

ENTRE LO POSIBLE Y LO CONVENIENTE: CASOS DE USO DE INTELIGENCIA ARTIFICIAL EN LA JUSTICIA LABORAL

SEBASTIÁN SOLTAU SALAZAR

Abogado por la Pontificia Universidad Católica del Perú (PUCP). Máster en Derecho (LL.M.) por The George Washington University. Profesor de la Maestría en Derecho Procesal de la PUCP. Gerente de Relaciones Laborales del BCP.



“I use the state of the art technology. Supposed to make for better living. But are we better human beings?”¹

(James, 2013)

I. INTRODUCCIÓN: DE LA PROFECÍA A LA AGENCIA

Escuchamos con frecuencia que la inteligencia artificial (IA) está en todas partes. En las rutas de Waze rumbo a la oficina, en la categorización de los correos en Outlook durante la jornada de trabajo, en las publicidades de Instagram relacionadas a lo que estábamos conversando durante el almuerzo y en las sugerencias de Netflix cuando, al final del día, pretendemos “desconectarnos”.

Y, a veces, pareciera que todo esto ha ocurrido sin que nos diéramos cuenta. La IA es algo que *nos ha pasado*; no es algo que, conscientemente, hayamos decidido incorporar en nuestras vidas.

Podríamos debatir por horas si el balance de esta aproximación (un tanto acrítica) a la IA, a la fecha, es positivo. En mi opinión, sí lo es.

Sin embargo, creo que es indispensable que todos empecemos a asumir una actitud distinta. A sabernos responsables de lo que se va a venir y contribuir a que la eventual adopción o rechazo del uso de nuevas tecnologías en ámbitos tan relevantes como el de la justicia, sea el producto de un debate robusto y no de la inercia.

No perdamos de vista, que la justicia no es TikTok o alguna otra aplicación con fines comerciales, y que el riesgo inherente a la introducción de tecnologías como la IA podría trascender al acto de juzgar y afectar elementos funcionales esenciales del estado de derecho y los sistemas judiciales (European Commission for the Efficiency of Justice [CEPEJ], 2018, p. 15).

1 “Yo uso la tecnología de última generación. Se supone que nos dará una vida mejor, pero... ¿somos mejores seres humanos?”

Mi invocación con esta ponencia, en línea con lo señalado por Gigerenzer (2022), es a que nos preocupemos por seguir siendo inteligentes en un mundo inteligente (*smart*), lo que significa entender el potencial y los riesgos de las tecnologías digitales, y tener la determinación de seguir en control en un mundo poblado por algoritmos (p. XI). Para eso, no es necesario aprender a programar.

No caigamos presos de profecías que, por su propia naturaleza, contribuyen a hacerse realidad² (“es cuestión de tiempo para que los jueces y abogados sean reemplazados por la IA”), ni de argumentos que sobresimplifican el reto que, como sociedad, estamos afrontando³.

Gigerenzer (2022) lo resume muy bien: en vez de temor o humo (*hype*), el mundo digital necesita a ciudadanos mejor informados y saludablemente críticos, que quieran mantener el control de sus vidas en sus propias manos (p. XXI).

II. TRES IDEAS CLAVE SOBRE LA IA (PARA NO TÉCNICOS)

1. La IA no funciona como la inteligencia humana

Incluso los más optimistas promotores de la IA, fuera y dentro del mundo legal, reconocen que esta no funciona como la inteligencia humana.

Por ejemplo, Marc Andreessen, socio de Andreessen Horowitz, una de las firmas de capital de riesgo más reconocidas de los EE. UU., ha dicho – refiriéndose a GPT-4 y otros grandes modelos de lenguaje (*Large Language Models*) – que esto no es, en sentido estricto, inteligencia humana, en la medida que, para comenzar, seguimos sin entender el funcionamiento de esta (Roberts, 2023, 11m45s).

En la misma línea, Walker (2019) indica que la IA no es “artificial” ni es “inteligente”, y que la inteligencia humana no es replicada, de manera demostrable, por algoritmos; en primer lugar, porque no la entendemos (p. 58)⁴.

2 Este tipo de profecías son una manifestación de lo que, en la filosofía, se denomina el problema de la “performatividad” (Pasquale, 2020, p. X).

3 Por ejemplo, Gigerenzer (2022) se refiere al argumento “IA vence a humano”, que va así: La IA ha vencido a los mejores humanos en ajedrez y Go. El poder computacional se duplica cada par de años. Por consiguiente, pronto las máquinas van a hacer todo mejor que los humanos (p. X).

4 Gigerenzer (2022) anota que, como la I en IA no tiene relación con la inteligencia como la conocemos, la expresión toma de decisiones automatizada (ADM por sus siglas en inglés) suele utilizarse en su reemplazo (p. 27).

Susskind (2019) responde a la pregunta “¿pueden las máquinas pensar, trabajar, exteriorizar sentimientos, crear, razonar y sentir como un juez humano?” con un rotundo no (p. 250); sin embargo, también alude a la “falacia de la IA”, entendida como la idea de que la única manera para lograr que las máquinas tengan un mejor desempeño que los abogados humanos es copiar la forma en que estos trabajan, ignorando así que los sistemas de IA actuales no buscan imitar o replicar el razonamiento humano (p. 272). Volveremos a esta falacia en la sección V.

Ahora bien, aunque es cierto que aún no sabemos lo suficiente sobre el funcionamiento de la inteligencia humana, parece haber cierto consenso en que hay dos factores clave que la distinguen de la IA: la conciencia y el sentido común.

Con relación a lo primero, la CEPEJ (2019) ha resaltado que las IAs “fuertes” de la literatura de ciencia ficción no existen. Este tipo de sistemas, equipados no solo con inteligencia, sino también con conciencia, siguen siendo puramente ficticios (p. 31).

Refiriéndose a los emblemáticos triunfos de sistemas de IA en el ajedrez, Gigerenzer (2022) precisa que el programa de ajedrez que usa fuerza bruta no sabe que está siendo más listo que un jugador humano. De hecho, ni siquiera sabe que está jugando algo llamado ajedrez; simplemente es bueno jugándolo (p. 27).

Con respecto a lo segundo, Marcus y Davis (2019) explican que el sentido común es el tipo de conocimiento básico que esperamos que las personas ordinarias posean y que, paradójicamente, es algo que todo el mundo sabe, mas nadie parece saber qué es exactamente ni cómo construir máquinas que lo tengan (pp. 149-150).

Por su parte, para Gigerenzer (2022) el sentido común es conocimiento compartido sobre las personas y el mundo físico, habilitado por el cerebro biológico y que requiere solo una experiencia limitada (p. 74).

Según el mismo autor, la inteligencia humana consiste en representar el mundo, creando modelos causales y adscribiendo intenciones a otros seres vivos, algo que la IA – específicamente, las llamadas redes neuronales profundas (*deep neural networks*) – no pueden hacer (p. 90). Así, por ejemplo, estas redes no saben que un texto o una imagen sobre una persona o un objeto representan algo en el mundo real. Y es que no tienen un concepto de las cosas (p. 60).

Es por ello, que autores como Han (2021) califican al conocimiento que proporciona la *big data* – de la que se alimentan los sistemas de IA basados en datos – de rudimentario (p. 58). A criterio de este autor, “la correlación representa la forma más baja de saber” y se ubica dos niveles por debajo de la verdadera comprensión⁵, que solo se produce al nivel de concepto (pp. 57-58).

2. El boom de la IA en los últimos tiempos es el boom de los sistemas de IA basados en datos

No existe un concepto único de la IA. Como se trata de un conjunto de métodos científicos, teorías y técnicas que buscan replicar, a través de una máquina, las habilidades cognitivas de los seres humanos (CEPEJ, 2018, p. 69) hay quienes consideran que puede llegar a ser un concepto vacío (Amunátegui, 2020, p. 13). Otros, ante la indefinición o duda, prefieren darle un alcance que, a mi criterio, es demasiado amplio (Gigerenzer, 2022, p. XXII)⁶, mientras que otros, en tono burlón, lo califican de “meme del marketing” (Walker, 2019, p. 59)⁷.

