversias del acuerdo comercial de América del Norte”, *Revista Jurídica Iura*, 4(1), pp. 29-44. DOI: [10.22497/IURA.42.4203](https://doi.org/10.22497/IURA.42.4203)
- Navas Navarro, S. (2017): “Derecho e Inteligencia Artificial desde el diseño. Aproximaciones”. En: Navas Navarro, S. ed. *Inteligencia artificial Tecnología Derecho*. [e-book] México: Tirant lo Blanch. Disponible a través de: Tirant Online [Consultado 15-11-2021]
- Nieva Fenoll, J. (2018): *Inteligencia artificial y el proceso judicial*. [e-book] Madrid: Marcial Pons Ediciones Jurídicas y Sociales. Disponible a través de: E-Libro <https://elibro.net/es/lc/uvalparaiso/titulos/127422> [Consultado 09-06-2021]
- Organisation for Economic Co-operation and Development (2019): *Recommendation of the Council on Artificial Intelligence*. OECD Legal Instrument.
- Oswald, M., Grace, J, Urwin, S. y Barnes, G. C. (2018): “Algorithmic risk assessment policing models: lessons from the Durham HART model and ‘Experimental’ proportionality”, *Information & Communications Technology Law*, 27(2), pp. 223–250. DOI: 10.1080/13600834.2018.1458455.
- Otero, C. (2018): “Donotpay, una app para demandar a alguien desde el móvil”, *As*, 18 de octubre. Disponible en:

https://as.com/meristation/2018/10/18/betech/1539887121_135941.html [Consultado 26-11-2021]

___ Pascual, D. A. (2017): *Inteligencia artificial: un panorama de algunos de sus desafíos éticos y jurídicos* (Tesis inédita). Universidad de Girona.

___ Pillado González, E. (2021): “Algoritmos predictivos del comportamiento y proceso penal de menores”. En: Barona Vilar, S. (ed.) *Justicia algorítmica y neuroderecho*. [e-book] Valencia: Tirant lo Blanch. Disponible a través de: Tirant Online <https://latam-tirantonline-com.pucv.idm.oclc.org/cloudLibrary/ebook/show/9788413972022?showPage=0> [Consultado 25-11-2021]

___ Reyes Olmedo, P. (2020): “Límites a la tecnología: la ética en los algoritmos”. En: Robledo Granero, H. ed. *Inteligencia Artificial y Derecho, un reto social*. [e-book]. Ciudad autónoma de Buenos Aires: Albremática.

___ Rincón Cárdenas, E. y Martínez Molano, V. (2021): “Un estudio sobre la posibilidad de aplicar la inteligencia artificial en las decisiones judiciales”, *Revista Direito GV*, 17(1), pp. 1-29. DOI: [10.1590/2317-6172202101](https://doi.org/10.1590/2317-6172202101)

___ Russell, S. J. y Norvig, P. (2004): *Inteligencia artificial: Un enfoque moderno*. 2ª ed. Madrid: Pearson.

___ Salter, S. y Thompson, D. (2017): “Public-Centred Civil Justice Redesign: a case study of the British Columbia Civil Resolution Tribunal”, *McGill Journal of Dispute Resolution*, (3), pp. 113-136. Disponible en: <https://ssrn.com/abstract=2955796> [Consultado 14-09-2021]

___ Simón Castellano, P., (2021): “Inteligencia artificial y Administración de Justicia: ¿Quo vadis, justitia?”. *Revista D'Internet, Dret y Política*, (33), pp. 1-15. DOI: 10.7238/idp.v0i33.373817.

___ Solar Cayón, J. I. (2021): “Inteligencia Artificial y Derecho”. En: Solar Cayón, J. I. ed., *Dimensiones éticas y jurídicas de la inteligencia artificial en el marco del Estado de Derecho*. [e-book] Madrid: Editorial Universidad de Alcalá. Disponible a través de E-Libro <https://elibro.net/es/lc/uvalparaiso/titulos/172948> [Consultado 01-11-2021]

___ Tonry, M., (2014): “Legal and ethical issues in the prediction of recidivism”, *Federal Sentencing Reporter*, 6(3), pp. 167-176. Disponible en: https://scholarship.law.umn.edu/faculty_articles/525 [Consultado 24-11-2021]

___ Torres Manrique, J. I. (2017): “Breves consideraciones acerca del aterrizaje de la inteligencia artificial en el derecho y su influencia en la realización de los derechos fundamentales”, *Pensamiento americano*, 10(19), pp. 210-227. DOI: [10.21803/2Fpenamer.10.19.480](https://doi.org/10.21803/2Fpenamer.10.19.480).

— Vargas Pávez, M. (2013): “Hacia la desjudicialización de la ejecución civil”, *Revista Chilena de Derecho*, 40(1), pp. 135-156. DOI: [10.4067/S0718-34372013000100006](https://doi.org/10.4067/S0718-34372013000100006)

— Vivar Vera, J. (2021): “La sentencia penal, el juez y el algoritmo: ¿Las nuevas tecnologías serán nuestros próximos jueces?”, *Revista chilena de Derecho y Tecnología*, 10(1), pp. 231-269. DOI: [10.5354/0719-2584.2021.58785](https://doi.org/10.5354/0719-2584.2021.58785).

— Zerilli, J., Knott, A., Maclaurin, J. y Gavaghan, C. (2019): “Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard?”, *Springer Nature*, 32, pp. 661-683. DOI: [10.1007/s13347-018-0330-6](https://doi.org/10.1007/s13347-018-0330-6)