



**Universidad
Europea**

PROYECTO FIN DE GRADO

**Los desafíos en la implementación de la
inteligencia artificial en la criminología en
Europa**

AUTOR: Víctor Arnedo Lozano

**DIRECTOR:
Álvaro Falcón Pulido**

**GRADO EN CRIMINOLOGÍA
FACULTAD DE LAS CIENCIAS SOCIALES Y DE LA
COMUNICACIÓN**

UNIVERSIDAD EUROPEA

Víctor Arnedo Lozano

**Los desafíos en la implementación de la inteligencia artificial
en la criminología en Europa**

UNIVERSIDAD EUROPEA

**FACULTAD DE CIENCIAS SOCIALES Y DE
COMUNICACIÓN**

TUTOR: ÁLVARO FALCÓN PULIDO

Madrid, a 8 de julio de 2024

Dedicatoria

Me gustaría dedicar este trabajo, primero, a mi familia, tanto a mis padres y a mis hermanos, que han estado durante este largo recorrido a mi lado día a día, y que me han estado apoyando en todo momento.

A la otra parte de mi familia, a mis primas y primos, tías y tíos, y abuelas y abuelos, que se han preocupado por mí y se han ofrecido a ayudarme en lo que pudiesen.

Tampoco me puedo olvidar de mis compañeros de clase, con los cuales he recorrido este camino, compartiendo tiempo y trabajos, ayudándome tanto fuera como dentro de la universidad.

Y, por último, a los profesores que he tenido a lo largo de la carrera, ya que me han transmitido su conocimiento y me han ofrecido su ayuda en todo momento.

Agradecimientos

Me gustaría agradecer este trabajo, primero a mis padres, ya que, sin ellos, yo no habría sido capaz de estar aquí y de poder realizar la carrera de mis sueños, por lo que les estoy muy agradecido por todo lo que han hecho por mi hasta ahora.

Segundo, a mi tutor del trabajo de fin de grado, Álvaro Falcón Pulido, por aconsejarme y guiarme durante este trabajo que tanto he disfrutado hacer.

A mi grupo de amigos del colegio, ya que han sido muchos años los que llevamos juntos, y que me han apoyado en todas las decisiones que he tomado, preguntando e interesándose por mi carrera y por mi progreso.

Por último, a mi grupo de amigos de la universidad, ya que con ellos he compartido recuerdos que espero recordar con nostalgia en el futuro.

Y, en especial, me gustaría agradecerle a una amistad que he hecho durante la carrera, que me ha ayudado tanto en lo académico, como en lo personal. A su lado, he conseguido crecer y madurar como persona, gracias a su apoyo incondicional y su forma de ver las cosas, por lo que estoy profundamente agradecido con ella.

Resumen

La inteligencia artificial ha surgido como un nuevo fenómeno revolucionario que se está incorporando en todos los ámbitos. Desde la medicina, hasta la educación, la implementación de la inteligencia artificial es una realidad que cada día se expande más. Pero al igual que se usa para el bien común, también puede ser usado por delincuentes para mejorar la eficacia y el rendimiento de ciertos delitos, al igual que abre la puerta a nuevos delitos. Sin embargo, este trabajo se centrará en su implementación en la criminología, tanto en el ámbito policial como en el judicial, la cual destaca en la predicción de delitos. Aunque su regulación es bastante escasa, hay algunas guías éticas sobre su uso para reducir los problemas legales con las futuras leyes. Se estudiará el caso de *Wisconsin v. Loomis*, en el cual uso como herramienta de apoyo una evaluación de una IA para la toma de decisiones. La conclusión que sacamos después de estudiar el caso es que el mayor desafío que conlleva el uso de inteligencia artificial en sentencias es la propia transparencia de ésta, que deja en un estado de indefensión al acusado al no saber cómo se ha llegado a ese resultado.

Palabras clave: Inteligencia artificial, implementación, criminología, predicción de delitos, *Wisconsin v. Loomis*

Abstract

Artificial intelligence has emerged as a revolutionary new phenomenon that is being incorporated into all fields. From medicine to education, the implementation of artificial intelligence is an ever-expanding reality. But just as it is used for the common good, it can also be used by criminals to improve the efficiency and performance of certain crimes, as well as opening the door to new crimes. However, this paper will focus on its implementation in criminology, both in law enforcement and in the judiciary, which excels in crime prediction. Although its regulation is rather scarce, there are some ethical guidelines on its use to reduce legal problems with future laws. The case of *Wisconsin v. Loomis* will be studied, in which an AI evaluation was used as a support tool for decision making by the judge. The conclusion we draw after studying the case is that the biggest challenge involved in the use of artificial intelligence in sentencing is the transparency of the AI itself, which leaves the defendant in a state of defencelessness as they do not know how the result has been reached.

Keywords: Artificial intelligence, implementation, criminology, crime prediction, *Wisconsin v. Loomis*

ÍNDICE

1. INTRODUCCIÓN	11
1.1 PROBLEMA DE INVESTIGACIÓN	12
1.2 PREGUNTA DE INVESTIGACIÓN	13
1.3 OBJETIVOS	13
1.3.1 OBJETIVO GENERAL	13
1.3.2 OBJETIVOS ESPECÍFICOS	13
1.4 JUSTIFICACIÓN: LA RELEVANCIA, ORIGINALIDAD Y CONTRIBUCIÓN CIENTÍFICA AL CONOCIMIENTO ACADÉMICO	14
2. FUNDAMENTACIÓN TEÓRICA	15
2.1 INTRODUCCIÓN A LA INTELIGENCIA ARTIFICIAL	15
2.1.1 ¿QUÉ ES LA INTELIGENCIA ARTIFICIAL?	15
2.1.2 FUNDAMENTOS DE LA INTELIGENCIA ARTIFICIAL: APRENDIZAJE AUTOMÁTICO Y PROCESAMIENTO DE DATOS	18
2.2 APLICACIONES DE LA INTELIGENCIA ARTIFICIAL EN LA CRIMINOLOGÍA	20
2.2.1 USOS CRIMINALES DE LA INTELIGENCIA ARTIFICIAL	24
2.3 ÉTICA Y DESAFÍOS DE LA INTELIGENCIA ARTIFICIAL	26
2.3.1 CONSIDERACIONES ÉTICAS	26
2.3.2 DESAFÍOS Y RIESGOS DE LA IA	30
2.4 MARCO LEGAL DE LA INTELIGENCIA ARTIFICIAL EN EUROPA	32
2.5 TEORÍAS CRIMINOLÓGICAS	36
2.6 FORMULACIÓN DE HIPÓTESIS: RESULTADOS ESPERADOS	37
3. METODOLOGÍA DE INVESTIGACIÓN	38

3.1 METODOLOGÍA.....	38
3.2 CONSIDERACIONES ÉTICAS	39
3.3 LIMITACIONES DEL ESTUDIO	40
3.4 CONTRASTE DE HIPÓTESIS	40
4. ANÁLISIS DEL CASO	42
5. CONCLUSIONES	51
5.1 LA AMPLITUD Y LIMITACIONES DE LA INVESTIGACIÓN.....	53
5.2 FUTURAS LÍNEAS DE INVESTIGACIÓN	54
6. REFERENCIAS BIBLIOGRÁFICAS	55

Índice de Siglas y Abreviaturas

Sigla	Inglés	Español
IA	Artificial intelligence	Inteligencia artificial
ML	Machine Learning	Aprendizaje automático
DL	Deep Learning	Aprendizaje profundo
ADN	Deoxyribonucleic acid	Ácido desoxirribonucleico
CCTV	Close Circuit Television	Circuito Cerrado de Televisión
SPOA	Oral Accusatory Penal System	Sistema Penal Oral Acusatorio
IBM	International Business Machines	Máquina de Negocios Internacionales
FGN	Attorney General's Office	Fiscalía General de la Nación
COMPAS	Correctional Offender Management Profiling for Alternative Sanctions	Perfilación de condenados en centros penitenciarios para sanciones alternativas
UNICRI	United Nations Interregional Crime and Justice Research Institute	Instituto Interregional de las Naciones Unidas para Investigaciones sobre la Delincuencia y la Justicia
UE	European Union	Unión Europea
HLEG	High Level Expert Group	Grupo de expertos de alto nivel
ONU	United Nations	Organización de las Naciones Unidas
UNESCO	United Nations Educational, Scientific and Cultural Organization	Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura

OCDE	Organization for Economic Co-operation and Development	Organización para la Cooperación y el Desarrollo Económicos
ISO	International Organization for Standardization	Organización Internacional para la Estandarización
PSI	Presentence Investigation Report	Investigación Previa a la Sentencia

1. INTRODUCCIÓN

En el último siglo, la tecnología ha revolucionado por completo la vida del ser humano, desde las relaciones sociales hasta el sector laboral. En los últimos tiempos, ha surgido un sistema que tendrá un impacto significativo en el mundo: el de las inteligencias artificiales.

Las inteligencias artificiales se alzan como la mejor herramienta para tratar y procesar toda la información actual, la cual hoy en día pueden resultar hasta cierto punto excesiva, y aun así obtener un resultado que satisfaga las necesidades del usuario.

Son muchos los campos donde se ha implementado ya esta nueva tecnología, desde en la educación, hasta la ciberseguridad, y en la medicina. Pero no es necesario irse a hospitales o grandes empresas para hacer uso de la IA, ya que la tenemos más cerca de lo que pensamos. Según Pega (2018) solo el 33% sabe realmente que está usando la inteligencia artificial en su día a día, y el 77% de nosotros usamos un dispositivo el cual usa inteligencia artificial.

Aunque estos son solo algunos ejemplos donde ya se utiliza inteligencia artificial, en este trabajo nos centraremos en el uso que se le da a la inteligencia artificial en la criminología. Desde la prevención de delitos, a ser una herramienta usada por jueces para ayudarse en sus sentencias, la inteligencia artificial se abre camino por muchas ramas.

Pero, al igual que los Cuerpos y Fuerzas de Seguridad del Estado de todos los países pueden usar esta herramienta, los delincuentes también pueden hacerlo, lo que abre una nueva ventana para crímenes nunca vistos.

Mientras se buscan nuevas formas de combatir y controlar los ciberdelitos, surge un nuevo problema más difícil de controlar: la suplantación de identidad a partir de fotos que se encuentran en las redes sociales, hasta estafas utilizando la voz de una persona que físicamente se encuentra a salvo, desnudos de personas a partir de una foto, creación de canciones con la voz de famosos (incluso fallecidos) que no han cantado, o poner palabras en la boca de un famoso que nunca ha dicho.

Este estudio propone varios enfoques para abordar los desafíos de la inteligencia artificial, su regulación en Europa y su consideración como una nueva forma de criminalidad.

1.1 PROBLEMA DE INVESTIGACIÓN

El problema que se tratará en esta investigación surge debido a las incógnitas que surgen con la aplicación de la inteligencia artificial en ámbito de la criminología, el policial, y el judicial.

La falta de regulación de la inteligencia artificial deja vacíos legales, los cuales todavía no han sido resueltos completamente, debido a la que la aparición de esta herramienta ha sido muy reciente, y los legisladores no han tenido tiempo para regularlo debidamente.

De esta manera surgen nuevos delitos, entre otros, Cabrales (2024) apunta que la inteligencia artificial se está usando para la discriminación, la violación de derechos de propiedad intelectual, el acoso cibernético, las estafas, los fraudes, el robo o suplantación de identidad, así como la propagación de virus y malware.