En mi opinión, para comprender a un nivel básico el potencial y las limitaciones de la IA, más útil que buscar o tratar de construir un concepto único de esta, que – como hemos visto – no existe, es entender la diferencia entre los sistemas de IA basados en reglas y los sistemas de IA basados en datos. Es en esta última categoría que, sin ahondar en detalles técnicos, se insertan el aprendizaje automático (*machine learning*), el aprendizaje profundo (*deep learning*), las redes neuronales profundas (*deep neural networks*) y el procesamiento del lenguaje natural (*natural language processing*).

Amunátegui (2020) explica de manera muy clara que hay escuelas o aproximaciones filosóficas distintas detrás de cada tipo de sistema. Por un lado, la escuela simbolista (para los sistemas basados en reglas), que insiste “en codificar proposiciones lógicas en programas cada vez más extensos y complejos” (p. 27). Para los simbolistas, “pensar es razonar formalmente, y puesto que dicho

5 Según Han (2021), los tres niveles del saber son la correlación (de la que se valen los sistemas de IA actuales), la causalidad y el concepto (pp. 57-58).

6 Gigerenzer (2022) utiliza el término IA en un sentido amplio para abarcar cualquier tipo de algoritmo llamado a hacer lo que hace la inteligencia humana (p. XXII).

7 Walker (2019) califica a la IA de “meme del marketing” porque, en la actualidad, el término claramente se ha salido del control de las pedagogías estrictas de la ciencia de la computación, ingresando prolíficamente (y con frecuencia rentablemente) al imaginario popular de los negocios (p. 59).

razonamiento formal podía ser expresado en términos computables, la construcción de la máquina pensante era íntimamente dependiente de la capacidad de codificar ideas” (p. 27). Por otro lado, la escuela conectivista (para los sistemas basados en datos), que sostiene que “lo relevante es la capacidad de aprender, de formar correlaciones con el medio” (p. 27). Para los conectivistas, “las ideas y abstracciones surgirán del proceso de aprendizaje en sí, más adelante, cuando la máquina tuviese redes suficientemente densas” (p. 27).

Ahora bien, Susskind (2019) precisa que los sistemas de IA basados en reglas – también llamados sistemas expertos o de IA psicológica (Gigerenzer, 2022, p. 26) – eran sistemas que eran explícitamente programados para realizar ciertas actividades, a través de – en esencia – árboles de decisiones y flujogramas contruidos por desarrolladores humanos y que, en la actualidad, los sistemas que están de moda son los sistemas de IA basados en datos, que en vez de seguir reglas articuladas de manera explícita, “aprenden” de colecciones gigantes de datos pasados (pp. 264-265).

En la misma línea, la CEPEJ (2019), ha señalado que, hoy en día, la idea ya no es escribir reglas de razonamiento que sean un espejo del razonamiento humano, sino dejar que sistemas de aprendizaje automático (*machine learning*) identifiquen por sí mismos modelos estadísticos en los datos y los emparejen con resultados específicos (p. 33).

Y esto se debe a que, como lo enfatizan Marcus y Davis (2019), los sistemas de IA basados en datos – y, más específicamente, el aprendizaje profundo (*deep learning*) – han estado en el centro de prácticamente todos los avances de la IA en los últimos años, puesto que, en cada caso, la combinación de *big data* con aprendizaje profundo (*deep learning*) y *hardware* más rápido ha sido una fórmula ganadora (p. 10)⁸.

Pero, ¿qué es lo que hace un sistema de IA basado en datos? Para los fines de esta ponencia, basta que sepamos que estos sistemas son capaces de extraer patrones complejos y “aprender” de grandes volúmenes de datos de manera eficiente y, con frecuencia, con un alto nivel de precisión predictiva (CEPEJ, 2019, p. 31).

8 Amunátegui (2020) coincide con que los algoritmos de aprendizaje profundo (*deep learning*) son los que han determinado “el nuevo auge de la inteligencia artificial” (p. 34).

En otras palabras, son sistemas que “aprenden” de grandes colecciones de datos de ejemplos históricos, mediante la aplicación de un análisis estadístico (Deakin y Markou, 2020, p. 2) y que no aspiran a replicar la labor interpretativa de un ser humano, sino a reemplazarla por la correlación de grandes cantidades de información (CEPEJ, 2019, p. 37).

Cabe resaltar que esta es la manera en que ChatGPT y otras herramientas basadas en grandes modelos de lenguaje (*Large Language Models*) como GPT-4, responsables por el reciente *boom* de la IA generativa⁹, funcionan.

Al respecto, Wolfram (2023) aclara que el objetivo general de ChatGPT es continuar un texto de una manera “razonable”, en función a lo que ha visto durante el entrenamiento que ha tenido, que consistió en mirar billones de páginas de textos de Internet y otras fuentes. De lo que se trata, entonces, es de tratar de predecir cuál es la palabra que debe venir después en función a las palabras previas (Rowghani, 2023, 1m45s), que es algo muy distinto a lo que hacemos los seres humanos cuando nos comunicamos con alguien.

Volviendo a una de las ideas del subacápite previo, ChatGPT no sabe que está escribiendo un texto. De hecho, ni siquiera sabe que las personas y objetos mencionados durante su interacción con el usuario de turno existen en algo llamado el mundo real; simplemente es, a estas alturas, relativamente bueno escribiendo textos.

2.3. Los sistemas de IA basados en datos preocupan por el posible sesgo (*bias*), la opacidad y las alucinaciones

El potencial de los sistemas de IA basados en datos para transformar diversas industrias – entre ellas, la legal (Thomson Reuters, 2024) – y mejorar la vida de las personas en ámbitos tan relevantes como la salud (Bean, 2024) y la educación (Brown, 2024) está fuera de toda discusión.

No obstante, dicho potencial no borra el hecho de que existan aspectos de estos sistemas que generen preocupación; en especial, cuando lo que se propone es su aprovechamiento para complementar o asumir la toma de decisiones que podrían afectar los derechos e intereses de las personas.

9 La IA generativa es una subdivisión de la IA que empodera a las máquinas para crear contenido inédito, incluyendo texto, código, voz, imágenes, videos, procesos y hasta estructuras 3D de proteínas (Deloitte, 2023, p. 4).

El primero de ellos es el posible sesgo (*bias*).

Como bien resalta la CEPEJ (2019), los algoritmos, cualquiera sea su nivel actual de sofisticación, siguen circunscribiéndose a las interacciones establecidas en la etapa de aprendizaje y, por consiguiente, a su pasado. El contenido y la calidad de los flujos de datos utilizados en la composición de los cálculos es, por lo tanto, fundamental para entender los resultados obtenidos e identificar posibles sesgos analíticos (p. 35).

En otras palabras, como apunta Amunátegui (2020), “las tecnologías actuales de inteligencia artificial son completamente dependientes de los datos y su eficacia está determinada por ellos” (p. 46).

Según Christian (2020), el problema con un sistema que puede, en teoría, aprender prácticamente cualquier cosa de un conjunto de ejemplos es que este se encuentra a merced de los ejemplos con los cuales es entrenado (p. 26). Además, en palabras del mismo autor, si cierto tipo de datos está subrepresentado o incluso ausente de los datos de entrenamiento del sistema, pero presente en el mundo real, cualquier cosa puede pasar (p. 29).

Así, pues, al igual que los seres humanos, los algoritmos pueden discriminar a las mujeres, a las personas de raza negra o a otros grupos marginados, si de los datos con los que fueron entrenados se desprenden tales sesgos (Gigerenzer, 2022, p. 122)¹⁰.

