LA INTELIGENCIA ARTIFICIAL EN EL PROCESO PENAL COMO ELEMENTO DISRUPTOR DE LA IMPARCIALIDAD, LA TRANSPARENCIA Y LA PRESUNCIÓN DE INOCENCIA

ARTIFICIAL INTELLIGENCE IN THE CRIMINAL PROCEDURE AS A DISRUPTIVE ELEMENT OF IMPARTIALITY, TRANSPARENCY AND PRESUMPTION OF INNOCENCE

Juan Manuel Rosas Caro[1]
Universidad de San Martín de Porres, Perú

Recibido: 20/06/2024 - Aceptado: 19/09/2024

Resumen
La presente investigación busca dilucidar los efectos que la aplicación de la inteligencia artificial tendrá en el proceso penal, por este objetivo, se realiza un análisis con respecto al funcionamiento de la IA y sobre sus capacidades plenas. Finalmente, se procede al análisis de casos de estudio sobre aplicación de la IA en el proceso penal, a partir de esto, se determina el nivel de vulneración que representa el uso irresponsable y con exceso de confianza de la inteligencia artificial para la imparcialidad, transparencia y presunción de inocencia, en el marco del proceso penal.

Palabras clave: Inteligencia artificial; proceso penal; transparencia; imparcialidad; presunción de inocencia.

Abstract
The present investigation seeks to elucidate the effects that the application of artificial intelligence will have on the criminal procedure, for this objective, we carry out an analysis regarding the functioning of AI, and it's full capabilities. Finally, we proceed with the analysis of case studies regarding the application of AI in the criminal procedure, through this, we determine the level of vulneration that represents the irresponsible and overconfident usage of artificial intelligence for impartiality, transparency and presumption of innocence, in the frame of the criminal procedure.

[1] juan_rosas2@usmp.pe. ORCID: https://orcid.org/0009-0009-5422-9496

Keywords: Artificial intelligence; criminal procedure; transparency; impartiality; presumption of innocence.

I. Introducción

El uso de la inteligencia artificial en el ámbito del derecho resulta una inevitabilidad, en la medida en la que la sociedad civil siempre se encuentra en la búsqueda de nuevos métodos y tecnologías que permitan maximizar la eficacia y la eficiencia con la cual se imparte justicia.

Pero la introducción de una tecnología que promete facilitar y automatizar, en una medida considerable, las diversas áreas de la actividad procesal, no puede ser aplicada sin que existan problemáticas y retos con relación a la integridad y observancia de los principios que rigen el proceso penal. El presente artículo abordará múltiples dimensiones del fenómeno de la aplicación de IA generativa a la actividad procesal de la creación de resoluciones judiciales, puesto que es una de las áreas que ha sido sujeta a un fuerte marketing con respecto a las funcionalidades de la inteligencia artificial, siendo considerada la idea de reemplazar a jueces humanos con los mencionados softwares.

Existe fuerte debate con respecto a la viabilidad de estos proyectos de automatización de la función jurisdiccional, siendo que se han creado programas de inteligencia artificial que pueden dar consejos legales, una vez más se debe mencionar que la apariencia de factibilidad de una transición hacia la completa automatización de los servicios legales en general está basada principalmente en una campaña de marketing, en la sobreestimación de las capacidades de la inteligencia artificial y en la generalizada falta de conocimiento con respecto a la verdadera esencia de la tecnología.

Siendo que la postura que se va a manejar en este artículo es que no es posible, en el estado actual de la tecnología, discutir seriamente una automatización de la función jurisdiccional, es decir la generación de sentencias judiciales, en la medida en que tal acto vulneraría múltiples garantías fundamentales con respecto a la imparcialidad, la transparencia y la presunción de inocencia.

De acuerdo con el listado que se ha presentado, estos ejes temáticos serán abordados en relación con la practicidad de la aplicación de la inteligencia artificial generativa como suplemento y complemento de la actividad racional humana en la función jurisdiccional, puesto que el advenimiento de los usos jurídicos de esta automatización tecnológica pueden tener

resultados incluso distópicos, la aplicación de la inteligencia artificial deberá introducirse de forma proporcional a las verdaderas capacidades que tiene, no pretendiendo poner todo el peso de la labor de un sistema judicial funcional sobre una tecnología aún con mucho espacio para desarrollarse.

Es ciertamente, imposible predecir si algún día será posible que una auténtica inteligencia artificial pueda completamente reemplazar a un juez humano, por tanto, inclusive cuando exista la inteligencia artificial plenamente realizada su aplicación deberá ser tomada con pinzas y se deberán tener muy en cuenta los dilemas y obstáculos que presentan su instauración.

II. Objetivo

El objetivo principal de este artículo es examinar críticamente el uso de la inteligencia artificial (IA) en el ámbito jurídico, específicamente en la función jurisdiccional. Se describen las capacidades reales de la IA, desmitificando las exageradas promesas de marketing, y se analizan las problemáticas que surgen al intentar aplicar esta tecnología al sistema de justicia. La argumentación se enfoca particularmente en cómo la implementación de la IA en la generación de sentencias judiciales podría afectar negativamente principios fundamentales del proceso penal, como la imparcialidad, la transparencia y la presunción de inocencia. Además, se busca argumentar que, en el estado actual de la tecnología, no es viable ni deseable una completa automatización de la función jurisdiccional, proponiendo en su lugar un enfoque más equilibrado donde la IA actúe como complemento y no como sustituto de la actividad racional humana en el ámbito judicial.

III. Metodología

El método utilizado corresponde al de una investigación cualitativa, según Olvera García (2015), en tanto se pretende profundizar en las características del fenómeno de la aplicación de la inteligencia artificial en el ámbito jurídico desde la perspectiva del derecho procesal penal y las garantías constitucionales de la administración de justicia penal.

Asimismo, la metodología jurídica usada es la investigación socio jurídica, según Castro Cuba (2019), puesto que se describen las implicancias de un fenómeno tecnológico emergente que afecta la administración de justicia y el funcionamiento del sistema legal. Se analiza de forma crítica y

desde una perspectiva de garantías procesales la implementación de la inteligencia artificial en la función jurisdiccional como fenómeno socio-jurídico, con especial énfasis en sus efectos sobre la imparcialidad, la transparencia y la presunción de inocencia en el proceso penal.