Por otro lado, como transmite Ulnicane (2022), existe una preocupación por la protección de los derechos humanos. En la predicción de delitos, debido a que como trata con datos generales, se teme que la IA caiga en la estigmatización de los individuos, lo cual puede llevarlos a un proceso de estereotipia errónea, y, por tanto, crear una desigualdad entre las personas. Así como en las sentencias, pudiéndose dictar sentencias en base a datos generales, y no en base al individuo.

Aunque en Europa no se ha implementado todavía lo suficiente, en Estados Unidos sí, por lo que será objeto de estudio para comprobar que problemas has surgido y qué soluciones se han propuesto.

1.2 PREGUNTA DE INVESTIGACIÓN

¿Qué desafíos conlleva la implementación de la inteligencia artificial en el campo de la criminología en Europa en la actualidad?

1.3 OBJETIVOS

1.3.1 Objetivo general

Identificar los desafíos que conlleva la implementación de la inteligencia artificial en el campo de la criminología en Europa en la actualidad.

1.3.2 Objetivos específicos

- Definir qué es la inteligencia artificial y sus fundamentos.
- Determinar las principales aplicaciones en el campo de la criminología.
- Evaluar los desafíos éticos y legales que conlleva el uso de la inteligencia artificial.
- Examinar el marco legal existente en Europa que regule la inteligencia artificial.
- Analizar cómo la implementación de la inteligencia artificial como herramienta de apoyo en una sentencia afecta los derechos del acusado.

1.4 JUSTIFICACIÓN: LA RELEVANCIA, ORIGINALIDAD Y CONTRIBUCIÓN CIENTÍFICA AL CONOCIMIENTO ACADÉMICO

Como hemos observado, existe una preocupación respecto al fenómeno de la inteligencia artificial, tanto con su uso indebido, como con los desafíos y los problemas que conlleva respecto a los derechos de los ciudadanos.

A pesar de suponer un obstáculo para los derechos fundamentales y para la seguridad de las personas, no se han realizado tantos estudios respecto al tema. De ese modo, hago énfasis en la relevancia de este trabajo, cuya intención es contribuir al conocimiento del área en cuestión, mejorando la educación y formación de aquellos que pretendan investigar este mismo fenómeno.

Por otro lado, el análisis crítico de los problemas puede ayudar a identificar áreas de mejora y atención específicas. Esto es fundamental para el progreso continuo hacia prácticas y tecnologías más seguras y eficientes. La investigación puede servir como base para mejorar el marco legal y regulatorio relacionado con el uso de IA en la criminología. Esto podría llevar a leyes y políticas más fuertes que protejan mejor los derechos de las personas y garanticen que la tecnología se use de manera ética.

Así mismo, se espera que este trabajo tenga su impacto tanto en la política, como en lo ético y social. Ya que al explorar las implicaciones que tiene su uso en la criminología, puede influenciar en la creación de nuevas leyes que ayuden a regular este fenómeno y a reducir las consecuencias de su mal uso, entre los cuales se encuentran, como hemos mencionado con anterioridad, en el impacto que tiene en los derechos fundamentales de una persona acusada,

Al combinar enfoques y conocimientos de diversas áreas, este trabajo de investigación puede fomentar el avance en múltiples disciplinas. En este caso, en la criminología y la IA se pueden incorporar elementos del derecho, la ética, la sociología, la informática y la psicología, lo que permite una comprensión más completa del problema.

2. FUNDAMENTACIÓN TEÓRICA

2.1 Introducción a la inteligencia artificial

2.1.1 ¿Qué es la inteligencia artificial?

El primer problema al que nos enfrentamos a la hora de definir lo que es la inteligencia artificial que no existe una única definición de inteligencia artificial como mencionan Zarina et al. (2019). Son muchos los autores que intentan definir esta rama, y aunque parten de una misma base, los conceptos en algunos autores cambian y muchos otros son definidos más a fondo o con un lenguaje más técnico.

Siguiendo a Manning (2020), fue el profesor Stanford John McCarthy el que acuñó el término Inteligencia artificial por primera vez en 1955, definiéndolo como “la ciencia y la ingeniería capaz de crear máquinas inteligentes”. También dio una definición de lo que es la inteligencia, entendiéndola cómo la capacidad de aprender y ejecutar técnicas adecuadas para resolver problemas y alcanzar objetivos, adecuados al contexto en un mundo incierto y en constante cambio.)

Con esta definición de inteligencia, podemos descartar aquellos “robots” que hayan sido preprogramados, porque son flexibles, consistentes, y precisos, pero no inteligentes, ya que las órdenes ya las tenían impuestas, por lo que no son capaces de aprender, y no podrían adaptarse cuando suceda algún problema. Delimitamos así lo que entendemos por inteligencia artificial separándolo de máquinas o programas automáticos o automatizados.

Sin embargo, esta definición queda un poco obsoleta a día de hoy, por lo que a continuación se compararán diferentes definiciones para mostrar cómo ha ido evolucionando el concepto de inteligencia artificial.

Para seguir completando la definición de inteligencia artificial, desde una perspectiva más actual e internacional, podemos acudir a la Comisión Europea. Según la Comisión Europea (2018) la inteligencia artificial se define como aquellos sistemas que muestran un comportamiento inteligente, que son capaces de analizar su entorno, y que emprenden acciones de manera independiente para lograr objetivos específicos. Los sistemas que incorporan IA pueden funcionar únicamente con software, como asistentes

de voz, software de análisis de imágenes, motores de búsqueda, sistemas de reconocimiento facial y de voz.; o la IA puede estar integrada en dispositivos de hardware, por ejemplo, robots avanzados, coches autónomos, drones o aplicaciones de “Internet of Things”.

El término “Internet of Things” usado en la definición de la Comisión Europea es importante definirlo, debido a que engloba al contexto actual y que es parte de nuestro día a día. Son Tan & Wang (2010) los que definen el término "Internet de las cosas" como la forma en que los dispositivos físicos, los vehículos, los electrodomésticos y otros objetos están conectados entre sí a través de sensores, software y tecnología de red que les permite recopilar y compartir datos. Estos dispositivos pueden interactuar con otros sistemas a través de Internet, lo que facilita la automatización de tareas, el control remoto y la recopilación de datos.

La relevancia de tener una definición de inteligencia artificial de la Comisión Europea radica en su función como entidad reguladora y promotora de políticas a nivel europeo. Esta definición no solo proporciona un marco conceptual claro para entender qué se considera inteligencia artificial, sino que también sienta las bases para el desarrollo de normas y estándares que guiarán el uso y el desarrollo futuro de la inteligencia artificial.

En adición, Afanasyev (2108) entiende la IA como un tipo de sistema inteligente capaz de tomar decisiones por sí mismo. Este sistema representa la dirección del desarrollo de funciones informáticas relacionadas con la inteligencia humana, tales como: razonamiento, entrenamiento y resolución de problemas. En otras palabras, la IA es la transferencia de las capacidades humanas de actividad mental al plano de la informática y las tecnologías de la información, pero sin los defectos propios del ser humano. (pp.28-34)

Esta definición, a diferencia de las anteriores, destaca la capacidad de la IA de tomar decisiones y realizar tareas humanas como el razonamiento y la resolución de problemas sin los defectos humanos. Esto destaca la capacidad de la IA para superar las limitaciones humanas, aumentar la eficiencia y la precisión en una variedad de aplicaciones y transformar radicalmente la informática y las tecnologías de la información.

Pero las definiciones de IA no sólo se centran en un término, sino que algunos autores dividen la IA según sus objetivos. Entre ellos, hay que destacar a Marr (2018), que distingue entre tres tipos diferentes de inteligencia artificial, en función de los objetivos que se pretendan alcanzar con su uso. Normalmente, la gente invierte en el desarrollo de la inteligencia artificial con uno de estos tres fines:

El primero es el de crear sistemas que funcionen sin entender cómo funciona el pensamiento humano ("inteligencia artificial débil"), los ejemplos más famosos de este fin sería Siri de Apple o Alexa de Amazon. Sistemas que la gran mayoría tiene en sus casas y que utiliza en su vida cotidiana, cumple funciones sencillas y que no requieren mucha complejidad.

El segundo fin tiene como objetivo el uso del pensamiento humano como modelo, pero no necesariamente como objetivo final, por ejemplo, un diagnóstico médico o la prevención de delitos mediante la predicción. Es en este fin en el que nos centraremos en este trabajo, debido a su utilidad y los problemas que pueden generar, serán el objeto de estudio de este trabajo.

Por último, se intentan crear sistemas que piensen exactamente como las personas ("inteligencia artificial fuerte"), actualmente no existe ningún ejemplo debido a la complejidad del proceso. No obstante, se ha podido observar como en las definiciones, ya asemejan la inteligencia artificial al cerebro humano.

Cómo podemos observar, es difícil quedarse con una sola definición, ya que cada una aporta algo diferente al concepto de inteligencia artificial, aun así, uno se puede hacer una idea general sin meterse en conceptos muy técnicos.

2.1.2 Fundamentos de la inteligencia artificial: Aprendizaje automático y procesamiento de datos

Una vez vistas varias definiciones de qué es la inteligencia artificial, vamos a profundizar en los fundamentos de la inteligencia artificial, o lo equivalente a su funcionamiento, para comprender qué hay detrás de esos procesos informáticos y entender mejor cómo llegan a los resultados.

Berk (2021) indica que una característica esencial de la inteligencia artificial débil es que depende de algoritmos, no de modelos. Los algoritmos son métodos para calcular cosas, y que normalmente, deben recibir entrenamiento para que sigan aprendiendo y corrigiendo posibles errores que surjan.

Detrás de esos algoritmos se encuentran dos subconjuntos diferentes, el primero según Jakhar & Kaur (2020) es el “machine learning” (ML), el cuál es un subconjunto de la IA, que incluye todos los enfoques que permiten a las máquinas aprender de los datos sin ser programadas explícitamente. La intención del ML es entrenar a las máquinas basándose en los datos y algoritmos proporcionados. Con los datos y la información procesados, las máquinas aprenden a tomar decisiones. El ML es dinámico, lo que significa que tiene la capacidad de modificarse a sí mismo cuando se expone a más datos.

El aspecto de "aprendizaje" del ML significa que los algoritmos intentan minimizar los errores y maximizar la probabilidad de que sus predicciones sean ciertas. En resumen, el ML es simplemente una técnica para realizar la IA.

Hayward y Maas (2021) señalan en su estudio que, en la última década, tres desarrollos han propiciado el auge del "aprendizaje automático", que es un enfoque de IA más dinámico y menos "frágil", y estos son: los avances en "Big Data", la potencia de procesamiento y las innovaciones algorítmicas. Estos desarrollos van creciendo año tras año, lo que se traduce en una mejora progresiva de la IA.

El aprendizaje automático consiste en que el sistema aprenda gradualmente por sí mismo, las reglas correctas, o útiles, que necesita para realizar las tareas con eficacia. Y lo que es más importante, lo hace a partir de datos de entrenamiento, en lugar de programar explícitamente esas reglas, como ocurre con los sistemas expertos. Janiesch

et al. (2021) mencionan algunas aplicaciones del “machine learning” en algunas herramientas de nuestro día a día, como, por ejemplo, la capacidad de reconocer objetos, la traducción de palabras, o el registro del lenguaje en una conversación.

El segundo subconjunto explicado por Jakhar y Kaur (2020) es el “Deep learning” (DL). El DL es un subconjunto del ML, e incorpora modelos computacionales y algoritmos que imitan la arquitectura de las redes neuronales biológicas del cerebro, es decir, redes neuronales artificiales). Cada vez que el cerebro recibe información nueva, intenta compararla con la ya conocida para tratar de darle sentido. El cerebro descifra la información mediante el etiquetado y la asignación de los elementos a diversas categorías, y DL emplea el mismo concepto.