El tema del sesgo (*bias*) nos lleva naturalmente a la segunda preocupación: la opacidad de los algoritmos.

Al respecto, Marcus y Davis (2019) explican que las redes neuronales están hechas de vastos despliegues de números, de los cuales virtualmente ninguno tiene sentido de manera intuitiva para los seres humanos ordinarios. Existe, pues, un misterio sin resolver sobre por qué las redes neuronales funcionan tan bien como lo hacen y una falta de claridad sobre las circunstancias exactas en las que no funcionan (p. 57).

En el mismo sentido, Christian (2020) apunta que, a medida que los modelos de aprendizaje automático (*machine learning*) se proliferan a lo largo de

10 Por ello, es muy importante que se sepa quién está representado en los datos y de qué manera, antes de usarlos para entrenar a sistemas que van a afectar a personas reales (Christian, 2020, p. 33).

la infraestructura de toma de decisiones del mundo, muchos se encuentran a sí mismos incómodos con lo poco que saben sobre lo que realmente está pasando dentro de esos modelos (p. 87).

Tanto Marcus y Davis (2019, p. 57) como Amunátegui (2020, p. 50) inciden en que estos sistemas no son capaces de dar explicaciones para sus respuestas, sean correctas o incorrectas, de la misma forma en que un ser humano las daría. En términos legales, podríamos decir que se trata de sistemas incapaces de motivar.

Nos encontramos, pues, frente a auténticas cajas negras, inescrutables a un nivel técnico (Deakin y Markou, 2020, p. 6), y no a algoritmos transparentes, que puedan ser entendidos, memorizados, enseñados y ejecutados por sus usuarios (Gigerenzer, 2022, p. 113).

Ahora bien, aunque suene contradictorio, esta opacidad o falta de transparencia, suele ser tomada como un signo de mayor sofisticación o fiabilidad.

Sobre el particular, Christian (2020) destaca que, en el campo de la IA, se suele pensar que los modelos más potentes son los que, en líneas generales, son los menos inteligibles, y que los más inteligibles son los menos precisos.

Gigerenzer (2022) coincide con esta apreciación, pues indica que la combinación de la fe en la complejidad (los problemas complejos siempre requieren soluciones complejas) y la fe en la opacidad (los algoritmos más precisos tienen que ser inherentemente incomprensibles) nos lleva a la creencia en el falso dilema de la precisión-transparencia: a más preciso el algoritmo, menos transparente va a ser (p. 129).

La tercera y última fuente de preocupación a la que me quiero referir brevemente son las alucinaciones, popularizadas desde el lanzamiento de ChatGPT y otras herramientas de IA generativa a finales del 2022, y que ya han metido en problemas a más de un abogado descuidado (Moran, 2023).

Al respecto, Marcus y Davis señalan que el aprendizaje profundo (*deep learning*) es frágil (*brittle*), en el sentido que puede dar resultados perfectos en una situación y absolutamente equivocados en otra. Asimismo, resaltan que las alucinaciones no son algo que ocurra esporádicamente, sino que siguen siendo un problema persistente, incluso años después de haber sido advertidas por primera vez (p. 58).

Vimos más arriba que los sistemas de IA no tienen conciencia ni tienen un concepto de las cosas y que, además, el proceso que siguen para, por ejemplo, escribir un texto es muy distinto al que seguimos las personas.

Siendo ello así, no debería sorprendernos que herramientas como ChatGPT de vez en cuando alucinen o, lo que es lo mismo, se inventen información y la transmitan con total seguridad, con tal de cumplir la tarea que les ha sido encomendada.

Y es que, al interactuar con un sistema de IA basado en datos, no podemos olvidarnos que, a fin de cuentas, por mucho que parezca que estamos interactuando con una persona, en realidad estamos frente a un modelo que, en el caso de ChatGPT y otras herramientas de IA generativa, ha sido entrenando para predecir la siguiente palabra, por lo que no sabe que no debería ser deshonesto (Rowghani, 2023, 3m10s) ni entiende lo que es la honestidad.

III. UN BREVE REPASO DE TRES INICIATIVAS EN LA REGIÓN: ¿CUÁNTO SABEMOS REALMENTE?

Como ya he sustentado mi escepticismo sobre los “jueces robot” en China en otro espacio (Soltau, 2020) y el propio gobierno de Estonia ha calificado de engañosos a artículos sobre sus planes en materia de IA y justicia (Ministry of Justice of the Republic of Estonia, 2022), he preferido centrarme en tres iniciativas que, aunque menos marqueteadas y disruptivas (por lo menos, sobre el papel), son mucho más cercanas a nuestra realidad. De hecho, la última de ellas corresponde al Poder Judicial peruano, aunque – como veremos – todo parece indicar que no superó la fase de piloto.

1. Prometea – Argentina

¿Quién desarrolló este sistema? La Fiscalía de la Ciudad Autónoma de Buenos Aires en el 2017 (Estevez, Linares y Fillottrani, 2020, p. 4).

¿En qué proceso se inserta? En el proceso de elaboración de dictámenes jurídicos por parte de dicha Fiscalía.

En todos los casos que llegan al Tribunal Superior de Justicia, la Fiscalía debe emitir un dictamen con una propuesta de solución, cuyo contenido no es vinculante para los jueces (Estevez, Linares y Fillottrani, 2020, p. 46).

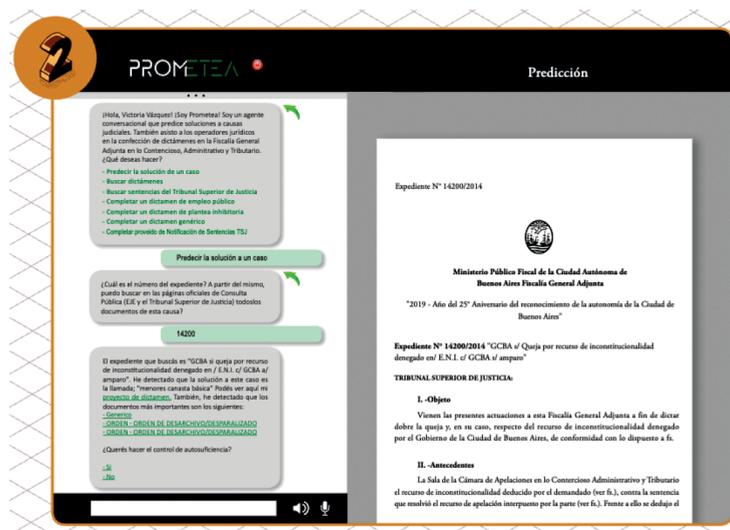
Según algunas fuentes, Prometea sólo se utiliza, en todo su potencial, en los ámbitos administrativo y tributario (ISICRI, 2021, 25m25s).

¿Qué problema busca resolver? La demora en la emisión de los dictámenes por parte de la Fiscalía, que antes de la implementación de Prometea podía tomar hasta tres meses (Estevez, Linares y Fillottrani, 2020, p. 11).

¿Cuál es su funcionalidad principal? Generar un proyecto de dictamen, a partir de la “lectura” de los documentos del caso concreto y su comparación con sentencias previas emitidas en casos similares (Estevez, Linares y Fillottrani, 2020, pp. 12-13).

Cabe resaltar que este proyecto tiene que ser revisado y complementado por el usuario, a través de la interacción con un asistente virtual (Estevez, Linares y Fillottrani, 2020, p. 51).

Según Estevez, Linares y Fillottrani (2020), “PROMETEA utiliza técnicas de IA para (...) realizar un análisis con inferencia predictiva de la solución jurídica a recomendar. Este análisis se basa en una correlación estadística entre palabras clave asociadas a cada proceso y patrones de sentencias previas” (p. 14).