IV. Las capacidades plenas de la inteligencia artificial.

Como preámbulo, es necesario describir cómo funcionan las inteligencias artificiales generativas actuales, en aras de poder determinar en qué medida es posible que apoyen a agilizar la función jurisdiccional, Kurzweil (1994), señala que:

> Los sistemas expertos tienen tres componentes primarios: a) una base de conocimiento estructurada con bases de datos relacionados con los conceptos propios del dominio; b) reglas de decisión que describen los métodos para tomar decisiones en un campo especializado, y c) máquina de inferencia, que también recibe el nombre de motor de inferencia, sistema que aplica las reglas de base de conocimientos a la toma de decisiones y es capaz de conducir el razonamiento para resolver un problema específico. (p. 504)

Esta descripción nos señala a la verdadera naturaleza del funcionamiento de las inteligencias artificiales generativas, siendo que se puede identificar las partes que componen los criterios mecánicos que rigen la producción que este software genera.

Se vuelve evidente que la inteligencia artificial se rige por medio de una serie de directrices que son establecidas como base para su toma de decisiones, el apartado de la “máquina de inferencia” refiere a la capacidad de aplicar estas reglas a una variedad de situaciones de forma automática, lo cual es en puridad la característica de la inteligencia artificial que le da una apariencia milagrosa y “humana” a los materiales que produce la inteligencia artificial.

Otro aspecto fundamental del funcionamiento de la inteligencia artificial es que su base metodológica está compuesta por unidades lógicas, conocidas bajo el nombre de “algoritmo”, en palabras de Nieva Fenoll (2018)

> La palabra clave en inteligencia artificial es “algoritmo”, que sería el esquema ejecutivo de la máquina almacenando todas las opciones de decisión en función

> de los datos que se vayan conociendo. Suelen representarse en los llamados “diagramas de flujo”, que son la descripción básica de ese esquema. (p.21)

Esto nos indica que la base del “pensamiento” de la inteligencia artificial se basa en repetir una serie de opciones que han sido preinstaladas en su software, no se niega la increíble complejidad que estas estructuras pueden llegar a tener, pero es claro que la repetición de una respuesta, que ha sido respondida por una mente humana, pero instalada en un sistema de inteligencia artificial, dista bastante de una máquina hecha y funcional en semejanza a la mente humana.

Con aras a ilustrar mejor el funcionamiento y las limitaciones de la IA, hemos de entender que los algoritmos que componen su pensamiento operativo siempre provienen de una fuente humana, es decir, mientras más amplia sea la base de datos de respuestas preestablecidas que tenga una IA, más complejas serán las situaciones que puedan resolver. Esto por la simple razón de que estará adaptado a una mayor cantidad de potenciales situaciones, pero la cuestión de fondo es que estas respuestas solo son repetidas, mas no creadas por la inteligencia artificial en cuestión.

Pero entonces se generan cuestionamientos en torno a la capacidad de extrapolación de la IA, existen múltiples ejemplos en los cuales la inteligencia artificial ha demostrado aplicar las bases que le han sido instaladas de forma no prevista por los humanos que crearon tales parámetros.

Nieva (2018) señala que, al igual que las emociones humanas, las máquinas también responden a las circunstancias que perciben según los parámetros de generalización que han establecido. Sin embargo, las máquinas actúan con una frialdad escénica, apartándose de lo que consideran peligroso y acercándose a lo que les proporciona protección o información útil. A diferencia de los humanos, las máquinas no necesitan un mecanismo de sorpresa para ajustarse, ya que recolectan y procesan toda la información que perciben por defecto.

Desde una perspectiva filosófica, es posible especular sobre estos temas, aunque Nieva sugiere que esto puede ser más un ejercicio de entretenimiento intelectual que una necesidad práctica. Es importante recordar que, al igual que los humanos, las máquinas también pueden cometer errores al aplicar sus generalizaciones. Ejemplos históricos de estos errores incluyen la censura automática en Facebook de la famosa foto de la niña vietnamita quemada por napalm por considerarla desnuda, y la eliminación de la declaración de independencia de los Estados Unidos al clasificarla como discurso de odio.

Lo previamente dicho nos lleva a analizar las cuestiones más prácticas de la realidad aplicada de la inteligencia artificial, puesto que esta no comparte la misma flexibilidad de la mente humana, es muy complicado hacer que la IA entienda de matices que se producen a partir de la complejidad de la historia y las relaciones humanas.

Esto puede ser visto de dos formas, la primera es desde la incapacidad de la IA para la hipocresía, en el ejemplo citado se menciona que una IA de la empresa Facebook reconoció a la declaración de independencia de los EEUU como un mensaje de odio, y esto se debe a que contiene menciones racistas y coloniales con respecto a los indios americanos. En la actualidad, este tipo de expresiones no son bien vistas (con justa razón), por tanto, a la IA se le equipó con un algoritmo que permite identificar estos discursos de odio y, si observamos el documento histórico, caeremos en cuenta de que efectivamente está presente el uso de un lenguaje insensible.

> El contenido del texto eliminado, que constaba de los párrafos 27-31 de la Declaración, rompía las reglas de la red social por ir en contra de los estándares sobre el discurso de odio, según el aviso que Facebook envió a la página del portal.
> Desafortunadamente, Jefferson, al igual que la mayoría de los colonos británicos de su época, no tenía una visión totalmente amistosa de los nativos americanos.
> Si bien The Vindicator no está seguro de qué fue exactamente lo que desencadenó el programa de filtrado de Facebook, dijeron en una publicación que probablemente fue una parte del escrito que se refería a "salvajes indios despiadados". (Camacho, 2018)

Como se puede advertir, la IA hizo su trabajo correctamente, puesto que se le instaló el algoritmo que le instruye lo que constituye un discurso de odio, siendo que a todas luces las expresiones en el mencionado documento histórico son compatibles con lo que hoy en día categorizaríamos como “intolerante”.

Esto nos señala a una potencial incorruptibilidad de la IA, en el sentido de que si una persona humana hubiera analizado el mismo texto, hubiera tomado en cuenta la necesidad de matizar lo expresado con las conveniencias políticas de la relativización histórica, pero la realidad es que esto nos indica la falta de capacidad que tiene la IA de adaptarse y de unificar criterios, claramente esto es debido a una limitación tecnológica, por tanto

hemos de señalar que la IA no tiene esa capacidad de tomar una decisión tomando en cuenta factores multidimensionales.

Pujol (2022) define el aprendizaje automático como un subcampo de la informática enfocado en programas capaces de aprender de la experiencia, lo que les permite mejorar su rendimiento con el tiempo. En términos generales, se refiere a un conjunto de técnicas que identifican patrones en los datos sin la necesidad de instrucciones explícitas sobre qué buscar. Estos algoritmos buscan desarrollar una aproximación computacional al razonamiento inductivo, utilizando datos o experiencias pasadas para predecir el futuro con un grado de certeza, normalmente empleando el lenguaje de las probabilidades.