Cómo podemos observar, los autores ya acercan el proceso que usa la inteligencia artificial al del cerebro humano y las redes neuronales de éste. Por lo tanto, se está equiparando el procesamiento de información de ambos, la capacidad de aprendizaje, la adaptabilidad, y la toma de decisiones, entre otras muchas tareas del cerebro humano.

Jakhar y Kaur (2020) explican que "Deep" es un término técnico que se refiere al número de capas de una red neuronal artificial. Hay tres tipos de capas: la capa de entrada (recibe los datos de entrada), la capa de salida (produce el resultado del procesamiento de los datos) y la capa oculta (extrae los patrones dentro de los datos). Una red neuronal artificial profunda se diferencia de la superficial (una sola capa oculta) por tener un gran número de capas ocultas, lo que significa que es capaz de realizar tareas más complejas. Mientras los datos pasan de una capa oculta a otra, las características más simples se recombinan y recomponen en características complejas.

En pocas palabras, la DL funciona excepcionalmente bien con datos no estructurados y tiene mayor precisión que el ML, pero requiere un enorme volumen de datos de entrenamiento, además de hardware y software caros.

2.2 Aplicaciones de la inteligencia artificial en la criminología

La criminología ha experimentado un cambio significativo en la forma en que entendemos, prevenimos y combatimos el crimen. La inteligencia artificial está siendo usada para optimizar estrategias de aplicación de la ley, pronosticar tendencias criminales e identificar patrones delictivos. Como veremos a continuación, a través de algoritmos avanzados y análisis de datos masivos los cuáles se han mencionado en el apartado anterior, la IA posibilita abordar de una manera diferente la delincuencia de hoy en día.

Entre los diferentes ámbitos en los que se usan estos algoritmos, Berk (2021) señala los siguientes. Dentro de las comisarías de policía, estos algoritmos se suelen usar como una herramienta de apoyo en asuntos forenses, para la gestión de casos en laboratorios de criminalística, la comparación de ADN y la elaboración de perfiles metabólicos.

Por otro lado, la IA también es usada para la organización y análisis de pruebas sobre delitos específicos ocurridos; la elaboración de estadísticas de delincuencia y patrones de delincuencia espaciotemporales.

Pero su uso no sólo se limita fuera de la comisaría, sino que además se usa para la organización, redacción y recuperación de información, como el reconocimiento facial, a partir de vídeos de CCTV o cámaras corporales; transcripciones de voz a informes escritos; y evaluaciones del rendimiento por zonas geográficas (por ejemplo, comisarías) como en COMPSTAT; entre otras muchas funciones.

Esto demuestra que la inteligencia artificial juega un gran papel en las comisarías, y es que es una herramienta más que se debería usar para el trabajo, ya que resulta de gran ayuda para algunas funciones.

Por otro lado, Danaher (2018) hace una separación entre las tareas de la policía y la de las inteligencias artificiales. La primera distinción se encuentra en el propio trabajo policial, el cual implica una amplia gama de tareas puntuales (patrullar, rellenar formularios, asignar "un código de delito", etc.), en el que la IA puede tener un efecto considerable. Sin embargo, el trabajo de un agente de policía va más allá de estas tareas y abarca funciones más amplias, como el "community policing", la investigación,

reconfortamiento, detención, etc., y estas actividades pueden ser más difíciles de sustituir.

La segunda distinción es entre el uso de la IA como "herramienta", "aliada" y "usurpadora". Cuando la tecnología es una simple herramienta policial, se utiliza para ayudar en ciertas tareas específicas que forman parte del trabajo de un agente de policía.

Por ejemplo, investigaciones recientes han demostrado cómo el ML puede ayudar con los vínculos probatorios, como reconocer el calibre y el modelo de un arma a partir de grabaciones de audio de disparos (Raponi et al., 2022), cotejar los residuos de disparos en la escena del crimen con las características químicas de la munición similar (Gallidabino et al., 2019) o determinar qué tipo de zapato dejó una huella determinada en la escena del crimen (Kong, 2017). Hasta cierto punto, las herramientas de IA pueden incluso desempeñar un papel en la identificación y señalización de indicios de delitos mediados por otros sistemas de IA que se explicarán más adelante (King et al., 2019).

Cuando la IA opera como "aliada", ciertos aspectos de la tecnología funcionan de forma autónoma, pero siguen requiriendo la aportación y el análisis humanos, como los algoritmos de predicción de delitos. Y, por último, cuando la tecnología actúa como "usurpadora", no se requiere ninguna aportación humana.

Al categorizar los planos donde se puede usar la inteligencia nos muestra cuán de ayuda nos puede llegar a ser en día a día en una comisaría como una herramienta de apoyo, cómo la inteligencia artificial abre un abanico de nuevas formas de actuar.

Uno de los fines con los que se usaba la inteligencia artificial era el uso para la predicción de delitos. En este aspecto Ligeti (2019) en su artículo menciona que, mediante algoritmos que procesan ingentes cantidades de datos, hacen predicciones sobre dónde y cuándo es probable que se cometan delitos. Además de si es probable que un sospechoso o acusado, huya o cometa más delitos y se convierta en reincidente, con la consecuencia de que los tribunales penales pueden denegar la libertad bajo fianza u optar por penas severas (justicia predictiva).

Esta posibilidad que nos ofrece la inteligencia artificial podría llegar a revolucionar lo que entendemos por la predicción policial, y es una aplicación que ya se está poniendo en uso en diversos países.

En cuanto a la investigación policial, Rigano (2019), se enfoca en la recopilación de pruebas, señala que los sistemas de IA pueden ser de gran utilidad para las autoridades policiales, incluso más allá de los ejemplos mencionados de actuación policial predictiva. En el análisis, por ejemplo, del ADN o de perfiles de redes sociales "produce grandes cantidades de datos complejos en formato electrónico", que pueden contener patrones útiles que el análisis humano no podría captar.

Todo ello muestra lo implementadas y desarrolladas que se encuentran las inteligencias artificiales dentro de las policías de todo el mundo, aunque sí es cierto que aún queda camino por recorrer y margen de mejora, pero esto indica que la inteligencia artificial está siendo bien recibida.

En Estados Unidos, por ejemplo, Puente & Chang (2019) indican que la policía californiana utiliza un programa informático llamado PredPol para predecir dónde y cuándo es más probable que se produzcan determinados delitos. Y en el caso de Europa, Završnik (2021) menciona que algunos cuerpos de policía europeos también recurren a un programa informático similar, Precobs (Pre Crime Observation System).

Con estos ejemplos se muestra el gran interés que tienen las instituciones en todo el mundo en introducir la inteligencia artificial para cuestiones que suponen un problema, como la predicción de crímenes.

Hasta ahora solo hemos visto la aplicación de la inteligencia artificial en la policía, pero la realidad es que también se está llegando a implementar en el ámbito judicial, como una herramienta de apoyo a los jueces para la toma de sus decisiones.

Portillo et al. (2022) estudió el uso de la inteligencia artificial en la policía de Colombia, como el Sistema Penal Oral Acusatorio (SPOA) y el Watson Machine Learning, creado por IBM, dos herramientas tecnológicas usadas por el personal judicial de la Fiscalía General de la Nación (FGN). La primera es una base de datos de gran tamaño, mientras que la segunda permite el análisis de datos no estructurados, como el relato de los hechos de un delito violento.

Los resultados mostrados por Portillo et al. (2022) indicaron que las personas que respondieron creían que tanto SPOA como Watson eran útiles y efectivos para conectar casos de delitos violentos con el proceso de perfilación criminal. Sin embargo, se mencionaron algunas desventajas de los programas, como la falta de organización o que, debido a su complejidad, no se utilizan a gran escala en organizaciones competentes.

De los sistemas usados en sentencias, el más revolucionario lo describe Lagioia et al. (2023), cuyo estudio se centra sobre un sistema utilizado en Estados Unidos, y que ha cogido mucha fuerza, llamado COMPAS (Correctional Offender Management Profiling for Alternative Sanctions). Es un instrumento de evaluación de riesgos y medidas necesarias, utilizado en el sistema de justicia penal para evaluar los perfiles de riesgo de los acusados: riesgo de reincidencia, riesgo de violencia y riesgo de no comparecencia ante el tribunal.

En su artículo Lagioia et al. (2023) explican que las evaluaciones realizadas por el sistema COMPAS son tenidas en cuenta por los jueces para decidir la concesión del beneficio de libertad condicional u otras medidas. El riesgo se evalúa mediante algoritmos estadísticos mediante el ML y se cuantifica en puntuaciones de riesgo. Dichas puntuaciones se calculan a partir de múltiples datos, que incluyen factores estáticos-históricos (como antecedentes penales, edad de la primera detención, relaciones criminales) y factores dinámicos-criminógenos (como estabilidad residencial, situación laboral, vínculos con la comunidad, abuso de sustancias, inclusión y relaciones sociales, y situación familiar), así como las respuestas a 137 preguntas de respuesta múltiple.

Para entenderlo más en profundidad, Avella & Sanabria-Moyano (2022) explican que COMPAS tiene dos modelos de riesgo principales: el primero de Reincidencia general, y el segundo de Reincidencia violenta. La Evaluación de Riesgo de Reincidencia General se utiliza para predecir nuevos delitos; y la Evaluación de Riesgo de Reincidencia Violenta se centra en la probabilidad de cometer delitos violentos, es decir, asesinato, homicidio, violación, robo y agresión con agravantes.

Este es el modelo más usado de inteligencia artificial en los juicios, mostrando que es capaz de ser usado en la mayoría de los ámbitos. Sin embargo, hay que tener en

cuenta que este modelo es una herramienta de apoyo, y que la decisión final la debe tener el juez.

La inteligencia artificial es capaz de manejar una gran cantidad de datos que un humano no podría, sin embargo, hay otros aspectos que la inteligencia artificial no es capaz de tener en cuenta, cómo los principios jurídicos, las circunstancias atenuantes y agravantes, el impacto social, o las circunstancias del delincuente. Por ende, es crucial que el juez tenga en cuenta estos aspectos, y las recomendaciones de otros profesionales, como psicólogos o criminólogos, a parte de la IA, para poder tomar la mejor decisión.

2.2.1 Usos criminales de la inteligencia artificial

Sin embargo, el uso de la inteligencia artificial no sólo se limita al policial o al judicial, sino que al ser accesible para todo el mundo también es aprovechado por los delincuentes, para facilitar la comisión de crímenes que ya existían como herramienta, y la creación de nuevos crímenes a raíz del uso de la IA.

Ya son numerosos los estudios que han puesto de manifiesto hasta qué punto las organizaciones delictivas pueden beneficiarse de la IA. Un ejemplo de ello se encuentra en el artículo de Ligeti (2019), en el cual menciona, que el tráfico de drogas está resultando más fácil, y mucho menos arriesgado para los delincuentes, ya que las sustancias ilegales se trasladan de un lugar a otro mediante drones. Lo mismo ocurre con los atentados terroristas, que pueden llevarse a cabo colocando materiales explosivos en máquinas dotadas de IA.

Otro de los delitos que menciona Ligeti (2019) en los que se está empezando a implementar la inteligencia artificial es el fraude en línea, especialmente el "spear phishing", que se refiere a la "estafa por correo o comunicaciones electrónicos dirigidas a una persona, organización o empresa específica. Aunque el phishing mediante correos electrónicos descaradamente falsos no suele tener éxito, los sistemas de IA pueden crear y enviar correos electrónicos fraudulentos adaptados al destinatario, al que pueden convencer para que siga un enlace malicioso y/o comparta sus datos con el estafador.