Interfaz de Prometea (Estevez, Linares y Fillottrani, 2020, p. 41).

¿Cuáles han sido sus resultados, en términos de eficiencia y efectividad? Según Estevez, Linares y Fillottrani (2020), Prometea habría generado un ahorro promedio de 54 minutos en la elaboración de un dictamen. El tiempo promedio

de elaboración, considerando los distintos tipos de procesos (amparos habitacionales, entre otros), se habría reducido de 72 a 18 minutos. Esto representa una eficiencia del 75% (p. 21).

En lo que respecta a la eficacia, entendida – en este caso – como coincidencia entre el sentido de la decisión del Tribunal Superior de Justicia y el dictamen emitido por la Fiscalía con el apoyo de Prometea, Estevez, Linares y Fillottrani (2020) reportan que, en el 2017, tal coincidencia existió en 92.2% de los casos, mientras que, en el 2018, se presentó en el 100% de los casos (p. 26).

2. PretorIA – Colombia

¿Quién desarrolló este sistema? El Laboratorio de Innovación e Inteligencia Artificial de la Universidad de Buenos Aires (IA LAB de la UBA), en coordinación con la Corte Constitucional de Colombia y la Universidad del Rosario (IA LAB de la UBA, 2020, 0m24s) en el 2019, aunque bajo el nombre “Prometea” y con muchas más funcionalidades, que las que tiene actualmente (Saavedra y Upegui, 2021).

Según Saavedra y Upegui (2021), es desde julio de 2020, mes en el que el sistema es presentado por primera vez de forma pública bajo el nombre “PretorIA”, que la Corte Constitucional de Colombia se apropia de él y valiéndose de un equipo propio e interno, termina de darle la forma que tiene actualmente (pp. 34-35).

¿En qué proceso se inserta? En el proceso de selección de las tutelas (el equivalente a los amparos en el Perú) que llegan a la Corte Constitucional de Colombia para revisión (Saavedra y Upegui, 2021, p. 5), sea de oficio o con motivo de la interposición de algún recurso (Venegas, 2021, 2m44s).

¿Qué problema busca resolver? La carga excesiva de trabajo derivada de la selección de las tutelas que deben ser revisadas por la Corte Constitucional de Colombia (Saavedra y Upegui, 2021, p. 22), que seguramente se traduce en demoras, no solo en la selección de las tutelas, sino en la generalidad de las tareas, jurisdiccionales y administrativas, de dicha corte.

Por ejemplo, en el 2019, se remitieron a la Corte Constitucional de Colombia más de 600 mil expedientes de tutela (Saavedra y Upegui, 2021, p. 22); es decir, poco menos de 1700 expedientes por día.

¿Cuál es su funcionalidad principal? Etiquetar los expedientes de tutela que llegan a la Corte Constitucional de Colombia para revisión y generar un documento “que detalla qué categorías de importancia están presentes en el texto y cuáles no” (ISICRI, 2021, 16m46s) (por ejemplo, si el caso se refiere a un menor de edad o a una mujer embarazada), lo que agiliza la selección.

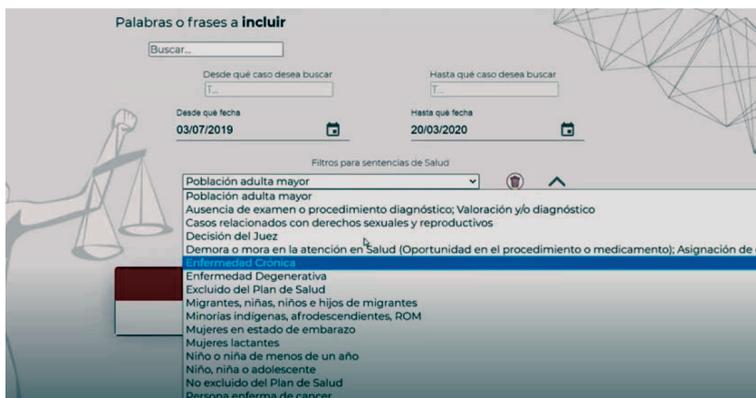
PretorIA no selecciona los casos. Esa es una tarea que sigue en manos de personas (ISICRI, 2021, 30m20s). Tampoco interviene en la etapa de elaboración de las sentencias, aunque haya fuentes que indiquen lo contrario (Resolución Administrativa No. 000040-2021-CE/PJ, 2021, p. 3 del “Plan para Implementar Proyecto Tucuy Ricuy”).

Cabe resaltar que existen versiones aparentemente contradictorias sobre las técnicas de las que se vale esta herramienta.

Por un lado, Víctor Saavedra (ISICRI, 2021) señala que, en el caso de PretorIA, “no es *machine learning* [aprendizaje automático], sino clasificación basada en reglas”, toda vez que “el proceso de clasificación o etiquetado no lo hace la máquina sobre la base de lo que ha aprendido, sino que lo hace a través de categorías que han sido modelizadas por el agente humano” (18m33s).

Por otro lado, Juan Corvalán (IA LAB de la UBA, 2020) afirma que PretorIA utiliza “una técnica de *machine learning* de caja blanca que se llama *topic model*” (5m9s).

Esta diferencia podría deberse a la mutación de PretorIA desde su concepción en el 2018 (bajo otro nombre) hasta su lanzamiento oficial (Saavedra y Upegui, 2021).



Interfaz de PretorIA (IA LAB de la UBA, 2020)

¿Cuáles han sido sus resultados, en términos de eficiencia y efectividad? No he podido encontrar fuentes confiables sobre los resultados de PretorIA, en términos de eficiencia y efectividad.

Si bien Estevez, Linares y Fillottrani (2020) incluyeron un recuadro titulado “PROMETEA EN LA CORTE CONSTITUCIONAL DE COLOMBIA” en un informe sobre la implementación de Prometea en Argentina, los datos que contiene sobre la cantidad de expedientes de tutela que llegan a la Corte Constitucional de Colombia (por ejemplo, indica que la corte recibe, en promedio, 2700 acciones de tutela por día) y las funcionalidades del sistema (por ejemplo, se habla de la selección de casos urgentes – no del etiquetado o pre-selección – en 2 minutos) no son compatibles con lo que encontré en otras fuentes. Además, el *link* a la fuente de este recuadro (alojada en la página web de DPI Cuántico – Derecho para Innovar) estaba roto (pp. 68-69).

En un artículo publicado en 2023, López et al. citan una infografía que supuestamente habría estado alojada en la página web del IA LAB de la UBA y que contendría datos sobre la eficiencia y eficacia de PretorIA¹¹ (p. 7). No obstante, el *link* a esta infografía también estaba roto.

A mediados de 2020, Juan Corvalán (Venegas, 2021) reconocía que PretorIA estaba “en fase de testeo con acciones reales” y que estaban “chequeando hasta qué punto el sistema rinde tal cual nosotros lo habíamos entrenado” (3m13s) y, a finales de 2021, Víctor Saavedra (ISICRI, 2021) indicaba que “si bien la Corte Constitucional ya ha llevado a producción este sistema, todavía no tiene claro cómo se va a insertar definitivamente en el el proceso” (22m8s).

Las idas y venidas sobre las funcionalidades de PretorIA y la no definición y publicación de métricas de éxito, a más de un año de su lanzamiento oficial, me hacen sospechar que, más allá de los titulares, que han venido bajando en intensidad y frecuencia desde el 2019, quizás esta herramienta no ha dado los resultados esperados o, peor aún, ha caído en desuso.

11 Según esta infografía, el sistema habría sido entrenado para identificar 33 categorías de importancia y en 13 de ellas habría presentado una tasa de acierto del 95% y en 9 de ellas de más del 90%. Por otro lado, en lo que respecta a la eficiencia, se habría comprobado que una persona se tardaría 27 minutos en hacer lo que PretorIA es capaz de hacer en segundos (López et al., 2023, p. 7).