Las IAs generativas que están bajo análisis funcionan bajo estos sistemas de aprendizaje automático, como hemos mencionado, estos softwares de IA funcionan en base a criterios que son instalados y que, a su vez, entrenan a la IA con respecto a cómo debe proceder de forma independiente cuando se le presente una serie de datos para analizar, deberá seleccionar aquellos datos en concordancia con el modelo mediante el cual ha sido entrenado.

> Recordemos que para explotar un modelo de aprendizaje automático este tiene que estar previamente entrenado. Es decir, necesitamos reunir un conjunto de documentos y etiquetarlos previamente en función de si son relevantes o no para el caso. Una vez hecho esto, durante la fase de entrenamiento, el algoritmo buscará patrones en los diferentes grupos que hacen que los documentos sean considerados relevantes o no y aprenderá a distinguir entre ambos grupos. (…)
> Uno de los principios en los que se basa el aprendizaje automático es que, dada una gran variedad de casos, esperamos que el algoritmo de aprendizaje extraiga "reglas" y patrones que sean transferibles a nuevos casos. (Pujol, 2022, p. 35)

A partir de esta disquisición podemos confirmar lo dicho previamente, que no existirá capacidad generativa de la IA, sin que se hayan construido criterios y reglas que orienten la producción, este aprendizaje consiste en seleccionar, de forma independiente, los datos correctos de la base que la IA posee, en relación con el "prompt" que se esté solicitando.

Con la finalidad de aterrizar estas ideas en su aplicación para el derecho, un "juez robot" necesariamente tendría que tener instalado una base de datos de jurisprudencia, estilos de redacción y estructuras argumentativas que provengan de jueces

humanos, muy probablemente de una colectividad de ellos; a partir de esta base de datos se deberá entrenar a la IA mediante una muestra adecuada calificación de los datos que están alojados en su base, es decir, que se le enseñará como discriminar entre la gran cantidad de opciones que tiene de acuerdo con los casos concretos que se le presenten. Inclusive yendo más allá de esto, también aprendería a combinar los criterios dentro de su base de datos para entregar un producto que represente una respuesta aparentemente nueva al caso postulado por el "prompt".

Pujol (2022) señala que el objetivo del aprendizaje predictivo supervisado es aprender una función de decisión que permita al algoritmo predecir correctamente las etiquetas de nuevos datos no vistos. Sin embargo, surge la pregunta de cómo se puede confiar en que estas predicciones son realmente precisas, dado que los datos futuros pueden ser muy variados. Esto introduce el concepto de generalización, que se refiere a la capacidad de un método de aprendizaje automático para predecir con precisión datos futuros. Lograr un error de generalización muy bajo es el objetivo final de cualquier tecnología de predicción.

La generalización es, esencialmente, la creación de un sesgo inductivo en el criterio de selección por relevancia de la IA, este sesgo permitirá al sistema circunscribir correctamente cuales datos de su base son pertinentes de acuerdo con el objetivo se demarca a través de la petición.

La correcta delimitación de lo que es llamado frontera de decisión se da mediante la creación de parámetros con las características adecuadas para lograr la selección de datos que se busca que la IA pueda realizar de forma independiente.

> En la primavera de 2014, un ejecutivo de IBM probó un nuevo programa llamado "Debater", un descendiente de Watson que emplea parte de la tecnología del propio Watson aplicada al procesamiento textual para realizar minería de argumentación. Sea cual sea el tópico, la tarea de Debater consiste en "detectar alegatos relevantes" y arrojar "sus predicciones principales sobre alegatos a favor y alegatos en contra". En un ejemplo de resolución, compruébese lo que hizo Debater luego de introducir como insumo el tópico "La venta de videojuegos violentos a menores debería estar prohibida":
>
> 1. Escaneó 4 millones de artículos de Wikipedia,
> 2. Devolvió los 10 artículos más relevantes,
> 3. Escaneó las 3.000 oraciones contenidas en esos 10 artículos,

> 4. Detectó aquellas oraciones que contenían “alegatos posibles”,
> 5. “Identificó limitaciones para los alegatos posibles”,
> 6. “Evaluó la polaridad favorable y desfavorable de los alegatos posibles”,
> 7. “Construyó una maqueta [demo] discursiva con las principales predicciones sobre alegatos”,
> 8. ¡Entonces estuvo “listo para entregar”! (Ashley, 2023, P. 61)

Las IAs basadas en el aprendizaje de modelos, si van a ser aplicados en el ejercicio de la función jurisdiccional deben contar con una frontera de decisión que esté nutrido por la heurística de las decisiones judiciales. En el ejemplo citado, podemos apreciar cómo funciona de forma pormenorizada la “creación” de argumentos por parte de una IA, siendo que se puede colegir que las reglas que utiliza para identificar los datos relevantes son fruto de criterios que han sido enseñados a la IA y que puede poner en práctica de forma independiente, con grados variables de éxito, pero este éxito se verá definido por las características propias de la “frontera de decisión” (Pujol, 2022, p. 41), que permita una adecuada discriminación de los datos.

Entonces, como conclusión de esta sección, podemos señalar que la IA en su estado actual, no es capaz de equiparar la capacidad humana de razonar de forma jurídica, primero, porque requiere de entrenamiento y el valor de los resultados que ofrezca dependen directamente de que tan bien pueda apegarse o aplicar estos criterios humanos que le son enseñados; segundo, puesto que trabaja con bases de datos de decisiones judiciales o modelos de razonamiento creados previamente a partir de mentes humanas, razonablemente, no tendrá la capacidad plena de creación que tienen las personas, por tanto, toda creación de la IA descansará completamente en las opiniones o razonamientos de humanos reales.

El verdadero mérito de la IA yace en su capacidad de analizar con gran rapidez y eficiencia una gran cantidad de datos jurídicos, decisiones judiciales, normas y razonamientos de litigio. Esta tarea de navegar el enorme océano de datos jurídicos que existe le es más sencilla a la IA que a un ser humano, es en mérito a esta realidad que la IA encuentra un uso adecuado como organizador automatizado de información.