Cómo ya se ha mencionado antes, el desarrollo de la IA puede dar lugar a nuevos delitos. En un informe de Interpol (2019) junto a UNICRI, comentan que debería realizarse un estudio sobre los nuevos delitos que implican el uso malicioso de la IA y la robótica.

Sin embargo, también hay que tener en cuenta que los sistemas de IA pueden convertirse en víctimas de delitos y es probable que se necesiten nuevas definiciones y normas para regular estas situaciones. Por ejemplo, los sistemas de IA pueden ser sabotados por terceros de modo que estos sistemas se vean impedidos de alcanzar sus objetivos y/o inducidos a cometer un delito. Se podría pensar en personas que intencionadamente perturban el software de coches de conducción autónoma, provocando así accidentes que estaban totalmente fuera del control del programador y usuario del vehículo.

Por otra parte, la IA es un poderoso instrumento para conductas peligrosas que podrían penalizarse en el futuro. Zarina et al. (2019) se enfocan en que los sistemas de IA pueden encargarse, por ejemplo, de la creación y difusión de noticias falsas. Aunque esto ya representa un problema complejo en la sociedad contemporánea, conductas similares no suelen constituir delito, salvo contadas excepciones.

La IA tiene el potencial de intensificar este fenómeno, hasta el punto de representar una amenaza desalentadora y sin precedentes para nuestras democracias, ya que es probable que el nivel de precisión de las máquinas dificulte incluso al usuario más atento distinguir la verdad de la ficción. Es una de las razones por lo que merece la pena examinar si las conductas en cuestión deberían tipificarse como delito para reducir sus efectos potencialmente devastadores.

Por último, Zarina et al. (2019) menciona las posibles interacciones entre la IA y la criptografía, con especial atención a las tecnologías que se basan en la criptografía, como blockchain, sugiriendo más atención en el futuro. Señalan que es difícil regular las implicaciones jurídicas -incluidas las penales- de blockchain y los criptoactivos en sí mismos, y que las posibles combinaciones con la IA pueden plantear cuestiones aún más complejas, ya que puede facilitar la comisión de delitos existentes o nuevos, y requerir potencialmente la introducción de disposiciones penales a medida. Pero que, aun así, existe una falta de un análisis en profundidad de la cuestión en literatura.

2.3 Ética y desafíos de la inteligencia artificial

2.3.1 Consideraciones éticas

Al igual que en la mayoría de los ámbitos, la inteligencia artificial como una herramienta también debe ser usada éticamente, respetando los derechos de las personas. En este apartado se explorarán las consideraciones éticas que surgen del desarrollo y la implementación de la IA.

Según Ulnicalne (2022), la necesidad de desarrollar un marco ético para la IA se ha destacado en todos los documentos de política de IA de la UE. En 2018, la Comisión creó un grupo independiente de 52 expertos de alto nivel (HLEG) sobre IA para desarrollar las directrices éticas.

La Comisión Europea (2019) propone un conjunto de 7 requisitos clave que deben cumplir los sistemas de IA para ser considerados fiables. Una lista de evaluación específica pretende ayudar a verificar la aplicación de cada uno de los requisitos clave:

Agencia y supervisión humanas: Los sistemas de IA deben capacitar a los seres humanos, permitiéndoles tomar decisiones informadas y fomentando sus derechos fundamentales. Al mismo tiempo, deben garantizarse mecanismos que lo supervisen adecuadamente, lo que puede ser logrado mediante enfoques basados en el ser humano dentro el bucle, el ser humano en el bucle, y el ser humano al mando.

Robustez técnica y seguridad: Los sistemas de IA deben ser resistentes y seguros. Tienen que ser seguros, garantizar un plan de emergencia en caso de que algo vaya mal y ser precisos, fiables y reproducibles. Es la única manera de garantizar que también se puedan minimizar y prevenir los daños no intencionados.

Privacidad y gobernanza de datos: además de garantizar el pleno respeto de la privacidad y la protección de datos, también deben garantizarse mecanismos adecuados de gobernanza de datos, teniendo en cuenta la calidad y la integridad de los datos, y garantizando un acceso legitimado a los mismos.

Transparencia: los datos, el sistema y los modelos de negocio de la IA deben ser transparentes. Los mecanismos de trazabilidad pueden ayudar a conseguirlo. Además, los sistemas de IA y sus decisiones deben explicarse de forma adaptada a la parte

interesada. Los seres humanos deben ser conscientes de que están interactuando con un sistema de IA, y deben ser informados de las capacidades y limitaciones del sistema.

Diversidad, no discriminación y equidad: Deben evitarse los prejuicios injustos, ya que podrían tener múltiples implicaciones negativas, desde la marginación de grupos vulnerables hasta la exacerbación de los prejuicios y la discriminación. Para fomentar la diversidad, los sistemas de IA deben ser accesibles para todos, independientemente de cualquier discapacidad, e implicar a las partes interesadas pertinentes a lo largo de todo su círculo vital.

Bienestar social y medioambiental: Los sistemas de IA deben beneficiar a todos los seres humanos, incluidas las generaciones futuras. Por ello, debe garantizarse que sean sostenibles y respetuosos con el medio ambiente. Además, deben tener en cuenta el medio ambiente, incluidos otros seres vivos, y su impacto social y societal debe considerarse cuidadosamente.

Rendición de cuentas: Deben establecerse mecanismos que garanticen la responsabilidad de los sistemas de inteligencia artificial y sus resultados. La auditabilidad, que permite la evaluación de algoritmos, datos y procesos de diseño, desempeña un papel clave en este sentido, especialmente en aplicaciones críticas. Además, debe garantizarse una reparación adecuada y accesible.

Una vez visto Europa, si nos vamos a un enfoque más general, la Organización de las Naciones Unidas, integrada por 193 países, tiene la intención de ser la entidad internacional que respalde y facilite la cooperación mundial en cuanto a la inteligencia artificial.

La ONU estableció un nuevo enviado de tecnología en 2022. En el mismo año, la UNESCO (2022) y los países miembros también aprobaron voluntariamente un marco de ética de la inteligencia artificial. Los miembros se comprometen, entre otras muchas, evaluar el impacto ambiental de la IA, a realizar evaluaciones de impacto ético para la IA, garantizar que se promueva la igualdad de género, y que no se utilice para la vigilancia masiva.

Una vez observados todos los puntos, teniendo en cuenta lo que se sabe de la inteligencia artificial, la fuente principal y lo que ayuda a progresar a la inteligencia

artificial son los datos, ya sean de internet, de las empresas, o de las personas. Y, por lo tanto, es el punto crítico que es objeto de debate.

En su artículo Zarina et al. (2019) menciona que hay dos problemas éticos con la IA: el primero hace referencia a la ética de los datos, y la segunda sobre problemas relacionados con la toma de decisiones de la IA basadas en datos generalizados.

Según este autor, en cuanto a ética de los datos, menciona que se está desarrollando y adquiriendo relevancia, como demuestran documentos pertinentes de Estados, empresas e instituciones sociales.

Definir los límites del acceso ético a los datos es un problema complejo que afecta a diversas partes interesadas, como ciudadanos, el Estado, las empresas, instituciones públicas, etc., y requiere una solución global. Por lo que las restricciones excesivas al acceso a los datos pueden frenar el desarrollo de las tecnologías de IA.

Respecto al segundo problema ético que se plantea, está relacionado con la ética y la humanidad de las decisiones de la IA. Zarina et al. (2019) hace mención del profesor Arkin, que se dedica al desarrollo de robots para necesidades militares. Señala que su investigación tiene importantes riesgos éticos, posibles cuando se utilizan sus desarrollos con fines delictivos.

Es decir, la responsabilidad no sólo cae en la IA, sino en quién la usa. A partir de este uso de la IA, lleva inevitablemente al problema de la elección ética.

Este problema es especialmente relevante para los vehículos no tripulados controlados por IA. Kingston (2016) describió una situación hipotética en 2023 en la que un coche controlado por IA que circulara por las calles de la ciudad atropellara a un peatón, y se preguntaba sobre la responsabilidad penal.

Esta cuestión fue el objeto de investigación de Casmi (2018), que recogió un estudio a gran escala llamado "Máquina Moral" ("Ética para el coche") hecho por el MIT en 2016, dentro de la cual se creó un sitio web especial, donde el usuario del coche piloto simuló situaciones con diversas situaciones de fuerza mayor. Los escenarios ofrecían la oportunidad de elegir en la carretera, en caso de emergencia, qué vidas sacrificar en primer lugar en caso de accidente, cuya tragedia es ya inevitable. Estos

datos serían usados posteriormente en los coches tripulados para saber qué decisión tomaría la IA en caso de encontrarse en tal situación.

Representantes de la empresa automovilística alemana "Mercedes-Benz" señalaron a su vez que sus coches darían prioridad a los pasajeros. Ante esa declaración, respondió al Ministerio de Transporte de Alemania que hacer tal elección en base a una serie de criterios sería ilegal, y, en cualquier caso, el fabricante será el responsable.

No es una excepción el ámbito de la sanidad, donde la introducción de la IA en el proceso de tratamiento y diagnóstico de enfermedades oncológicas ha tenido consecuencias dispares.

En el estudio de Kolenov (2018) se centró en la empresa "IBM", la cual ha desarrollado la AI médica "WatsonHealth", utilizado en 230 hospitales de todo el mundo para el tratamiento de 13 tipos de cáncer a 84 000 pacientes. "WatsonHealth" ofrece tratamientos que pueden realizar basado en los datos de anteriores casos, teniendo en cuenta síntomas, operaciones que se pueden realizar, y los casos exitosos de dichas operaciones.

Sin embargo, Kolenov (2018) relata que se filtraron documentos internos de dicha empresa, la cual era una de los mayores fabricantes y proveedores de hardware y software del mundo, IBM. Estos documentos indican que WatsonHealth, la IA médica desarrollada por ella, cometía errores médicos. Por lo tanto, WatsonHealth ofrece tratamientos incorrectos que pueden llevar a la muerte del paciente.

Y de aquí surge el problema legal, ¿el responsable de la muerte es la empresa que desarrolló la IA?, ¿del médico que se apoyó en la IA? Así pues, es esencial insistir en la necesidad de un marco ético claro, riguroso y eficaz en el diseño, la construcción, la producción, el uso y la modificación de la IA.

2.3.2 Desafíos y riesgos de la IA

Cómo se ha visto en el apartado anterior, uno de los desafíos de la regulación de la IA es la responsabilidad legal de ésta. Ligeti (2019) en su estudio recoge, entre los juristas que se ocupan de las cuestiones de la personalidad jurídica de la IA, tres enfoques clave:

- dotar a la IA de una personalidad jurídica correspondiente a la humana;
- la dotación a la IA de una personalidad jurídica similar a la de una persona jurídica;
- la dotación a la IA de personalidad jurídica limitada.

Desde mi punto de vista es inadecuado involucrar a los diseñadores y desarrolladores de IA, cuando el resultado de la decisión final de la IA depende en gran medida de la situación de su aplicación y de las tareas que se le asignen, por lo que el resultado final no puede ser conocido por los creadores.

Coincido con la opinión de Nevejans (2016) de que la personalidad jurídica de la IA no puede equipararse a la condición humana o jurídica de una persona jurídica. Una persona con personalidad jurídica actúa sobre la base de procesos mentales, guiada por creencias subjetivas. Detrás de las acciones de una persona jurídica hay individuos, sin los cuales las actividades de una persona jurídica son imposibles. La persona jurídica, a su vez, actúa de forma independiente, sin conciencia ni sentimientos.