3. Tucuy Ricuy – Perú

¿Quién desarrolló este sistema? El IA LAB de la UBA, en coordinación con el Poder Judicial peruano (Resolución Administrativa No. 000040-2021-CE/PJ, 2021).

¿En qué proceso se inserta? En el proceso de atención de las denuncias por violencia familiar enviadas por las comisarías a los Juzgados de Familia (IA LAB de la UBA, 2021).

¿Qué problema busca resolver? La demora en la atención de las denuncias por violencia familiar.

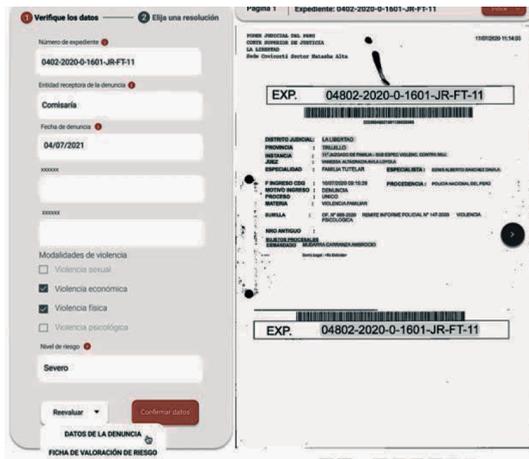
Según el plan para implementar este proyecto, “desde el año 2015 hasta el año 2017 se ha registrado un crecimiento superior al mil por ciento (1148%) respecto al ingreso de nuevas causas en materia de violencia familiar” (Resolución Administrativa No. 000040-2021-CE/PJ, 2021, p. 2 del “Plan para Implementar Proyecto Tucuy Ricuy”).

Cabe resaltar que este plan no contiene información sobre el tiempo promedio que toma atender – entiéndase, emitir las medidas de protección que correspondan – una denuncia por violencia familiar.

¿Cuál es su funcionalidad principal? Clasificar las denuncias en función al nivel de riesgo (sin riesgo, leve, moderado o severo) y generar proyectos de los documentos a emitir, incluidos – de corresponder – los que se refieren al dictado de las medidas de protección (IA LAB de la UBA, 2021).

Hay un punto técnico que me genera dudas sobre este sistema: ¿ha sido entrenado – con data histórica de casos de violencia familiar – para detectar el nivel de riesgo o lo que hace es extraer ese dato de la ficha de valoración del riesgo que viene llenada a mano? Juan Corvalán da pie a interpretar ambas cosas en un video de presentación de un prototipo avanzado de este sistema¹².

12 En un video del IA LAB de la UBA (2021), primero habla sobre la necesidad de “seguir reentrenando al sistema porque claramente para detectar la violencia o el riesgo severo que proviene de las comisarías es muy importante alimentar al sistema con el historial de casos que provienen de las comisarías” (2m18s), pero luego dice que “la ficha de valoración del riesgo viene llenada a mano” y que “en el sistema se pueden modificar los datos, pero no es que el sistema haga automáticamente una reevaluación” (4m40s).



Interfaz de Tucuy Ricuy (IA LAB de la UBA, 2021)

¿Cuáles han sido sus resultados, en términos de eficiencia y efectividad? No he podido encontrar fuentes sobre los resultados de Tucuy Ricuy, en términos de eficiencia y efectividad.

En teoría, según una nota de prensa del Poder Judicial, el testeo en juzgados piloto iba a comenzar en setiembre de 2021 (Plataforma digital única del Estado Peruano, 2021); sin embargo, no he podido encontrar fuentes recientes sobre el proyecto, que no se limiten a referirse a otras fuentes del 2021.

En resumen, tenemos:

- Un sistema que genera proyectos de dictámenes para fiscales, con buenos resultados en términos de eficiencia y efectividad;
- un sistema que etiqueta expedientes para agilizar su selección, sin resultados reportados; y,
- un sistema que clasifica denuncias en función al nivel de riesgo y genera proyectos de los documentos a emitir, sin resultados reportados.

En todos los casos, el sistema no tiene la palabra final (no presenta el dictamen, no selecciona el expediente, no emite la medida de protección). Siempre hay una revisión o validación por parte de una persona.

IV. LO POSIBLE: UNA LLUVIA DE IDEAS, RÁPIDA Y SIN FILTRO, DE CASOS DE USO

Como hemos visto, los sistemas de IA basados en datos son la razón de ser del “verano de la IA” (Amunátegui, 2022, p. 12) que estamos viviendo.

Por ello, es razonable que exista la expectativa de que cualquier herramienta, que sea marquetada como apalancada en IA, haya “aprendido” a realizar una o más actividades a partir de su entrenamiento con grandes cantidades de datos y que, además, lo haga con algún nivel de autonomía.

De igual modo, cuando se difunde que una persona ha incorporado la IA en algún ámbito de su vida (por ejemplo, el profesional), la percepción que se genera, casi automáticamente, es que la IA está siendo utilizada para hacer algo que no se podía hacer antes o que, en todo caso, no se podía hacer con la misma eficiencia y efectividad con otras herramientas que ya estaban disponibles.

No obstante, es común que automatizaciones simples (como las que se pueden crear con Excel o con plataformas *no-code* para la construcción de flujos de trabajo) sean difundidas como casos de uso de IA o que personas se vuelvan fugazmente famosas por ser las primeras en utilizar la IA para hacer operaciones aritméticas (LP Pasión por el Derecho, 2023) o buscar una fuente (Ámbito, 2023).

Es cierto que la IA puede ser aprovechada para una infinidad de cosas por cualquier persona. Por ejemplo, ChatGPT y otras herramientas de IA generativa pueden ser utilizadas para redactar, resumir y comparar textos de todo tipo, y también para investigar diversos temas, siempre y cuando el usuario tenga claro que, dado el riesgo de que la herramienta alucine, sigue siendo su responsabilidad verificar la veracidad de la información proporcionada. Todos estos usos podrían ser beneficiosos, qué duda cabe, para un juez laboral y el equipo que lo apoya.

Dicho esto, los siete casos de uso que listo a continuación, pensando únicamente en lo que ya es técnicamente posible, se vinculan específicamente a algunas actividades que el órgano jurisdiccional está llamado a realizar en ciertas etapas de un proceso judicial laboral¹³:

- Un sistema que genere proyectos de autos admisorios.

13 Para efectos de este ejercicio, me he basado en las actividades y etapas que se desprenden de la configuración del proceso judicial laboral prevista en la Ley 29497, Nueva Ley Procesal del Trabajo.

En su versión más extrema, sería el sistema el que decidiría si la demanda es admitida o no.

- Un sistema que, para efectos de la evaluación de una solicitud cautelar, estime la probabilidad de que la demanda sea declarada fundada o infundada.

En su versión más extrema, sería el sistema el que decidiría sobre la configuración del presupuesto de la verosimilitud del derecho.

- Un sistema que, al inicio de la etapa de conciliación, estime la probabilidad de que la demanda sea declarada fundada o infundada, y el monto que se ordenaría pagar en una eventual sentencia.

Este sistema incluso podría ser puesto a disposición de las partes antes de la audiencia de conciliación e ignorar su predicción podría ligarse a ciertas consecuencias (p. ej. el pago de intereses, costos y costas).

- Un sistema que, durante la etapa de conciliación, genere fórmulas conciliatorias.
- Un sistema que, con posterioridad a la etapa de confrontación de posiciones, proponga una lista de los puntos controvertidos o hechos necesitados de actuación probatoria.

En su versión más extrema, sería el sistema el que fijaría los puntos controvertidos o hechos necesitados de actuación probatoria.

- Un sistema que genere proyectos de sentencias (de primera instancia, de vista y casatorias).

En su versión más extrema, sería el sistema el que emitiría las sentencias.