Ashley (2023) argumenta que las aplicaciones jurídicas de computación cognitiva, al ayudar a los humanos a enmarcar, probar y evaluar hipótesis jurídicas, se integrarán con ellos en una colaboración paradigmática. Los humanos son más aptos para generar hipótesis interesantes que, si se confirman, tendrán

implicaciones estratégicas o tácticas en las posiciones jurídicas que adopten y en cómo las justifican. Las computadoras, por su parte, pueden analizar rápidamente grandes volúmenes de texto en busca de evidencia relevante para estas hipótesis.

Estas aplicaciones podrían involucrarse con los usuarios en un proceso iterativo de reformulación de hipótesis, tanto a través de términos explícitos como seleccionando ejemplos judiciales que las confirmen o contradigan. La computación cognitiva no ofrecerá una respuesta definitiva, sino conclusiones tentativas que resuman la evidencia a favor o en contra de la hipótesis formulada. Construirá argumentos basados en la evidencia disponible, incluyendo listas de ejemplos que parezcan apoyar la consulta, así como contraejemplos que la contradigan y fallos parciales que solo satisfagan parcialmente los antecedentes normativos de la hipótesis.

En última instancia, los usuarios humanos deberán leer los ejemplos, contraejemplos y fallos parciales seleccionados. No obstante, la aplicación jurídica estructurará su presentación en torno a la hipótesis, enfocando al lector en los documentos más relevantes desde un punto de vista sustantivo. Además, la aplicación resumirá los documentos relevantes de manera que clarifique su relación con la consulta o hipótesis.

Entonces, de esto se puede colegir, que las capacidades plenas de la IA no permiten que actúe de forma independiente de la raza humana, puesto que no cuenta con las capacidades mentales, propias de la humanidad, que permitan crear desligándose de los parámetros que le han sido inculcados. En cambio, la IA es completamente dependiente de dichos criterios, siendo que toda “opinión” que emita está directamente extraída de su base de datos en función de su “frontera de decisión”, como corolario, el futuro inmediato de la IA yace como apoyo a las personas al poder automatizar la discriminación de relevancia de una gran cantidad de datos de naturaleza jurídica, en definitiva, representa una gran ayuda y ahorro de tiempo, que se puede dedicar a aplicar verdadero razonamiento que es tan necesario a la hora de tomar decisiones judiciales y argumentarlas.

V. Peligros con respecto a la imparcialidad en las decisiones judiciales al aplicar de forma irrestricta la IA en la actividad jurisdiccional

Existen registros reales sobre la aplicación de inteligencia artificial, y como es pertinente para la materia bajo análisis, se tiene que su utilización se ha visto afectada por sesgos negativos

que han tenido como repercusión el daño de la posible reputación de imparcialidad que podrían haber cimentado estos sistemas.

> Nos dispusimos a evaluar una de las herramientas comerciales de Northpointe, Inc. Para descubrir la precisión subyacente de su algoritmo de reincidencia y para probar si el algoritmo estaba parcializado contra ciertos grupos.
> Nuestro análisis de la herramienta de Northepointe, llamada COMPAS (que significa Manejo del perfilamiento de delincuente correccional para sanciones alternativas), encontró que los detenidos negros eran mucho más susceptibles que los acusados blancos a ser juzgados incorrectamente de tener un mayor riesgo de reincidencia, mientras que los acusados blancos eran más susceptibles que los acusados negros de ser incorrectamente señalados como de bajo riesgo.
> Revisamos más de 10,000 acusados penales en el condado de BROWARD, Florida, y comparamos su predicción de tasa de reincidencia con la tasa que verdaderamente ocurrió en un periodo de dos años. Cuando la mayoría de los acusados están en la cárcel, ellos responden a un cuestionario de COMPAS. Sus respuestas son alimentadas al software de COMPAS para generar múltiples puntajes incluyendo predicciones de "riesgo de reincidencia" y "riesgo de reincidencia violenta". (Larson et al., 2016)

El funcionamiento de la IA COMPAS señala a un verdadero problema con respecto a la aplicación de estas tecnologías en la forma en cómo se juzga a las personas, inclusive tomando en cuenta sus características personales y sus trasfondos. Esto significa una grave afectación del derecho al juez imparcial, cuyo contenido indica que toda decisión judicial debe verse libre de toda injerencia subjetiva o que proceda de prejuicios que sean propios de la persona que está juzgando, a pesar de que estemos hablando de una inteligencia artificial, podemos apreciar que dentro de su "frontera de decisión" se han instalado criterios humanos de discriminación que contaminan el funcionamiento automático e independiente, generando de forma indirecta una vulneración a las garantías procesales.

Según la Corte interamericana de derechos humanos en la sentencia del caso Barreto Leiva vs. Venezuela:

> La Corte Interamericana ha establecido que la imparcialidad exige que el juez que interviene en una contienda particular se aproxime a los hechos de la causa careciendo, de manera subjetiva, de todo prejuicio y, asimismo, ofreciendo garantías suficientes de índole objetiva que permitan desterrar toda duda que el justiciable o la comunidad puedan albergar respecto de la ausencia de imparcialidad. La imparcialidad personal o subjetiva se presume a menos que exista prueba en contrario. Por su parte, la denominada prueba objetiva consiste en determinar si el juez cuestionado brindó elementos convincentes que permitan eliminar temores legítimos o fundadas sospechas de parcialidad sobre su persona. (2009)

Entonces, a partir de esta definición de los contenidos de la imparcialidad, podemos delimitar que para poder librar a un juzgador de toda sospecha de parcialización y de aplicar un juzgamiento dejándose influenciar por sus prejuicios, se vuelve preciso que exista una observancia de las garantías procesales, en especial en lo que refiere a la adecuada motivación de las resoluciones judiciales, pero cuando observamos el funcionamiento coadyuvado de la IA con los jueces humanos podremos apreciar que existirá una sobre confianza en las capacidades de la IA y, que por ser una máquina, en su imparcialidad.

La existencia de prejuicios racistas utilizados como criterios de juzgamiento al momento de analizar las probabilidades de reincidencia, en el caso que estamos observando, representa un peligro especialmente vigente, puesto que existe una actitud de los juzgadores frente a las predicciones que produce este software, la sobre confianza que se deposita en la IA también representa un fallo que lleva a la ruptura de la imparcialidad, aunque esto sea de forma indirecta.

Pero esta situación nos lleva a preguntarnos como sucede esto, es difícil de creer que un equipo legal y científico que desarrolla una herramienta que en buena medida va a definir la libertad y el perfilamiento de individuos, sea deliberadamente instalada con criterios de frontera de decisión que sean explícitamente racistas, puesto que esto significaría un quiebre del estado de derecho, además que sería algo detectable y que generaría un importante sentimiento de desconfianza en el sistema judicial, creando un ambiente de inseguridad jurídica.