Desde otro punto de vista, Bikeev et al. (2019) opinan que la idea de reconocer a la IA como sujeto de derecho contradice ideas sobre el sujeto de derecho como el valor socio-jurídico, la dignidad, la voluntad jurídica autónoma, y que también entra en conflicto con la composición de la relación jurídica, la composición del delito y es insignificante en el marco de la institución de la representación.

Al mismo tiempo, Bikeev et al. (2019) hacen un apunte, y es que la IA no tiene las características necesarias y suficientes del sujeto de derecho, en particular, no tiene el potencial de adquirir y ejercer de forma independiente derechos subjetivos y

obligaciones legales, de asumir la responsabilidad legal, de tomar decisiones legales de forma independiente, no tiene sus propios intereses y aspiraciones legales, etc.

Son muchos los juristas que están a favor de exigir responsabilidades penales a la IA, pero también los que piensan que es imposible imputarle un delito a la IA, ya que no tiene personalidad jurídica.

Si cada país acaba regulando la cuestión de la inteligencia artificial y las pruebas penales con arreglo a sus propios principios, normas y tal vez incluso estándares técnicos, la diversidad de regímenes diferentes puede obstaculizar la cooperación judicial, por lo que cabe preguntarse si no sería apropiado un enfoque coordinado a escala internacional. Así como con los ciberdelitos, se debería crear leyes internacionales, no sólo reduciéndose a la sanción del delito, sino a la cooperación entre países para seguir desarrollando estas medidas y garantizar el desarrollo ético y legal de la IA.

Otro desafío que surge tiene relación con un punto anterior, respecto al uso de la IA como herramienta de apoyo en la toma de decisiones judiciales.

Algunos autores, como Veale et al. (2019) opinan que se vulnera el derecho a la defensa y a un juicio justo. Si las investigaciones se basan en técnicas de IA, el demandado debería estar en condiciones de comprender cómo se han recopilado las pruebas, mientras que la denegación del descubrimiento en relación con el programa, el código o los datos que rigen el sistema de IA representaría una clara infracción del principio de igualdad de armas entre las partes.

Por ejemplo, una solución que propone Ligeti (2019) es que se podría ayudar a fomentar la fiabilidad y la transparencia de las técnicas de IA, pidiendo y proporcionando una certificación independiente de la fiabilidad del sistema de IA. El juez podría designar a un perito para verificar el proceso algorítmico o la red neuronal de un determinado sistema de IA, siempre que las partes expresen sus dudas sobre la exactitud de los datos automatizados.

Por último, Berk (2021) señala otro riesgo que conlleva el uso de la IA, y este se relaciona con la precisión. Muchos algoritmos de riesgo de ML son dependientes del

tamaño de la muestra, por lo que, si no es suficientemente grande la cantidad de datos, la precisión puede sufrir.

Esto nos lleva a cuestionar la precisión del ML, aunque Berk (2021) menciona que el ML es preciso, por ejemplo, en cuanto a si un criminal va a reincidir o no, casos en los que acierta la gran mayoría, sí que no es tan preciso aquellos en los que no están tan claros. Dando lugar a los falsos positivos y a los falsos negativos.

En resumen, los desafíos de la inteligencia artificial en la criminología incluyen la protección de la privacidad y los derechos humanos, y la precisión y equidad en la toma de decisiones. Para garantizar que el uso de la IA en este campo contribuya a una justicia más eficiente y justa sin comprometer los principios éticos y legales fundamentales, es esencial que esté acompañado de una supervisión ética rigurosa y un marco regulatorio sólido.

2.4 Marco legal de la inteligencia artificial en Europa

Ante los problemas que nos hemos ido encontrando, las distintas entidades internacionales y gubernamentales ya han ido desarrollando medidas para regular el uso de la IA, y así tener controlado su progreso. En la actualidad, muchos países se dan cuenta, de forma bastante razonable, de que la regulación jurídica difícilmente puede considerarse como el único o principal mecanismo de regulación del uso de la inteligencia artificial.

En el estudio de Zarina et al. (2019) se indica que existe una práctica bastante extendida de iniciativas privadas que combinan los esfuerzos de varias empresas para desarrollar el uso de tecnologías de inteligencia artificial y consolidar los principios básicos del trabajo con dichas tecnologías. Por ejemplo, los principios de la IA de Asilomar (2017) tienen como objetivo crear inteligencia útil, mantener los valores humanos y la confidencialidad de los datos personales.

Tras una serie de trágicos incidentes relacionados con el uso de vehículos aéreos no tripulados, el estudio de Delvaux (2017) se centra en la Unión Europea, en la cual se ha debatido ampliamente la posibilidad de conceder estatuto jurídico a los robots y, en

consecuencia, la posibilidad de llevar ante la justicia a una persona electrónica (sujeto electrónico).

Así, la resolución del Parlamento Europeo (2020), junto con las recomendaciones de la Comisión de Regulación del Derecho Civil de la Robótica del Parlamento Europeo de 16 de febrero de 2017 "Normas civiles sobre robótica" tiene por objeto regular el estatuto jurídico de los robots en la sociedad humana a través de las siguientes acciones:

- La creación de una agencia europea especial para la robótica y la IA.
- El desarrollo de una definición reglamentaria de "robot inteligente autónomo".
- El desarrollo de un sistema de registro para todas las versiones de robots, así como un sistema para su clasificación.
- El desarrollo de requisitos para que los desarrolladores ofrezcan garantías de prevención de riesgos.
- El desarrollo de una nueva estructura de presentación de informes para las empresas que utilizan o necesitan robots, que incluirá información sobre el impacto de la robótica y la IA en los resultados económicos de la empresa.

Y es que esa idea de crear agencias europeas especializadas en IA está creciendo, son numerosos autores, como Zarina et al. (2019), que opinan que se debería crear una nueva institución independiente de derecho penal, diferente de acercamiento tradicional antropocéntrico. Una institución jurídica de este tipo requiere un enfoque completamente nuevo.

Desde su punto de vista, Zarina et al. (2019) consideran que el marco de esta institución se debería basar en la simbiosis de las características técnicas y de otro tipo de la IA, así como tipos alternativos de responsabilidad, como la desactivación, la reprogramación o la atribución de la condición de "delincuente", que servirán de advertencia para todos los participantes en las relaciones jurídicas. Creen que una solución de este tipo en el futuro puede minimizar los riesgos criminológicos del uso de la IA.

En 2019, 38 países miembros de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) acordaron adoptar un conjunto de principios no vinculantes que establecen algunos valores que deben sustentar el desarrollo de la IA.

Según estos principios que sugiere la OCDE (2019), los sistemas con este tipo de inteligencia deben ser transparentes y explicables; deben funcionar de forma sólida, segura y protegida; deben contar con mecanismos de rendición de cuentas; y deben diseñarse respetando el Estado de derecho, los derechos humanos, los valores democráticos y la diversidad.

Desde entonces, otras iniciativas políticas en todo el mundo se han inspirado en estos principios, que sirven como una especie de manifiesto de la política occidental en relación con la IA.

La Unión Europea ha puesto el foco estos últimos años en el estudio sobre las consecuencias del uso de la IA. Le solicitó a la dra. González, G. (2020), un estudio para analizar cómo la inteligencia artificial afecta los derechos fundamentales en el ámbito de la aplicación de la ley y la justicia penal en la Unión Europea. En este estudio se presenta el marco legal relevante, especialmente en lo que respecta a la protección de datos, y analiza las principales tendencias y discusiones políticas. Al final argumenta que la gravedad y magnitud de los problemas podrían requerir una intervención por parte de la Unión Europea.

La Organización Internacional de Estandarización (ISO), en 2023 ya ha establecido pautas sobre cómo las empresas deben llevar a cabo la gestión de riesgos, las evaluaciones de impacto y la gestión del desarrollo de la inteligencia artificial. Estas normas ayudan a las empresas a implementar regulaciones complicadas, ayudando a crear productos que funcionen en múltiples jurisdicciones a medida que los países empiecen a redactar sus propias leyes individuales para la IA.

Aunque no fue hasta 2021 cuando se propuso la primera ley que regularía la inteligencia artificial por el Parlamento Europeo, y que tendría un impacto en todos los países.

Según el Portal de Administración Electrónica (2023) la propuesta de la Comisión es una parte importante de la política de la UE para promover la creación y

adopción de una IA segura y legal que respete los derechos fundamentales en todo el mercado único. La propuesta adopta un enfoque orientado al riesgo y propone un marco legal coherente y transversal para la inteligencia artificial, con el fin de asegurar la certeza legal.

Ya el Parlamento Europeo (2024), anunció un año después que la Eurocámara había aprobado dicha ley que regularía la inteligencia artificial. Lo más destacable de esta ley es que regula algunas aplicaciones de IA que violan los derechos de los ciudadanos, como la categorización biométrica y el reconocimiento de emociones en el trabajo, las cuales quedan prohibidas. Se implementan obligaciones para los sistemas de alto riesgo, que incluyen evaluar y reducir el riesgo, mantener registros y ser transparentes. Medidas de apoyo a la innovación: Se crearán espacios de prueba para pymes y startups para desarrollar IA innovadora antes de su comercialización.

En la ley propuesta por el Parlamento Europeo (2024), para prevenir los excesos más perjudiciales de la IA, la ley indica que se podría responsabilizar a los malos actores mediante la imposición de grandes multas y la prevención de la venta y uso de tecnología contraria a la UE. La Unión Europea es la primera en iniciar la iniciativa, ya que es la única normativa exhaustiva sobre IA. Es posible que la UE se convierta en la regulación global sobre IA, ya que las compañías de otros países que deseen operar en este bloque comercial tendrán que modificar sus prácticas para cumplir con las leyes.

Hay que tener en cuenta que esta ley ha sido aprobada, pero que todavía no ha entrado en vigor, por lo que todavía no se puede estudiar su impacto de momento. Aun así, desde mi punto de vista, la regulación me parece bastante coherente y realizada con mucho cuidado, centrándose en los aspectos relevantes de la inteligencia artificial y afrontando los desafíos que provenían de su uso.

Quizá, a partir de esta ley, se podrá regular debidamente en España el uso de la inteligencia artificial, ya que como apunta Bahamontes (2024), los delitos relacionados con la inteligencia artificial aún no están claramente definidos en el Código Penal. Lo que lleva a que las situaciones que involucran IA están sujetas a regulaciones generales, como las relativas a delitos informáticos o contra la intimidad.

2.5 Teorías criminológicas

Por último, para explicar la comisión de delitos apoyándose con el uso de la IA, me baso en dos teorías que lo podrían explicar.

La primera es la teoría de la oportunidad, explicada por Felson & Clarke (2008), se enfoca en cómo las oportunidades para cometer delitos influyen en su ocurrencia. Esta teoría sostiene que no solo las disposiciones o motivaciones de los delincuentes son responsables del delito, sino también la disponibilidad de oportunidades para cometerlo. La falta de vigilancia, la disponibilidad de objetivos fáciles y las circunstancias que facilitan la comisión de delitos son algunas de las muchas razones por las que pueden surgir oportunidades delictivas.

Los delincuentes que utilizan IA buscan una gran cantidad de datos valiosos en línea, como información personal y financiera; si no hay sistemas de seguridad actualizados y funcionales, los delincuentes pueden explotar vulnerabilidades con herramientas de IA; la motivación de los criminales surge de la búsqueda en la IA para automatizar y mejorar la eficiencia de sus ataques, facilitando los ataques como el phishing avanzado, los deepfakes, el ransomware automatizado y el fraude financiero al analizar grandes volúmenes de datos y simular comportamientos humanos.