- Un sistema que precalifique la procedencia de los recursos de casación.

En su versión más extrema, sería el sistema el que calificaría la procedencia de los recursos de casación.

No he hecho referencia a sistemas que generen proyectos de o emitan resoluciones de mero trámite, como – por ejemplo – las que programan una audiencia, pues considero que para este tipo de automatizaciones no tiene sentido valerse de la IA.

Un comentario final con relación a los casos de uso listados.

La distinción entre los supuestos en los que los sistemas de IA proyectan o sugieren, y los supuestos en los que deciden, a los que he calificado de “versiones más extremas”, puede terminar siendo un tanto artificial.

Y es que, como advierte Gigerenzer (2022), existe el peligro de que los jueces caigan en la toma de decisiones defensiva, que consiste en preferir los proyectos o sugerencias de los sistemas de IA para no correr el riesgo de que luego sus decisiones sean cuestionadas (p. 129)¹⁴.

V. LO CONVENIENTE: EXISTE UN MARCO DE REFERENCIA, PERO... ¿SERÁ EL MISMO EN 5, 10 O MÁS AÑOS?

En 2018, la CEPEJ (2019) aprobó la “Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno”.

En este documento, refiriéndose al fenómeno de la justicia predictiva, la CEPEJ advierte sobre la necesidad de considerar si las soluciones de IA en la justicia son compatibles con los derechos individuales previstos en la Convención Europea de Derechos Humanos, tales como el derecho a un juicio justo (en particular, el derecho al juez natural predeterminado por ley, el derecho a un tribunal independiente e imparcial, y la igualdad de armas en los procedimientos judiciales) (p. 15). Asimismo, remitiéndose a los principios delineados por el Tribunal Europeo de Derechos Humanos, la CEPEJ pone énfasis en la necesidad de balancear la certeza legal, que hace a las decisiones más predecibles, con la vitalidad de la interpretación judicial (p. 23).

En 2019, la Organización para la Cooperación y el Desarrollo Económico (OECD por sus siglas en inglés) adoptó la “Recomendación del Consejo sobre la Inteligencia Artificial”, que resalta la importancia de la transparencia y la explicabilidad, conforme a las cuales, las personas y organizaciones responsables por el lanzamiento y operación de sistemas de IA deberían comprometerse a,

¹⁴ Por ejemplo, imaginemos que un sistema de IA genera un proyecto de sentencia que declara infundada la demanda de reposición por despido fraudulento interpuesta por un trabajador que ha sido despedido por apropiarse de bienes de su empleador. ¿Qué sucedería si el juez decidiera darle la contra al sistema y amparar la demanda, y luego ese mismo trabajador, una vez repuesto, fuese despedido nuevamente por la misma falta, con base en prueba casi plena (múltiples videos, etc.)? Para no exponerse a situaciones de este tipo, los jueces podrían comenzar a alinearse a los proyectos de la IA, con lo cual, en la práctica, estos dejarían de ser meros proyectos.

entre otras cosas, posibilitar que las personas afectadas por un sistema de IA puedan entender el resultado y cuestionarlo, con base en información simple y fácil de entender sobre los factores y la lógica que se siguió para la predicción, recomendación o decisión (pp. 8-9).

En 2021, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) (2022) adoptó la “Recomendación sobre la ética de la inteligencia artificial”, en la que se indica que:

La falta de transparencia también podría mermar la posibilidad de impugnar eficazmente las decisiones basadas en resultados producidos por los sistemas de IA y, por lo tanto, podría vulnerar el derecho a un juicio imparcial y a un recurso efectivo, y limita los ámbitos en los que estos sistemas pueden utilizarse legalmente. (p. 22)

Dicha recomendación precisa que:

En caso de que los sistemas de IA sean utilizados por el poder judicial, se necesitan suficientes salvaguardias para garantizar, entre otras cosas, la protección de los derechos humanos fundamentales, el estado de derecho, la independencia judicial y el principio de supervisión humana, así como para asegurar un desarrollo y una utilización de los sistemas de IA en el poder judicial que sean fiables, orientados al interés público y centrados en el ser humano. (p. 28)

Ahora bien, en el ámbito local, si bien la Ley 31814 (2023) declaró de interés nacional el fomento del desarrollo y uso de la inteligencia artificial para la mejora de la justicia (Art. 2), es muy probable que su reglamento (Resolución Ministerial No. 132-2024-PCM, 2024, literal g) del artículo 14.2 del proyecto de reglamento de la Ley 31814)¹⁵, inspirado en la Ley de Inteligencia Artificial de la Unión Europea (2024, literal b) del numeral 8 del Anexo III)¹⁶, vaya a calificar a los sistemas de IA para la justicia como de alto riesgo, lo que quiere decir que sus implementadores van a tener que cumplir con una serie de obliga-

15 Se considera que un sistema de IA es de riesgo alto cuando se utiliza como “apoyo en la toma de decisiones, interpretación normativa o de los hechos por una autoridad jurisdiccional”.

16 Se califica como sistema de IA de alto riesgo a los “sistemas de IA destinados a ser utilizados por una autoridad judicial o en su nombre para asistir a una autoridad judicial en la investigación e interpretación de hechos y del Derecho y en la aplicación del Derecho a un conjunto concreto de hechos, o a ser utilizados de forma similar en la resolución alternativa de litigios”.

ciones, que – en la práctica – van a desincentivar el desarrollo e implementación de estos sistemas.

Más allá de que, sabiendo lo que sabemos sobre el funcionamiento de los sistemas de IA basados en datos, parezca muy difícil (por no decir imposible) que estos cumplan con las exigencias de los instrumentos repasados, que – por cierto – han sido acogidas a nivel local al mencionar el desarrollo ético para una IA responsable como uno de los principios para el desarrollo y uso de la IA (Ley 31814, 2023, literal e) del Art. Único), es importante que entendamos que, detrás de este marco, hay una concepción de lo que son y deben ser el derecho y el proceso, que – queramos o no – encuentra un correlato en nuestra Constitución Política (1993).

Como se puede comprobar leyendo a autores como Pasquale (2020) y Susskind (2019), estudiosos de la relación entre la justicia y la tecnología pueden tener respuestas diametralmente distintas a las preguntas “¿qué es el derecho?” y “¿para qué sirve el proceso?”

Según Pasquale (2020), para el pragmatista más firme (*tough-minded*), la pregunta “¿qué es el derecho?” se puede reducir a una simple predicción de si la autoridad va o no a actuar para detener una acción planificada o penalizar una acción realizada, a pesar de que el derecho y el proceso legal son mucho más que eso (p. vi).

Para dicho autor, dado que el derecho es una institución humana que se ocupa principalmente de actividades humanas, es probable que la IA nunca vaya a ser capaz de explicar por qué un caso debe ser resuelto de cierta manera y qué tan amplio o acotado debe ser el alcance de una decisión, puesto que estas habilidades y competencias requieren juicio normativo y solo pueden ser realizadas de manera legítima por una persona que pueda entender subjetivamente lo que ella o él le está imponiendo a otros cuando se toman decisiones (p. vii).

Susskind (2019), por el contrario, menciona que, en cierto sentido, lo que los abogados hacen realmente es predecir el derecho (p. 254) y que, al final, lo que realmente importa en el derecho es la manera en que determinado problema legal va a ser resuelto por la autoridad (p. 256).

Ahora bien, en estos momentos, nuestra Constitución Política (1993) reconoce derechos que, en mi opinión, son auténticas barreras para la implementación de sistemas de IA que aspiren a reemplazar a los jueces humanos.

Me refiero al derecho al juez natural (Art. 139.3) y a la debida motivación de las resoluciones judiciales (Art. 139.5), por poner solo un par de ejemplos. Ese es el estado de las cosas.