> El perfilamiento racial y la evaluación de riesgo de delito basado en herramientas tradicionales ahora son prácticas

> familiares, y es plausible pensar que el significado social de estas prácticas se extendería inmediatamente al uso de herramientas de ML (Machine learning). La pregunta interesante es si el significado del chivo expiatorio perfecto del perfilamiento (vía el uso de una herramienta de ML que excluye a la raza como un factor predictivo explicito), o a aquello del chivo expiatorio imperfecto del perfilamiento (vía el uso de una herramienta tradicional). Sospechamos que sería algo más cercano a lo anteriormente mencionado.
>
> (…)
>
> Al menos cuando se vuelve claro y ampliamente conocido que una herramienta de ML ha aprendido un chivo expiatorio perfecto para membresía de un grupo racial conocido por haber sido sujeto de opresión sistemática y racialmente motivada, el uso de esa herramienta plausiblemente vendría a expresar un irrespeto distintivamente racial. (Davies & Douglas, 2022, pp. 110-112)

El citado argumento señala que la forma en la que una IA puede tener criterios de selección racistas sucede mediante la implantación de herramientas tradicionales, que representan el racismo común que tienen los humanos en forma de prejuicios sociales. El verdadero problema de fondo es como se identifican estos prejuicios dentro de los criterios instalados en la IA, puesto que no pueden ser dispuestos de forma plenamente explícita.

Por tanto, tenemos los conceptos de chivo expiatorio perfecto e imperfecto, este último refiere a un criterio de selección de la IA que captura ciertos elementos relacionados con el grupo marginado al momento de analizar su posibilidad de reincidencia, pero no llega al punto de considerar todas las características que socialmente son relacionadas con dicho grupo, por esto es considerada imperfecta, el nivel de afectación que tendrá con respecto al resultado (si resulta más gravoso o no que un chivo expiatorio perfecto) es materia de la especificidad de los mismos criterios que aplique.

En cambio, el chivo expiatorio perfecto implica la utilización de criterios por parte de la IA que sean plenamente idénticos a los que socialmente corresponden al grupo social marginado, en la medida en que se consideran estos factores que corresponden de forma ampliamente conocida a dicho sector sin aludir de forma alguna a esta relación, a pesar de que la sociedad podría reconocer dicha conexión si analizara los criterios.

Las herramientas tradicionales a las cuales se refiere el fragmento refieren a las categorías que son analizadas de forma

regular por el juzgador humano, siendo que son conductas o situaciones personales que tienen impacto en la decisión, como puede ser la edad, el trasfondo socio-cultural, historial de abuso de sustancias psicotrópicas, etc. Es razonable pensar que, si estas herramientas ya están contaminadas por un sesgo racial, al ser transferidas a la IA, esta va a desarrollar y extrapolar el mismo sesgo, inclusive de forma que sus criterios de decisión no mencionen alguna motivación discriminatoria.

Existe un peligro añadido con respecto al uso de IA cuyos criterios inevitablemente van a tener sesgos sociales que son inaceptables en la administración de la justicia, el hecho de que los criterios de aprendizaje, una vez que son aprendidos, por las mismas limitaciones de la IA no pueden superarse, y se genera un estado de estancamiento, considerando que los juzgadores tienden a asumir que los resultados producidos por la IA son realidad objetivas y no "hipótesis" posiblemente fallidas, ciertamente genera un potencial de estancamiento del derecho con respecto a criterios sociales que no corresponden con los valores del estado de derecho y que deben ser culturalmente superados, siendo esto posible únicamente mediante la autocrítica que los humanos constantemente hacemos sobre nuestras propias ideologías sociales.

Una constante actualización de los criterios de aprendizaje que informan la frontera de decisión puede ser una solución plausible, pero para que esto pueda considerarse, los juzgadores deben asumir una posición distinta con respecto a la IA, deben considerarla como falible e incluso más influenciable que una mente humana; pero mientras se mantenga una confianza desproporcionada sobre la IA y sus capacidades, difícilmente podremos solucionar de forma satisfactoria los problemas de parcialización que los productos de estos sistemas causan en las decisiones judiciales.

VI. La menor transparencia de las decisiones judiciales a través de algoritmos de inteligencia artificial en contraste con los juzgadores humanos.

La transparencia generalmente es contemplada como efectiva en tanto se respete la publicidad y la inteligibilidad de los procesos y decisiones judiciales. En especial cuando hablamos de los criterios que conforman la validez y legitimidad de las decisiones judiciales, si estos criterios de decisión son arbitrarios y pertenecen al fuero interno del juzgador, entonces la sociedad civil no va a considerar que el sistema judicial sea predecible ni justo, esto se traduce en un rechazo de todo el sistema.

Chiao (2022), señala que la transparencia es un atributo fundamental de cualquier régimen de sentencias legítimo, normalmente garantizado al hacer que las decisiones judiciales sean públicas e inteligibles. Sin embargo, la transparencia se ha convertido en una preocupación creciente en el uso de machine learning en funciones jurisdiccionales. Las herramientas algorítmicas son consideradas menos transparentes que los jueces humanos, tanto porque el código fuente subyacente puede estar protegido como un secreto empresarial, como porque la interpretación de los resultados de predicción generados por los datos de entrada puede ser extremadamente difícil, incluso para expertos. El autor sugiere que, aunque la falta de transparencia en estos aspectos es problemática para las herramientas algorítmicas, también lo es para las decisiones humanas tomadas por jueces. Por lo tanto, las preocupaciones sobre la transparencia tienen un significado ambiguo en los debates sobre el uso de algoritmos en las sentencias.

En el estado actual del ordenamiento jurídico se ha puesto un especial énfasis en reforzar la percepción de transparencia, mediante la obligatoria publicidad de las audiencias y en la necesidad del juez de seguir las reglas de la sana crítica compuesta por máximas de la experiencia, los conocimientos científicos y los principios de la lógica estos son los criterios forzosos bajo los cuales justificar la legitimidad de sus decisiones y argumentos.

Si deseamos lograr la automatización de las decisiones judiciales, entonces indefectiblemente debemos enfrentarnos al problema de la legitimización de estas decisiones bajo los criterios y métodos de juzgamiento que la IA va a aplicar, los cuales no serán los mismo que los juzgadores humanos usan. La legitimidad de una resolución judicial no solo se da un intento de ganar el consentimiento de los administrados, sino que tambien se basa en una necesidad de cohesión del ordenamiento jurídico y de integridad que impida la desnaturalización de los preceptos jurídicos que componen la práctica jurídica.