La segunda teoría sería la de la elección racional, Grautoff et al. (2011) explican que las personas cometen delitos después de realizar un proceso de toma de decisiones racional. De acuerdo con esta teoría, las personas evalúan los beneficios y los costos de sus acciones antes de decidir si cometer o no un delito. La probabilidad de éxito, las posibles recompensas y los riesgos de captura y castigo son factores considerados.

El uso de la IA en actividades delictivas puede entenderse como una estrategia para maximizar los beneficios y minimizar los riesgos, según la teoría racional de criminología. Por ejemplo, la IA puede aumentar significativamente las recompensas potenciales al mejorar la eficiencia y efectividad de las actividades delictivas. Además de reducir el riesgo de detección mediante técnicas avanzadas de evasión y anonimato, también automatiza procesos complejos para reducir los costos operativos.

2.6 Formulación de hipótesis: Resultados esperados

En esta sección se desarrollará la hipótesis de investigación, que surgirán de la revisión de la literatura, del marco teórico y de los objetivos propuestos.

Hipótesis 1: La implementación de la inteligencia artificial en la criminología ha ayudado a mejorar la predicción y prevención de delitos.

Hipótesis 2: La implementación de la inteligencia artificial en la criminología plantea desafíos éticos relacionados con la transparencia y la responsabilidad en la toma de decisiones automatizadas.

3. METODOLOGÍA DE INVESTIGACIÓN

3.1 METODOLOGÍA

La investigación desarrollada tiene un enfoque cualitativo, ya que tratará del estudio de un caso, que será analizado en profundidad para comprender mejor el fenómeno que se está estudiando. Este enfoque permite una exploración exhaustiva y detallada del fenómeno particular, además esta metodología es particularmente útil cuando se desea comprender todos los matices y complejidades de un caso específico, como es el que se va a realizar.

Se han descartado los diseños de tipo experimental, cuasiexperimental y correlacional, ya que encontramos el modelo exploratorio como el más eficaz para el diseño de nuestra investigación. Además de que en nuestra investigación no se incluyen suficientes variables para desarrollar estos diseños, siendo la muestra no aleatoria, y, por lo tanto, nuestro muestreo es no probabilístico.

Dentro del muestreo no probabilístico, nuestra investigación se compone del estudio de un caso. Como muestra se ha elegido un caso de Estados Unidos, en la que una decisión judicial fue tomada con apoyo de la inteligencia artificial, y causó diversas opiniones.

La elección de un solo caso permite realizar un examen completo y detallado del fenómeno. Se puede examinar minuciosamente cada aspecto en lugar de ofrecer una visión superficial y generalizada. Esto proporciona una comprensión más completa y rica del caso estudiado.

La fuente principal de este trabajo ha sido la investigación bibliográfica, debido a que la mayoría de este ha sido la consulta de fuentes secundarias como trabajos de investigación, libros, artículos científicos, y documentos académicos que estén relacionados con el tema de estudio.

3.2 CONSIDERACIONES ÉTICAS

Este trabajo cumple con los objetivos de la Agenda 2030 establecidos por las Naciones Unidas (2015), específicamente con los siguientes:

Cumple con el objetivo n° 4: *Educación de calidad*. El uso de la IA en criminología requiere empleados con habilidades avanzadas en tecnología y análisis de datos. Este objetivo se relaciona con la necesidad de capacitar a los criminólogos y otros profesionales de la justicia penal en el uso de tecnologías avanzadas. Los cursos sobre IA, análisis de datos y ciberseguridad deben incluirse en los programas educativos y de capacitación. Por lo tanto, se asegurará de que los futuros profesionales estén bien equipados para usar estas herramientas en sus trabajos de manera ética y efectiva.

Cumple con el objetivo n° 8: *Trabajo decente y crecimiento económico*. En criminología, el análisis de la IA puede ayudar a modernizar significativamente los sistemas de justicia penal. El análisis de patrones delictivos, la predicción de delitos y la gestión de recursos policiales son ejemplos de tareas en las que la IA puede ayudar a optimizar procesos y aumentar la eficiencia operativa. Esto no solo mejora la productividad de las fuerzas de seguridad, sino que también puede fomentar un entorno más seguro, lo que impulsa la economía al reducir los costos asociados con la delincuencia.

Cumple con el objetivo n° 9: *Industria, innovación e infraestructuras*. En criminología, el uso de la IA es un campo emergente de innovación que requiere una inversión significativa en investigación y desarrollo. Las fuerzas de seguridad y los sistemas judiciales pueden mejorar su capacidad tecnológica mediante la financiación de proyectos de investigación en este campo. La Unión Europea podría liderar la innovación en la justicia penal al invertir en estas tecnologías, creando herramientas sofisticadas que puedan ser utilizadas a nivel mundial y mejorando la infraestructura tecnológica del sector.

Cumple con el objetivo n° 16: *Paz, justicia, e instituciones sólidas*. La inteligencia artificial tiene el potencial de mejorar el estado de derecho y el acceso a la justicia. La IA puede mejorar la imparcialidad en los procedimientos judiciales identificando patrones de discriminación o sesgo en grandes cantidades de datos legales

y judiciales. Además, la IA puede ayudar a crear comunidades más seguras y justas al detectar y prevenir delitos de manera más efectiva. No obstante, es fundamental implementar estas tecnologías con cuidados éticos para garantizar que no se perpetúen injusticias ni se violen derechos fundamentales.

El caso analizado ha sido obtenido de manera legal, ya que es público, por lo que cualquier persona tiene acceso a él, evitando quebrantar las leyes y respetando la privacidad de dicho documento. Al igual se mantiene la imparcialidad y la objetividad en todo momento, para realizar en análisis más eficiente.

3.3 LIMITACIONES DEL ESTUDIO

Respecto a las limitaciones que me encuentro con el estudio, primero me encuentro con las pocas investigaciones que hay sobre el tema que estoy tratando, y sobre las que me puedan servir de base, y así realizar una investigación más eficaz.

Otro aspecto que limita mi estudio, son las pocas sentencias judiciales que se pueden analizar, y eso es debido a que la inteligencia artificial no está tan implementada en el ámbito jurídico, por no mencionar que, en Europa, judicialmente no está implementada, y, por tanto, tendré que investigar un caso fuera de Europa. Esto conlleva una legislación diferente, y por tanto, diferentes perspectivas de los derechos, dependiendo del país.

Por último, la poca y pobre legislación que regula la inteligencia artificial. Este aspecto también delimita la investigación, al dificultar el estudio a nivel legislativo y sus consecuencias debido a un mal uso. Aunque ya hay borradores de leyes que entrarán en vigor dentro de poco, no es suficiente para estudiar su impacto en la realidad.

3.4 CONTRASTE DE HIPÓTESIS

Para evaluar la validez de las hipótesis presentadas en la introducción del presente estudio, se analizarán en este apartado y se presentará evidencia empírica que respalde o refute estas afirmaciones.

Respecto a la hipótesis 1, se ha conseguido la suficiente información para respaldar esta hipótesis. La inteligencia artificial es capaz de analizar una gran cantidad de datos, y se ha demostrado que ha ayudado a mejorar la predicción y prevención de delitos en cierta medida. Sin embargo, esto supone una estigmatización para el individuo que se esté evaluando o la zona en sí, por lo que hay que tratar los datos con mucho cuidado.

Respecto a la hipótesis 2, se ha comprobado que la transparencia y la toma de decisiones automatizadas son unos de los muchos desafíos que presenta la implementación de la inteligencia artificial en la criminología. Presentan un riesgo para los derechos fundamentales, y debe ser manejado con delicadeza.

4. ANÁLISIS DEL CASO

Los objetivos de Desarrollo Sostenible propuestos por las Naciones Unidas (2015) que se mencionan en el apartado de consideraciones éticas están relacionados con el análisis y son cruciales para garantizar que todos los miembros de la sociedad se sientan completamente incluidos y verdaderamente iguales.

El caso de *Wisconsin v. Loomis*, según resume Beriain (2018), en febrero de 2013, la policía de Wisconsin detuvo a Eric Loomis cuando conducía un coche que supuestamente estaba involucrado en un tiroteo. Poco después, se declaró culpable de intentar huir de un policía y de conducir un vehículo sin el consentimiento de su propietario, por lo que fue condenado a seis años de prisión.

Lo que llamó la atención en esta sentencia fue que el tribunal de circuito que la dictó reconoció que la duración de la condena se había determinado en parte mediante el uso de una herramienta de evaluación de riesgos desarrollado por Northpointe, Inc., denominada *Correctional Offender Management Profiling for Alternative Sanction (COMPAS)*.

Como se había mencionado en la fundamentación teórica, para predecir el riesgo de reincidencia de una persona, COMPAS utiliza un análisis complejo utilizando datos de encuestas, específicamente de 137 preguntas separadas en varias secciones diferentes e información correspondiente a los antecedentes penales públicos individuales. El Estado de Wisconsin, es uno de los muchos donde se ha fomentado su uso en el ámbito policial.

Ante esta decisión, como se recoge en la sentencia de *Wisconsin v. Loomis* (2016), el acusado decidió apelar ante la Corte Suprema de Wisconsin, siendo el uso de una herramienta de inteligencia artificial, particular de evaluación de riesgos basada en evidencia en la sentencia, el corazón de la cuestión. El peticionario, Eric L. Loomis, apeló la denegación por parte del tribunal de circuito, de su moción posterior a la condena solicitando una audiencia de nueva sentencia.

El tribunal de apelaciones dio por válida la solicitud de apelación, debido a la cuestión específica de si el uso de una evaluación de riesgo de COMPAS en la sentencia viola el derecho de un acusado al debido proceso, ya sea porque la naturaleza de

propiedad de COMPAS impide que los acusados impugnen la validez científica de la evaluación de COMPAS, o porque las evaluaciones de COMPAS tienen en cuenta el género (*Wisconsin v. Loomis*, 2016).

Loomis afirma que la consideración por parte del tribunal de circuito de una evaluación de riesgo de COMPAS en la sentencia viola el derecho del acusado al debido proceso. Además, sostiene que el tribunal de circuito ejerció erróneamente su facultad discrecional al suponer que las bases fácticas de los cargos de lectura eran verdaderas.

Sin embargo, la Corte Suprema (*Wisconsin v. Loomis*, 2016) determinó que debido a que el tribunal de circuito explicó que su consideración de los puntajes de riesgo de COMPAS estaba respaldada por otros factores independientes, su uso no fue determinante para decidir si Loomis podía ser supervisado de manera segura y efectiva en la comunidad. Por lo tanto, el tribunal de circuito no ejerció erróneamente su facultad discrecional. Así, la Corte Suprema llegó a la conclusión de que la consideración por el tribunal de circuito de los cargos de lectura no fue un ejercicio erróneo de discreción porque empleó normas legales reconocidas.

Al parecer, Loomis estaba siendo investigado por su participación en un tiroteo, al cual negó su participación, pero se declaró culpable de otros dos delitos, el intento de huida de un oficial de tránsito, y operar un vehículo motorizado sin el consentimiento del propietario. Después de aceptar la declaración de culpabilidad, el tribunal de circuito ordenó una investigación previa a la sentencia, en la cual el informe de Investigación Previa a la Sentencia (PSI) incluía la evaluación de riesgos de COMPAS.

Como se explica en la sentencia de *Wisconsin v. Loomis*, (2016), en la PSI las puntuaciones de riesgo tienen por objeto predecir la probabilidad general de que las personas con un historial similar de delitos tengan menos o más probabilidades de cometer otro delito después de la puesta en libertad.

Sin embargo, la evaluación de riesgos de COMPAS no predice la probabilidad específica de que un delincuente individual vuelva a delinquir. En su lugar, proporciona una predicción basada en una comparación de información sobre el individuo con un grupo de datos similar. Para ello la evaluación que realiza COMPAS se basa en la información recopilada del expediente penal del acusado y en una entrevista con el

acusado. Así es capaz de evaluar áreas como empleo, vivienda y abuso de sustancias, entre otras, para mostrar los resultados.