Pero lo que nos deberíamos preguntar es si eso siempre va a ser así. ¿Qué ocurriría si una inmensa mayoría de las personas, en un momento dado, prefiriera que sus conflictos fueran resueltos de manera rápida y predecible, aunque ello implicase reemplazar el juicio y la persuasión por la combinación y la clasificación (Pasquale, 2020, p. xiv), o sustituir la comprensión del mundo real por la estadística (Marcus y Davis, 2019, p. 90)? ¿No sería lógico que esa preferencia eventualmente se traslade al derecho y al proceso, propiciando la transformación y hasta eliminación de instituciones y garantías que hoy parecen intocables?

Lo cierto es que, por lo menos yo, no me ánimo a descartar un futuro así, en un contexto de crisis global de acceso a la justicia (Task Force for Justice, 2019)¹⁷ y dudas legítimas, basadas en evidencia empírica, sobre el proceso de toma de decisiones de los jueces humanos (Kahneman, Sibony y Sunstein, 2021, pp. 15-16).

Puedo, pues, concebir un futuro, no demasiado lejano, en el que la gran masa de casos de baja complejidad, en los que no se define la manera en que se van a abordar los principales problemas sociales (Pasquale, 2019, p. xiv), sea resuelta utilizando métodos “inhumanos”, ahora que existe una alternativa viable para obtener los resultados sociales y económicos que, hoy en día, son dados por jueces y cortes (Susskind, 2019, pp. 286-287).

Además, Gigerenzer (2022) explica que las nuevas tecnologías no simplemente nos asisten, pues su efecto es mucho más profundo. Para beneficiarnos de su máxima eficiencia, tenemos que estar preparados para cambiar nuestro comportamiento y nuestro entorno (p. 54).

Esta adaptación ya se está notando en ámbitos como la medicina, en los que ya existen *chatbots* que obtienen mejores calificaciones que los doctores humanos, no solo en cuanto a la calidad de sus respuestas, sino a su empatía (Ayers et al., 2023). Y en la resolución de disputas, cada vez es más común que, por lo menos para conflictos del día-a-día y de baja cuantía, aceptemos soluciones predefinidas o automatizadas, algo que las plataformas de resolución de disputas

17 Según este reporte, dos tercios de la población mundial no tiene un acceso significativo a la justicia (p. 18).

en línea (ODR por sus siglas en inglés), vienen haciendo desde inicios de este siglo (Katsh y Rabinovich-Einy, 2017).

Refiriéndose al juez computador, Susskind (2019) indica que deberíamos esperar que las posturas no sólo cambien a lo largo del tiempo, sino que además varíen de un lugar a otro. Coincido con lo señalado por este autor, en el sentido que es muy probable que nuestros nietos tengan un punto de vista diferente al nuestro (p. 292).

VI. REFLEXIÓN FINAL

La incorporación de la IA en la justicia no es un tema del futuro. Es algo de lo que, como hemos visto, ya se habla mucho e incluso existen experiencias a nivel global y en nuestra región, con variados niveles de éxito.

La presión por al menos experimentar con estos sistemas sólo va a crecer en los próximos años por distintas razones.

Al respecto, Pasquale (2019) advierte que, primero, el canto de sirena de la austeridad en medio de ralentizaciones económicas va a tentar a los legisladores a tratar de reemplazar a trabajadores del sistema de justicia con *software* menos costoso y que, segundo, negocios tecnológicos con altos márgenes van a seguir reinvertiendo sus utilidades en hacerle *lobby* a los gobiernos para que estos se animen a usar sus productos (p. v).

Por su parte, la CEPEJ (2019) incide en que la expansión de la apuesta por la toma de decisiones automatizada está vinculada a factores psicológicos como la creencia en la objetividad de los algoritmos y el temor a quedarse atrás (*fear of missing out*), en comparación a otros países (p. 107).

Más allá de cuál sea el factor determinante de este movimiento y de las intenciones de quienes lo promueven, no podemos perder de vista, como apunta Christian (2020), que nos encontramos en un momento frágil de la historia, en el que el poder y flexibilidad de estos modelos los ha hecho irresistiblemente útiles para una gran cantidad de aplicaciones comerciales y públicas. No obstante, nuestros estándares y normas respecto de cómo usarlos apropiadamente todavía son emergentes. Es exactamente en este período, que deberíamos ser más cuidadosos y conservadores; más aún porque es poco probable que muchos de estos modelos vayan a cambiar sustancialmente después de que sean desplegados en el mundo real (p. 49).

Por eso, vuelvo sobre la necesidad de hacernos responsables de lo que se va a venir, para lo cual es indispensable que nos informemos y promovamos un debate robusto, que parta de la idea de que la IA no es buena ni mala *per se*, sino útil para ciertas tareas y menos útil para otras (Gigerenzer, 2022, p. XV) y que, al tratarse de máquinas, deberíamos procurar usarlas para librarnos de ciertas actividades, de tal manera que podamos dedicarle más tiempo a las preguntas más importantes (Roberts, 2023, 29m10s).

Cierro con estas palabras de Amy Webb (SXSW, 2024) en una presentación que todos los integrantes de la generación T o de la transición¹⁸ (es decir, todos nosotros) deberíamos ver:

Pues sí, el superciclo de la tecnología está aquí. Y sí, va a cambiar el curso de la historia de la humanidad. Pero no tenemos que someternos. No tenemos que renunciar a nuestra agencia. Nos podemos apoyar mutuamente durante esta gran transición. (15m5s)

REFERENCIAS BIBLIOGRÁFICAS

- Ámbito. (2 de febrero de 2023). *Polémica en Colombia con un juez que usó ChatGPT para una sentencia*. <https://www.ambito.com/mundo/inteligencia-artificial/polemica-colombia-un-juez-que-uso-chatgpt-una-sentencia-n5643297>
- Amunátegui, C. (2020). *Arcana Technicae. El derecho y la inteligencia artificial*. Tirant Lo Blanch.
- Ayers J., Poliak A, Dredze M, Leas, E., Zhu, Z., Kelley, J., Faix, D., Goodman, A., Longhurst, C. Hogarth, M. y Smith, D. (2023). Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. *JAMA Internal Medicine*, 183(6), 589-596. <https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/2804309>
- Bean, R. (6 de mayo de 2024). Can AI Be Applied To Revolutionize Healthcare And Medical Outcomes? Forbes. <https://www.forbes.com/sites/andybean/2024/05/06/can-ai-be-applied-to-revolutionize-healthcare-and-medical-outcomes/?sh=366acee621c1>

18 En palabras de Amy Webb (SXSW, 2024), colectivamente estamos pasando por algo trascendental en estos momentos, lo que nos convierte, a todos los que estamos vivos hoy, en la generación T o de la transición, pues somos parte de una gran transición, lo que significa que nuestra sociedad se va a ver muy distinta después de que esta transición haya completado su ciclo (13m56s).