Está preocupación en torno a la aplicación de la IA, principalmente tiene procedencia a partir de la insondabilidad de los criterios y mecanismos que componen el “pensamiento artificial”, en este trabajo se ha hecho un esfuerzo por simplificar los mecanismos y las formas en las cuales estos sistemas funcionan, pero esto ha sido efectuado de forma general, los criterios específicos de funcionamiento y de frontera de decisión de cada software de IA van a encerrar complejidades individuales que dificultarán el escrutinio de su actividad generativa, esto se vuelve especialmente peligroso para el ámbito judicial, puesto que dificulta cualquier tarea de

dilucidación o verificación de los criterios que se están empleando y que deciden sobre la libertad de los administrados.

Ryberg (2022), expresa que, el uso de algoritmos protegidos legalmente presenta un desafío significativo para la transparencia, como se evidencia en el caso *State v. Loomis*. En este caso, Eric Loomis fue acusado de un ataque a mano armada, y durante su proceso judicial, se utilizó el algoritmo COMPAS para evaluar su riesgo. El juez, basándose en esta evaluación, determinó que Loomis era un individuo de alto riesgo para la comunidad y, por tanto, negó la libertad condicional. Loomis apeló, argumentando que no había tenido la oportunidad de evaluar la precisión del algoritmo COMPAS, lo que planteaba un problema de debido proceso. Sin embargo, la Corte Suprema de Wisconsin rechazó el argumento, señalando que el funcionamiento interno del algoritmo estaba protegido como un secreto empresarial por Northpointe Inc., la empresa que lo comercializa. La corte subrayó que no se revelan los métodos utilizados para calcular los puntajes de riesgo ni cómo se sopesan los factores. A pesar de esto, la Corte Suprema del estado emitió advertencias sobre el uso del algoritmo COMPAS, lo que ha generado un amplio debate sobre la decisión judicial y el uso de algoritmos privados en el proceso de juzgamiento. Esta situación se describe como "opacidad causada legalmente".

En lo que refiere al problema del secretismo industrial con respecto al proceso pormenorizado de la generación de resultados de relevancia para casos penales, es sencillo llegar a la conclusión de que esto es absolutamente aberrante e inaceptable para el estado de derecho. Puesto que es de plano una actitud inquisitiva, rompiendo completamente con todo precepto del modelo acusatorio, puesto que el secreto en el proceso penal de antaño resultó siendo un factor deslegitimador del poder judicial en razón a que la sociedad civil siempre va a considerar una decisión que arrebata la libertad de un ciudadano y que, para mayor inri, se da sin transparencia, o sea en secreto, como una decisión arbitraria, que representa un claro abuso de autoridad.

El argumento jurídico de que el COMPAS es un servicio que el Estado contrata y que su funcionamiento es un secreto industrial debería ser considerado como un llamado de atención con respecto a los peligros de la privatización de la justicia penal. No es admisible que se busque vulnerar un derecho fundamental, como es el debido proceso penal, usando de argumento la primacía del derecho de propiedad, puesto que el secreto industrial es una manifestación de este derecho; si fuéramos a realizar una ponderación entre derechos fundamentales, no cabe duda que en el caso propuesto el

derecho al debido proceso prima por encima del derecho de propiedad, sobre todo porque no se puede apreciar que la transparencia con respecto al funcionamiento del servicio de COMPAS represente una verdadera afectación a los derechos de la empresa, sobre todo porque esa información protegida es necesaria para que el juez pueda cumplir con su función jurisdiccional.

Ryberg (2022), señala que la complejidad de los algoritmos, especialmente en el contexto del aprendizaje profundo dentro del machine learning, es una preocupación significativa en términos de transparencia. Dado que estos algoritmos pueden ser extremadamente complejos, resulta difícil para una persona común, como un procesado, comprender su funcionamiento. Este desafío se denomina a veces como “analfabetismo algorítmico”. Además, incluso los expertos en computación pueden encontrar complicado entender completamente cómo se determina el producto de un algoritmo, lo que dificulta tanto las predicciones de los resultados como las explicaciones de los procesos que conducen a ellos. Por lo tanto, la complejidad de un algoritmo puede considerarse una amenaza a la transparencia, incluso si su fórmula no está protegida como secreto industrial. Este fenómeno puede denominarse "opacidad causada por tecnicidad"

La complejidad de la tecnología representa una dificultad complicada de superar, en especial porque implica que el mundo está cambiando rápidamente y el nivel educativo de los profesionales, así como del hombre medio, no se está dando abasto para poder asegurar la garantía de la transparencia en los procesos judiciales.

La simplificación de los procesos de selección y el funcionamiento pormenorizado de las inteligencias artificiales se convierte en una necesidad para que sea posible su aplicación en el ámbito judicial, si bien la simplificación técnica propiamente dicha es imposible y en realidad para poder aumentar las capacidades de la IA, efectivamente se generará un aumento de la complejidad de los sistemas que le dan su potencialidad a los sistemas de IA.

Entonces, es necesario encontrar otra forma de hacer que los datos con respecto al funcionamiento de IA sean más transmisibles y entendibles para el público en general, no especializado o familiarizado con el funcionamiento de estas tecnologías. Este objetivo se puede lograr mediante un esfuerzo de adaptación de los términos técnicos, en la literatura especializada siempre, los autores, se sirven de analogías para poder transmitir de forma satisfactoria al lector la forma en la que estos sistemas de IA funcionan, es posible trazar paralelos

con el derecho, puesto que ambos temas sufren de una manifiesta predisposición por el oscurantismo, siendo que reina el uso de tecnicismo y de lenguaje insondable para las personas que no están inmersas en el ejercicio de estas disciplinas.

La literatura especializada, los propios análisis jurídicos del funcionamiento de la IA y sus efectos en el derecho son la vanguardia de este movimiento de "alfabetización" con respecto a los pormenores del funcionamiento de la IA y debemos mantener esta tendencia, puesto que estos esfuerzos académicos hacen que un fenómeno tecnológico como es la inteligencia artificial -que ha sido propenso a ser mitificado por la sociedad civil y el hombre medio- se vuelva parte de obligatoria y más sencilla comprensión para el mundo del derecho, se ha de concluir que será mediante estos esfuerzos de producción académica que será posible superar este reto de opacidad por tecnicidad y se obtendrá un pleno funcionamiento transparente de la IA en el derecho.