Las puntuaciones de riesgo COMPAS de Loomis indicaron que presentaba un alto riesgo de reincidencia en los tres gráficos de barras. Su PSI incluía una descripción de cómo debía utilizarse la evaluación de riesgos de COMPAS y advertía contra su uso indebido, instruyendo a que se utilice para identificar a los delincuentes que podrían beneficiarse de las intervenciones y para abordar los factores de riesgo que deben abordarse durante la supervisión. La PSI también advierte que no se debe utilizar una evaluación de riesgo COMPAS para determinar la severidad de una sentencia o si un delincuente está encarcelado (*Wisconsin v. Loomis*, 2016).

Por lo que, como se puede observar, en todo momento se le pone en conocimiento al juez y al tribunal que no base la sentencia sólo a la evaluación de COMPAS, que sólo es una herramienta que mide ciertos riesgos.

Además, en la sentencia se detalla que no sólo se ha utilizado COMPAS para dictar la sentencia, sino que se han tenido en cuenta diversos factores que descartaron su libertad condicional, como la gravedad del delito, sus antecedentes penales, su historial de vigilancia judicial, y las herramientas de evaluación de riesgos que se han utilizado.

Por lo que Loomis presentó una moción de alivio posterior a la condena solicitando una nueva audiencia de sentencia. Argumentó que la consideración por parte del tribunal de circuito de la evaluación de riesgo de COMPAS en la sentencia violó sus derechos al debido proceso.

Específicamente, Loomis afirma que el uso por parte del tribunal de circuito de una evaluación de riesgo de COMPAS en la sentencia viola el derecho de un acusado al debido proceso por tres razones: viola el derecho de un acusado a ser sentenciado sobre la base de información precisa, en parte porque la naturaleza de propiedad de COMPAS le impide evaluar su exactitud; viola el derecho del acusado a una sentencia individualizada; y utiliza indebidamente evaluaciones de género en las sentencias (*Wisconsin v. Loomis*, 2016).

El acusado ofreció el testimonio de un testigo experto, el Dr. David Thompson, con respecto al uso en la sentencia de una evaluación de riesgo de COMPAS. El Dr.

Thompson opinó que no se debe utilizar una evaluación de riesgos COMPAS para tomar decisiones relativas al encarcelamiento porque no se diseñó una evaluación de riesgos COMPAS para tal uso. Según el Dr. Thompson, la consideración por parte de un tribunal de circuito al dictar sentencia de las partes de evaluación de riesgos de COMPAS corre un tremendo riesgo de sobrestimar el riesgo de un individuo y sentenciarlos erróneamente o basar su sentencia en factores que pueden no aplicarse.

Primero abordaron el argumento de Loomis de que, la consideración por parte de un tribunal de circuito de una evaluación de riesgo de COMPAS viola el derecho al debido proceso de un acusado a ser sentenciado sobre la base de información precisa. Loomis alega inicialmente que la naturaleza de propiedad de COMPAS impide que un demandado impugne la validez científica de la evaluación del riesgo. En consecuencia, Loomis sostiene que debido a que se adjunta una evaluación de riesgo COMPAS a la PSI, al acusado se le niega el pleno acceso a la información de la PSI y, por lo tanto, no puede asegurarse de que se le está condenando sobre la base de información precisa (*Wisconsin v. Loomis*, 2016).

En la sentencia se recoge que Northpointe, Inc., el desarrollador de COMPAS, considera que COMPAS es un instrumento patentado y un secreto comercial. En consecuencia, no revela cómo se determinan las puntuaciones de riesgo ni cómo se ponderan los factores. Loomis afirma que debido a que COMPAS no divulga esta información, se le ha negado información que el tribunal de circuito consideró en la sentencia.

Loomis argumenta que está en la mejor posición para refutar o explicar la evaluación de riesgos de COMPAS, pero no puede hacerlo basándose únicamente en una revisión de las puntuaciones reflejadas en los gráficos de barras. Además, Loomis sostiene que, a menos que pueda revisar cómo se determinan las puntuaciones de riesgo, no se puede verificar la exactitud de la evaluación COMPAS.

Loomis tiene razón en que las puntuaciones de riesgo no explican cómo el programa COMPAS utiliza la información para calcular las puntuaciones de riesgo. Sin embargo, la Corte Suprema (*Wisconsin v. Loomis*, 2016) indica que en la Guía del Profesional de COMPAS de 2015 de Northpointe explica que las puntuaciones de riesgo se basan en gran medida en información estática (antecedentes penales), con un

uso limitado de algunas variables dinámicas (es decir, asociados criminales, abuso de sustancias). El tribunal de circuito y Loomis tuvieron acceso a la misma copia de la evaluación de riesgos. Loomis tuvo la oportunidad de cuestionar sus puntajes de riesgo argumentando que otros factores o información demuestran su inexactitud.

El informe COMPAS adjunto a la PSI de Loomis contiene una lista de 21 preguntas y respuestas con respecto a estos factores estáticos, tales como: ¿Cuántas veces esta persona ha sido devuelta a la custodia mientras estaba en libertad condicional? ¿Cuántas veces esta persona ha tenido un nuevo cargo/arresto mientras estaba en libertad condicional? ¿Cuántas veces ha sido detenida esta persona antes como adulto o menor de edad (sólo detención penal)?

De este modo, en la medida en que la evaluación del riesgo de Loomis se basa en sus respuestas a las preguntas y en los datos públicamente disponibles sobre sus antecedentes penales, Loomis tuvo la oportunidad de comprobar que las preguntas y respuestas enumeradas en el informe COMPAS eran exactas.

En la sentencia (*Wisconsin v. Loomis*, 2016) se menciona que algunos Estados que utilizan COMPAS han llevado a cabo estudios de validación de COMPAS concluyendo que es una herramienta de evaluación de riesgos suficientemente precisa. La División de Servicios de Justicia Penal del Estado de Nueva York realizó un estudio que examinó la eficacia y la precisión predictiva de la escala de reincidencia de una evaluación de COMPAS y concluyó que "la Escala de Reincidencia funcionó eficazmente y logró una precisión predictiva satisfactoria". A diferencia de Nueva York y otros estados, Wisconsin aún no ha completado un estudio de validación estadística de COMPAS para una población de Wisconsin.

Posteriormente, sin embargo, el Departamento de Correcciones y Rehabilitación de California publicó su informe final de 2010 sobre el estudio de validación COMPAS de California. El estudio de 2010 concluyó que, aunque no es perfecto, "COMPAS es un instrumento confiable".

En la sentencia (*Wisconsin v. Loomis*, 2016) se menciona que, específicamente, cualquier PSI que contenga una evaluación de riesgo de COMPAS debe informar al tribunal sentenciador sobre las siguientes precauciones con respecto a la exactitud de

una evaluación de riesgo de COMPAS: (1) se ha invocado la naturaleza de propiedad de COMPAS para evitar la divulgación de información relacionada con cómo se sopesan los factores o cómo se determinarán las puntuaciones de riesgo; (2) la evaluación de riesgos compara a los acusados con una muestra nacional, pero aún no se ha completado ningún estudio de validación cruzada para una población de Wisconsin; (3) algunos estudios de las puntuaciones de la evaluación de riesgos de COMPAS han planteado dudas sobre si clasifican desproporcionadamente a los delincuentes de minorías como personas con un mayor riesgo de reincidencia; y (4) las herramientas de evaluación de riesgos deben ser monitoreadas constantemente y actualizadas para su precisión debido a los cambios en las poblaciones y subpoblaciones.

Al proporcionar esta información a los tribunales sentenciadores sobre las limitaciones y precauciones que conlleva el uso de las evaluaciones de riesgo de COMPAS, permitirá a los tribunales evaluar mejor la precisión de la evaluación y el peso adecuado que debe darse a la puntuación de riesgo.

De manera similar, en la sentencia se recoge que la Guía del Practicante de COMPAS de 2015 explica que la evaluación del riesgo se trata de predecir el comportamiento grupal, no se trata de una predicción a nivel individual. Las escalas de riesgo son capaces de identificar grupos de delincuentes de alto riesgo, no un individuo de alto riesgo en particular. Un ejemplo puntual de posibles malentendidos que surgen del uso de datos grupales es que un individuo que nunca ha cometido un delito violento puede, sin embargo, ser etiquetado como de alto riesgo de reincidencia en la escala de riesgo violento.

En última instancia, la Corte Suprema no está de acuerdo con Loomis porque la consideración de una evaluación de riesgo de COMPAS en el momento de la sentencia, junto con otros factores de apoyo, es útil para proporcionar al tribunal sentenciador la mayor cantidad de información posible para llegar a una sentencia individualizada.

Pasamos ahora a abordar el argumento de Loomis de que el uso del género por parte de una evaluación de riesgos de COMPAS viola los derechos al debido proceso de un acusado.

Loomis afirma que debido a que las puntuaciones de riesgo de COMPAS tienen en cuenta el género, la consideración de un tribunal de circuito de una evaluación de riesgo de COMPAS viola el derecho al debido proceso de un acusado a no ser sentenciado por motivos de género.

Debido a la naturaleza propietaria de COMPAS, en la sentencia las partes disputan el método específico por el cual COMPAS considera el género. Loomis afirma que se desconoce exactamente cómo COMPAS utiliza el género, pero sostiene que COMPAS considera el género como un factor criminógeno. El Estado no está de acuerdo, alegando que el Departamento de Comercio utiliza la misma evaluación de riesgo COMPAS tanto en hombres como en mujeres, pero luego compara a cada infractor con un grupo "normal" de su propio género.

Independientemente de si el género se utiliza como factor criminógeno o únicamente para la normalización estadística, Loomis se opone a cualquier uso del género en el cálculo de las puntuaciones de riesgo de COMPAS. En respuesta, el Estado (*Wisconsin v. Loomis*, 2016) sostiene que es necesario considerar el género en una evaluación de riesgos COMPAS para lograr la precisión estadística. El Estado argumenta que debido a que los hombres y las mujeres tienen diferentes tasas de reincidencia y diferente potencial de rehabilitación, una evaluación de riesgo neutral, en cuanto al género, proporcionaría resultados inexactos tanto para hombres como para mujeres.

Por lo tanto, si la inclusión del género promueve la exactitud, sirve a los intereses de las instituciones y de los acusados, en lugar de a un propósito discriminatorio.

En la sentencia se determinó que el acusado no cumplió con su carga de demostrar que el tribunal sentenciador realmente se basó en el género como un factor en la sentencia. Por lo tanto, concluyeron que el uso de la evaluación de riesgos COMPAS en la sentencia no violó el derecho de Loomis al debido proceso.

Lo que me llamó la atención en la sentencia, es que se explica muy bien las funciones de la evaluación COMPAS, de sus limitaciones y cómo se debe entender.

Por ejemplo, como lo reconoce el Departamento Correccional, la PSI instruye que los puntajes de riesgo no se pueden usar: (1) para determinar si un delincuente está encarcelado; o (2) para determinar la severidad de la sentencia.

Cualquier Informe de Investigación Previa a la Sentencia que contenga una evaluación de riesgo de COMPAS presentado ante el tribunal debe contener un aviso por escrito en el que se enumeren las limitaciones. Además, este asesoramiento por escrito debe informar a los tribunales sentenciadores de las siguientes advertencias, tal como se analizan a lo largo de esta opinión:

- Se ha invocado la naturaleza de propiedad de COMPAS para evitar la divulgación de información relacionada con la forma en que se sopesan los factores o cómo se determinan las puntuaciones de riesgo.