- Brown, B. (Anfitrión). (1 de mayo de 2024). Lisa Gevelber on Technology and AI for Good [Episodio de Podcast]. En *Dare to Lead with Brené Brown*. Spotify. <https://open.spotify.com/episode/2q5yPAOFAw6xumiKNCKsUy?si=iLZgsl-wS9euMyl3CV6CEQ>
- Christian, B. (2020). *The Alignment Problem: Machine Learning and Human Values*. W. W. Norton & Company.
- Constitución Política del Perú. 29 de diciembre de 1993 (Perú).
- Deakin, S. y Markou, C. (2020). From Rule of Law to Legal Singularity en Deakin S. y Markou, C. (Ed.), *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence* (1ª ed., pp. 1-29). Hart Publishing.
- Deloitte (2023). *Generative AI. A guide for corporate legal departments*. Deloitte, <https://www.deloitte.com/content/dam/assets-shared/docs/services/legal/2023/dttl-legal-generative-ai-guide-jun23.pdf>
- Estevez, E., Linares, S. y Fillottrani, P. (2020). *PROMETEA: transformando la administración de justicia con herramientas de inteligencia artificial*. Banco Interamericano de Desarrollo, <https://publications.iadb.org/es/publications/spanish/viewer/PROMETEA-Transformando-la-administracion-de-justicia-con-herramientas-de-inteligencia-artificial.pdf>
- European Commission for the Efficiency of Justice (CEPEJ) (2019). *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*. Council of Europe, <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>
- Gigerenzer, G. (2022). *How to Stay Smart in a Smart World: Why Human Intelligence Still Beats Algorithms*. The MIT Press.
- Han, B. (2021). *No-cosas*. Taurus.
- ISICRI. (11 de octubre de 2021). *Charla: Análisis de los casos comparados en la región: Prometea (Argentina) y PretorIA (Colombia)* [Archivo de Video]. Youtube. <https://youtu.be/XBJ5r5yckzk?si=WxUB2AmwG0rCfXNC&t=2>
- James, J. (2013). State of the Art (A.E.I.O.U.) [Canción]. En *Regions of Light and Sound of God*. ATO Records.
- Kahneman, D., Sibony, O. y Sunstein, C. (2021). *Noise. A Flaw in Human Judgment*. Little, Brown Spark.
- Katsh, E. y Rabinovich-Einy, O. (2017). *Digital Justice. Technology and the Internet of Disputes*. Oxford University Press.
- Laboratorio de Innovación e Inteligencia Artificial de la Universidad de Buenos Aires (IA LAB de la UBA) (2021). *Tucuy Ricuy: lanzamos nuestro quinto sistema*

predictivo [Archivo de Vídeo]. IA LAB de la UBA. https://drive.google.com/file/d/1j4xq_2jTPYQ5-Mzg-xRWltdqIsvjKBQG/view

Laboratorio de Innovación e Inteligencia Artificial de la Universidad de Buenos Aires (IA LAB de la UBA) (6 de agosto de 2020). PretorIA: Inteligencia Artificial predictiva en la Corte Constitucional de Colombia [Archivo de Vídeo]. Youtube. https://youtu.be/kq_N3r2diKw?si=vO9rfzthgEacrUA1&t=2

Ley 31814 de 2023. Ley que promueve el uso de la inteligencia artificial en favor del desarrollo económico y social del país. 5 de junio de 2023. Diario Oficial El Peruano.

Ley de Inteligencia Artificial de la Unión Europea de 2024. 13 de marzo de 2024.

López, J., Becerra, V., Guzmán, M. y Landazuri, J. (2023). Inteligencia artificial en la justicia colombiana: ¿la solución a la congestión judicial? *Revista Lecciones Vitales*, 1(1), 1-11.

LP Pasión por el Derecho. (28 de marzo de 2023). URGENTE: Se dicta la primera sentencia en Perú con asistencia de inteligencia artificial (OPENAI-ChatGPT) [Exp. 00052-2022-18-3002-JP-FC-01]. <https://lpderecho.pe/primera-sentencia-peru-asistencia-inteligencia-artificial/>

Marcus, G. y Davis. E. (2019). *Rebooting AI: Building Artificial Intelligence We Can Trust*. Vintage Books.

Ministry of Justice of the Republic of Estonia. (16 de febrero de 2022). *Estonia does not develop AI Judge*. <https://www.just.ee/en/news/estonia-does-not-develop-ai-judge>

Moran, L. (30 de mayo de 2023). Lawyer cites fake cases generated by ChatGPT in legal brief. Legal Dive. <https://www.legaldive.com/news/chatgpt-fake-legal-cases-generative-ai-hallucinations/651557/>

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) (2022). *Recomendación sobre la ética de la inteligencia artificial*. UNESCO, https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

Organization for Economic Cooperation and Development (OECD) (2019). *Recommendation of the Council on Artificial Intelligence*. OECD, <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>

Pasquale, F. (2020). The Resilient Fragility of Law en Deakin S. y Markou, C. (Ed.), *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence* (1ª ed., pp. v-xvi). Hart Publishing.

Plataforma digital única del Estado Peruano. (24 de agosto de 2021). *Software Tucuy Ricuy permitirá a jueces dictar medidas de protección inmediata en casos de violencia contra la mujer*. <https://www.gob.pe/institucion/pj/noticias/513187-software->

tucuy-ricuy-permitira-a-jueces-dictar-medidas-de-proteccion-inmediata-en-casos-de-violencia-contr-la-mujer

Resolución Administrativa No. 000040-2021-CE/PJ de 2021. Resolución Administrativa que aprueba el Plan de Implementación del Proyecto “Tucuy Ricuy”. https://www.pj.gob.pe/wps/wcm/connect/3eff4f8041aa41489539bd5aa55ef1d3/RESO_LUCI%C3%93N+ADMINISTRATIVA+N%C2%B0+040-2021-CE-PJ.pdf?MOD=AJPERES&CACHEID=3eff4f8041aa41489539bd5aa55ef1d3

Resolución Ministerial N.º 132-2024-PCM de 2024. Proyecto de Reglamento de la Ley 31814, Ley que promueve el uso de la inteligencia artificial en favor del desarrollo económico y social del país. <https://cdn.www.gob.pe/uploads/document/file/6273073/5516872-proyecto-reglamento-ley-31814.pdf?v=1714659628>

Roberts, R. (Anfitrión). (10 de julio de 2023). Marc Andreessen on Why AI Will Save the World [Episodio de Podcast]. En *EconTalk*. Spotify. <https://open.spotify.com/episode/1e90Fp7rNxRAgPGflm6e7b>

Rowghani, A. (Anfitrión). (6 de marzo de 2023). The REAL Potential of Generative AI [Episodio de Podcast]. En *Y Combinator*. Spotify. <https://open.spotify.com/episode/6lR3oCFNS0fXae2RpyYUbB>

Saavedra, V. y Upegui, J. (2021). *PretorIA y la automatización del procesamiento de causas de derechos humanos*. Derechos Digitales, https://www.dejusticia.org/wp-content/uploads/2021/04/CPC_informe_Colombia.pdf

Soltau, S. (13 de junio de 2020). *Detrás de los titulares: dudas sobre los “jueces robot” en China*. Medium. <https://sebastian-soltau.medium.com/detr%C3%A1s-de-los-titulares-dudas-sobre-los-jueces-robot-en-china-d9e197fee2eb>

Susskind, R. (2019). *Online Courts and the Future of Justice*. Oxford University Press. SXS.W. (10 de marzo de 2024). *Amy Webb Launches 2024 Emerging Tech Trend Report | SXS.W 2024* [Archivo de Vídeo]. Youtube. https://youtu.be/5uLSDbh6M_U?si=kv9nZlVlcL1N8zSw&t=2

Task Force for Justice (2019). *Justice for All – Final Report*. New York Center on International Cooperation, https://bf889554-6857-4cfe-8d55-8770007b8841.filesusr.com/ugd/90b3d6_746fc8e4f9404abeb994928d3fe85c9e.pdf

Thomson Reuters (2024). *Future of Professionals Report*. Thomson Reuters, <https://www.thomsonreuters.com/content/dam/ewp-m/documents/thomsonreuters/en/pdf/reports/future-of-professionals-august-2023.pdf>

Venegas, A. (11 de marzo de 2021). Juan Corvalán y PretorIA: “La tecnología a veces es la única solución a la tutela de derechos”. Idealex.press. <https://idealex.press/>

juan-corvalan-y-pretoria-la-tecnologia-a-veces-es-la-unica-solucion-a-la-tutela-de-derechos/

Walker, J. (2019). *On Legal AI*. Full Court Press.

Wolfram, S. (14 de febrero de 2023). *What is ChatGPT Doing... and Why Does it Work?* Writings. <https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/>