VII. Los retos de la inteligencia artificial frente al principio de presunción de inocencia

El uso de la inteligencia artificial en la videovigilancia masiva es un tópico que se aleja un poco del marco de la IA generativa, pero definitivamente tiene ramificaciones que son imposibles de ignorar. Por tanto, previamente hemos de analizar cuál es el rol de la inteligencia artificial y sus algoritmos aplicados en la videovigilancia, siendo que genera un efecto de perfilamiento en tiempo real, que al igual que en el caso del COMPAS, que tiene como consecuencia una ruptura de la presunción de inocencia, puesto que se efectúan arrestos basándose en las características personales del individuo por su "peligrosidad".

La relevancia actual del problema que presenta la inteligencia artificial para la presunción de inocencia se ve evidenciada por su reciente y notoria aplicación en una forma trasngresora de dicho principio, según *The Next Wave*,

> La policía metropolitana confirmó que utilizaría tecnología de vigilancia masiva de reconocimiento facial en tiempo real, que se utiliza en conjunto con un software de inteligencia artificial que analiza los datos biométricos e identifica rostros que sean compatibles con individuos potencialmente peligrosos. (2023)

Además, mientras se desplegaba esta operación, según *CNN,* "la policía Metropolitana dijo que realizó 52 arrestos durante la coronación del Rey Charles III el sábado, mientras las fuerzas enfrentan un escrutinio en aumento por su actitud contra protestantes antimonárquicos" (2023)

De acuerdo con una declaración en el London Assembly[2], el 18 de Julio, "el reconocimiento facial en tiempo real fue usado como parte de una compleja operación que ayudo a desincentivar la criminalidad y mejorar la seguridad pública (durante la coronación)". (2023).

La importancia de hablar con respecto a la videovigilancia masiva se vuelve evidente, puesto que los datos generados mediante el uso de un reconocimiento facial y biométrico realizado mediante inteligencia artificial, tiene implicancias directas con el posterior juzgamiento que se dará con respecto a una causa penal que está iniciando permeada por una flagrante presunción de culpabilidad la cual se genera a través de un uso incorrecto de la inteligencia artificial, una vez más podemos señalar a la sobre confianza en estos softwares por parte de las fuerzas de orden y de los jueces con respecto a la objetividad de los "productos" de la IA.

Mendola (2016) discute cómo el software de predicción, al ser implementado masivamente, puede establecer factores de riesgo que resultan en efectos desproporcionadamente negativos para ciertos grupos y comunidades. Existen numerosos ejemplos en los que el análisis de datos y la aplicación de la ley basada en estos datos se perciben como discriminatorios, afectando desproporcionadamente a algunos sectores. En los Estados Unidos, por ejemplo, la adopción del programa COMPSTAT, diseñado para analizar y reducir la actividad criminal en Nueva York, ha proporcionado un apoyo significativo a la policía. Sin embargo, también se ha observado que la intervención policial está desproporcionadamente dirigida hacia ciertas comunidades, principalmente personas de color, pobres y residentes de suburbios. En el contexto del minado de datos, algunos intelectuales han identificado tres tipos de discriminación: la primera se relaciona con intenciones conscientes de perjudicar, incluso a miembros de la alta sociedad, que son más difíciles de detectar; la segunda se centra en los problemas del proceso de minado de datos en sí, como errores del sistema que podrían evitarse; y la tercera se refiere a los resultados indeseados cuando el minado de datos aumenta los poderes de discernimiento de los tomadores de decisiones, perpetuando la desigualdad incluso en ausencia de prejuicio, parcialidad o error.

[2] Los miembros de la asamblea de Londres investigan asuntos que afectan a Londres y realizan 10 preguntas al año al alcalde.

Esta conceptualización es propia de la actividad policial preventiva, que tiene el funcionamiento descrito en el evento de la coronación del rey Carlos III, el software de IA escanea las características de las personas y aplica su algoritmo de selección, siguiendo criterios que son un misterios sin regular, a partir de la predicción de que un individuo es "peligroso" y susceptible de cometer un delito, esto se toma como un hecho objetivo, procediéndose a aplicar el Ius puniendi de forma irrestricta, efectivamente creando una situación de absurdo adelantamiento de las barreras punitivas y criminalizando la personalidad de los individuos.

Según la *CNN* (2023), las autoridades policiales de Londres advirtieron que adoptarían una postura estricta frente a cualquier alteración del orden público durante las celebraciones, incluyendo manifestaciones. La legislación británica el "police, crime, sentencing and courts act" de 2022, amplió las facultades policiales para regular las protestas, tipificando como delito la provocación deliberada de molestias públicas por parte de los manifestantes.

Siendo que dichas operaciones de videovigilancia con subsecuentes arrestos, se llevó a cabo en el contexto de la vigencia de una normativa que expande el criterio de lesividad, para comprender conductas de protesta que no eran punibles previamente. Esta situación legislativa se coadyuva con la predisposición de asumir un valor probatorio desproporcionado a los resultados de la IA, haciendo muy posible que se incurra en una situación similar a la de la aplicación del software COMPAS, con la particularidad de representar una afectación a la presunción de inocencia.

En relación con el "police, crime, sentencing and court act", se tiene que la mera participación en actos de protesta por parte de individuos que hayan sido identificados como disruptores del orden implica susceptibilidad de ser procesados, según la CNN (2023), por delitos como infracción de la paz y conspiración para causar molestia pública. Lo cual constituye una situación de taxatividad dudosa, vulnerando la garantía de lex certa, criminalizando la peligrosidad de los individuos, apoyándose del perfilamiento en tiempo real que ofrece la Inteligencia Artificial para dar asidero a un procesamiento penal por las características personales del individuo.

El peligro a la presunción de inocencia se configura con mayor urgencia por la fuerte posibilidad de falsos positivos otorgados por las inteligencias artificiales usadas para determinar la peligrosidad de los individuos que están bajo videovigilancia masiva.

El aplicar las detenciones y potencialmente el Ius puniendi usando como única, o principal, justificación la predicción de un sistema de IA cuyos criterios pueden tener sesgos discriminatorios, cuyo funcionamiento pormenorizado está oculto por secreto industrial o por el oscurantismo de los tecnicismos o que, inclusive, puede tener un fallo en su frontera de decisión, llevando a que de un falso positivo al interpretar de forma errónea la selección de datos, es claramente una afectación a la presunción de inocencia, puesto que por lo mencionado, la predicción de tal sistema no genera el acervo probatorio con la exigencia necesaria como para enervar la presunción de inocencia.

Asumir que la predicción otorgada por este sistema es suficiente para efectuar un arresto es un acto peligrosamente cercano a la presunción de culpabilidad y considerando que este acto implica el comienzo de un proceso penal, como punto de partida se estaría ignorando una garantía constitucional propia del proceso penal al operar bajo dicha presunción de culpabilidad que se sobreentiende por tan pronto ejercicio del Ius puniendi.

VIII. Conclusiones

La inteligencia artificial generativa, y de otros tipos, se ve limitada por los criterios bajo los cuales ha sido entrenado, es decir, que sus decisiones independientes se verán definidas por su “frontera de decisión”, el cual es un concepto que refiere a las capacidades de selección que tiene la IA de acuerdo con los criterios que le han sido instalados.

Es en vista a esto que las capacidades plenas de la IA no son plenamente autónomas, puesto que lo que se “genera” proviene de un parafraseo y reorganización de materiales de la base de datos de la IA, cuyo proceso de selección se ve delimitado por la frontera de decisión que funciona en base a algoritmos, los cuales funcionan como proposiciones lógicas.

La imparcialidad de la IA es un tema muy delicado que implica el análisis de los criterios que informan su frontera de decisión, mediante los chivos expiatorios perfectos e imperfectos podremos dilucidar cuando las categorías que componen los parámetros de análisis de la IA son realmente discriminativas o buscan solapar la aplicación de algún sesgo prejuicioso durante el proceso de juzgamiento, puesto que estos chivos expiatorios refieren a los criterios que son compatibles y están directamente vinculados a los grupos que están

desventajados por el sesgo de la IA, este sesgo se aplica sin alguna mención de perfilamiento explícito.

La transparencia del uso de la IA en la función jurisdiccional sugiere la cuestión de dos retos evidentes, la falta de transparencia por secreto industrial y la falta de transparencia por analfabetismo técnico. El secreto industrial como protección del derecho de propiedad se ve inmediatamente superado en hipotética ponderación con respecto al derecho al debido proceso, por lo que no existe cabida para una falta de transparencia por esa razón en nuestro ordenamiento jurídico; la otra razón se debe a una falta de familiarización con la terminología y los pormenores de la tecnología, lo cual ha de ser solucionado mediante la simplificación de estos contenidos altamente técnicos, tarea que recae en el sector académico jurídico, la cual se está llevando a cabo de forma satisfactoria, pues la literatura especializada sirve al propósito de explicar de forma más entendible todo lo relacionado con el funcionamiento de la IA y su relevancia para el proceso penal.

Finalmente, tenemos la relación tensa entre la aplicación de la IA en el proceso penal y la presunción de inocencia, como hemos repasado, la utilización de IA predictiva sobre la ciudadanía ha generado aplicaciones irrestrictas del ius puniendi, efectivamente creando una situación de presunción de culpabilidad al tomar como un dato objetivo la predicción de una IA, esto es definitivamente algo inaceptable, al comprender que estos sistemas se ven definidos por criterios humanos que son falibles y que representan “hipótesis”, en contraste con el tratamiento más común que es tomar los productos de la IA como hechos objetivos, esto es una forma de pensar que debe ser extirpada del pensamiento jurídico, puesto que resulta en flagrantes vulneraciones a la presunción de inocencia.

Bibliografía

Ashley, K. (2023). Inteligencia artificial y analítica jurídica nuevas herramientas para la práctica del derecho en la era digital. Editorial Yachay legal; Editorial PUCP.

Camacho, L. (2018, 3 de agosto). ¿Hay 'discurso de odio' en la declaración de Independencia de EE. UU.? *El Tiempo.* https://www.eltiempo.com/tecnosfera/novedades-tecnologia/facebook-encuentra-racista-al-documento-de-independencia-de-ee-uu-239594

Caso Barreto Leiva vs. Venezuela. (2009). Corte interamericana de derechos humanos. https://www.corteidh.or.cr/docs/casos/articulos/seriec_206_esp1.pdf

Castro Cuba Barineza, I. (2019). Investigar en Derecho. Editorial de la Universidad Andina de Cusco.

Chiao, V. (2022). Transparency at sentencing, are human judges more trasnparent than algorithms? En J. Ryberg & J. Roberts (editores), Sentencing and artificial intelligence (pp. 34-56). Oxford university press.

Davies, B. & Douglas, T. (2022). Learning to discriminate: the perfect proxy problem in artificially intelligent sentencing. En J. Ryberg & J. Roberts (editores), Sentencing and artificial intelligence (pp. 97-121). Oxford university press.

Geschwindt, S. (2023, 9 de mayo). *Controversial AI tech deployed at King's coronation. The Next Wave*. Recuperado de: https://thenextweb.com/news/kings-coronation-controversial-ai-tech-deployed-alongside-record-setting-5g-network

Kennedy, N., Edwards, C., Isaac, L. & Goodwin, A. (2023, 6 de mayo). *'Something out of a police state': Anti-monarchy protesters arrested ahead of King Charles' coronation. CNN*. Recuperado de: *https://edition.cnn.com/2023/05/06/uk/king-charles-anti-monarchy-protest-arrests-ckc-gbr-intl/index.html*

Kurzweil, R. (1994). La era de las máquinas inteligentes, México, CONACYT/Equipo sirius mexicana.

Larson, J., Mattu, S., Kirchner, L. & Angwin, J. (2016, 23 de mayo). How We Analyzed the COMPAS Recidivism Algorithm. *Propublica.* https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm

London Assembly. (2023). *Facial Recognition Technology at the Coronation.* Recuperado de*: https://www.london.gov.uk/who-we-are/what-london-*

assembly-does/questions-mayor/find-an-answer/facial-recognition-technology-coronation

Mendola, M. (2016) One step further in the surveillance society: the case of predictive policing. Tech and Law Center.

Nieva Fenoll, J. (2018). Inteligencia artificial y proceso judicial. Editorial Marcial Pons.

Olivera García, J. (2015). Metodología de la investigación jurídica: para la investigación y la elaboración de tesis de licenciatura y posgrado. Editorial UAEM & Maporrua.

Pujol Vila, O. (2022). Aspectos tecnológicos y empresariales de la inteligencia artificial. En P. García (director), Claves de Inteligencia Artificial y derecho (pp. 15-59). Editorial La Ley.

Ryberg, J. (2022). Sentencing and algorithmic transparency. En J. Ryberg & J. Roberts (editores), Sentencing and artificial intelligence (pp. 13-33). Oxford university press.