- Debido a que las puntuaciones de la evaluación de riesgos de COMPAS se basan en datos de grupos, pueden identificar grupos de delincuentes de alto riesgo, no un individuo de alto riesgo en particular.

- Algunos estudios sobre las puntuaciones de la evaluación de riesgos de COMPAS han planteado dudas sobre si clasifican desproporcionadamente a los delincuentes pertenecientes a minorías como personas con mayor riesgo de reincidencia.

- Una evaluación de riesgos de COMPAS compara a los acusados con una muestra nacional, pero aún no se ha completado ningún estudio de validación cruzada para una población de Wisconsin. Las herramientas de evaluación de riesgos deben ser monitoreadas y actualizadas constantemente para garantizar su precisión debido a los cambios en las poblaciones y subpoblaciones.

- COMPAS no fue desarrollado para su uso en la sentencia, sino que fue diseñado para ser utilizado por el Departamento Correccional para tomar determinaciones con respecto al tratamiento, la supervisión y la libertad condicional.

En particular, Loomis no argumenta que los otros factores que el tribunal consideró en la sentencia fueran insuficientes para respaldar la sentencia que recibió. De hecho, en los alegatos orales, el abogado de Loomis reconoció que no impugnaría la

sentencia impuesta si careciera de cualquier referencia a la evaluación de riesgos de COMPAS.

Como se discutió anteriormente, si se usa correctamente con conocimiento de las limitaciones y precauciones, la consideración por parte de un tribunal de circuito de una evaluación de riesgo COMPAS en la sentencia no viola el derecho de un acusado al debido proceso. En este caso, el tribunal de circuito era consciente de las limitaciones.

En última instancia, la Corte Suprema concluye que si se usa correctamente como se establece en este documento, la consideración por parte de un tribunal de circuito de una evaluación de riesgo de COMPAS en la sentencia no viola el derecho de un acusado al debido proceso y que el tribunal de circuito no ejerció erróneamente su discreción aquí.

5. CONCLUSIONES

Los algoritmos predictivos pueden ayudar a mantener nuestras sociedades más seguras y reducir la reincidencia. Sin embargo, este objetivo no debe perseguirse a costa de una protección adecuada de los derechos de los acusados, incluido su derecho fundamental a garantías procesales. En este artículo, se han puesto en duda las normas de autenticación sobre procesos automatizados, demostrando que estos son las principales preocupaciones cuando se utilizan como herramientas de apoyo en un proceso judicial.

Además, creo que la única manera realista de lograr este objetivo es dar al acusado el derecho a una explicación de los algoritmos, incluido el acceso a sus códigos fuente. Es evidente que solo los algoritmos de código abierto pueden garantizar la exactitud científica y el respeto adecuado de los derechos involucrados, lo que permite al acusado defenderse.

También hay que resaltar que, en el proceso judicial del caso, como se menciona, en todo momento se le da a conocer el juez el funcionamiento de COMPAS, los datos sobre los que se basa, los riesgos que conlleva la evaluación, y el peso que debería tener a la hora de ser tenido en cuenta en la sentencia.

Con todas estas medidas, considero que el juez es lo suficientemente consciente de los riesgos que conlleva, y que solamente con esa evaluación no es suficiente como para dictaminar una sentencia en base únicamente a ese criterio.

Después del análisis, se me hace indispensable sugerir medidas más estrictas respecto al uso de la inteligencia artificial en el ámbito judicial, como, por ejemplo, excluir las evaluaciones de riesgo que mantienen su metodología en secreto, para así proteger los derechos de los acusados; o restringir su uso hasta que se disponga de más estudios sobre su efectividad y su precisión. De este modo sería una forma más adecuada de contrarrestar los inconvenientes de estas evaluaciones, y asegurarse de que se respetan los derechos de todos.

También me gustaría cuestionar la precisión de la inteligencia artificial, ya que, aunque es capaz de analizar un gran volumen de datos, esto no significa que siempre tenga una precisión exacta, ya que el volumen de datos que tiene que manejar debe estar

siempre actualizado. Por tanto, si no lo está, puede incurrir a informes o evaluaciones erróneas,

Otro de los problemas principales que quiero destacar del uso de las evaluaciones en la predicción de delitos, es la estigmatización del individuo. Como se ha podido observar, COMPAS se basa en datos generales, y hace su evaluación respecto a un grupo de personas con las mismas características, por lo que, si es tenido en cuenta en un juicio, se puede caer en un error enorme al etiquetar a una persona dentro de un grupo, perdiendo así el derecho a un juicio individualizado.

Aunque el caso elegido no ha tenido suceso en Europa, con este análisis se quiere demostrar lo que pasaría si la inteligencia artificial se hubiese implementado en Europa, y se hubiese utilizado en el ámbito judicial. Ha sido utilizado como un pretexto de lo que podría haber sucedido aquí, para regular los fallos que surgen con esta implementación, y así cuando se implante en Europa, que sea más preciso, transparente, y fiable.

Cabe destacar que este trabajo ha sido realizado con una bibliografía en su mayor parte actualizada, lo que indica la veracidad de información y de las ideas de los autores, que están lo más cerca posible de la realidad actual.

Por lo tanto, nuestra conclusión es que los sistemas judiciales deberían abogar por el uso de este tipo de algoritmos, ya que son una herramienta más de apoyo. Sin embargo, aunque esta política no coincida bien con los intereses de los desarrolladores, se debería abogar por la transparencia de los métodos usados para la evaluación, pudiendo darle una oportunidad al acusado de defenderse, al poder observar cómo se ha llegado al resultado final.

Por otro lado, al igual que hacen en Estados Unidos, si el juez solicita el uso de esta herramienta, se le deberá notificar sobre cómo se ha llegado al resultado, sobre los datos sobre los que se ha basado, las limitaciones que tiene, y cómo debe entenderlo el juez, teniendo este último la última palabra.

Así que, desde mi punto de vista, primero habría que regular adecuadamente este fenómeno, y observar cómo influye en la sociedad. Una vez observado, y si existen consecuencias positivas, animo a los gobiernos a invertir y desarrollar estos sistemas de

inteligencia artificial, debido a los grandes beneficios que suponen, con su debido cuidado a la hora de tratar los datos.

5.1 LA AMPLITUD Y LIMITACIONES DE LA INVESTIGACIÓN

La amplitud que tiene este trabajo alcanza la parte fundamental para comprender lo que es y cómo funciona la inteligencia artificial. Si nos centramos en el ámbito de la criminología, se explica gran parte del uso que se le da en el ámbito, aunque no o cubre completamente, debido a que las bases de datos donde he buscado la información no recogen todos los artículos.

Otra barrera que me he encontrado ha sido el idioma, ya que seguramente hubiese encontrado más información si hubiese buscado en otras bases de datos, sin embargo, mi conocimiento me obliga a limitarme a buscar tanto en español como en inglés.

En cuanto a las limitaciones que me he encontrado con el trabajo, el principal es que el caso analizado es un caso de Estados Unidos y no de Europa. Por lo que la jurisprudencia es diferente a la de España y de Europa, lo que dificulta la síntesis del estudio del caso.

Añadiendo que existen pocos casos en los que se use la inteligencia artificial, debido a que es un fenómeno muy reciente que todavía está en desarrollo, y no está implementado en muchos sitios debido a los desafíos que conlleva.

Por último, la falta de regulación ha dificultado su estudio y el impacto que ha tenido en la sociedad, llegando a perderse información valiosa para el trabajo. Aunque si es verdad que se ha podido analizar la nueva ley que entrará en vigor que regula la inteligencia artificial, no será posible estudiar sus efectos sobre los sistemas de IA.

5.2 FUTURAS LÍNEAS DE INVESTIGACIÓN

Para futuras investigaciones, habría que tener en cuenta la nueva ley que sacará el Parlamento Europeo, para estudiar los efectos que tienen sobre los nuevos sistemas de inteligencia artificial, en la sociedad, y en las empresas.

Además, para las futuras investigaciones se deberían estudiar más casos, cuando la inteligencia éste más desarrollada y sea más transparente, para poder realizar una mejor evaluación de su funcionamiento y sobre cómo llegan al resultado. Así se podría estudiar su implantación en Europa y observar cómo evoluciona.

Asimismo, se deberían estudiar casos un poco más actuales, que se asemejen más a la realidad de hoy en día, para que la información sea más creíble y certera, debido a los cambios que pueden haber surgido con la jurisprudencia del país que se estudia. Al igual que se podría estudiar el caso de un solo país, para tener la legislación actual y poder realizar un análisis lo más próximo a la realidad.

6. REFERENCIAS BIBLIOGRÁFICAS

- Afanasyev, A. (2018). Artificial intelligence or intelligence of subjects of detection, disclosure and investigation of crimes: what will win? *Library CSL. Scientific journal*, № 3(38), 28-34.
- Avella, M., & Sanabria-Moyano, J. (2022). Uso del algoritmo COMPAS en el proceso penal y los riesgos a los derechos humanos. *Revista Brasileira de Direito Processual Penal*, 8(1), 275-310.
- Bahamontes, A. (2024). Delitos de inteligencia artificial en España: un panorama en evolución. *TecFuturo*. [04/06/2024]. Disponible en: <https://tecfuturo.es/delitos-de-inteligencia-artificial-en-espana-un-panorama-en-evolucion/>
- Beriain, I. D. M. (2018). Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the Wisconsin v. Loomis ruling. *Law, Probability and Risk*, 17(1), 45-53.
- Berk, R. A. (2021). Artificial intelligence, predictive policing, and risk assessment for law enforcement. *Annual Review of Criminology*, 4, 209-237.
- Bikeev, I., Kabanov, P., Begishev, I., & Khisamova, Z. (2019). Criminological risks and legal aspects of artificial intelligence implementation. In *Proceedings of the international conference on artificial intelligence, information processing and cloud computing*. 20, 1-7.
- Cabrales, J. M. (2024). La inteligencia artificial y la nueva delincuencia. *El Derecho*. [04/06/2024]. Disponible en: <https://elderecho.com/la-inteligencia-artificial-y-la-nueva-delincuencia>
- Casmi, E. (2018). Opinion of millions of people: autonomous cars have to push the elderly and the young to save. *The Network edition "Cnews"*. [16/05/2024]. Disponible en: https://www.cnews.ru/news/top/2018-10-26_mnenie_millionov_chelovek_bespilotnye_avto_dolzhny
- Comisión Europea (2018). Artificial intelligence for Europe. Communication from the commission. COM(2018) 237 final. Brussels 25.4.2018.
- Comisión Europea (2019). Ethics guidelines for trustworthy AI. Comisión Europea. [11/06/2024]. Disponible en: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

c3QlMjBBSSUyMGFuZCUyMFJvYm90aWNzJTIwUmVwb3J0LnBkZg&ntb=1

- ISO (2023). Information technology — Artificial intelligence — Management system. ISO/IEC 42001:2023(E). 1. Disponible en: <https://webstore.iec.ch/publication/90574>
- Jakhar, D., & Kaur, I. (2020). Artificial intelligence, machine learning and deep learning: definitions and differences. *Clinical and experimental dermatology*, 45(1), 131-132.
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685-695.
- King TC, Aggarwal N, Taddeo M et al. (2019) Artificial intelligence crime: an interdisciplinary analysis of foreseeable threats and solutions. *Science and Engineering Ethics*, 26, 89–120.
- Kingston, J. K. (2016) Artificial intelligence and legal liability. *Research and Development in Intelligent Systems XXXIII. Conference Paper*, 269-279. https://doi.org/10.1007/978-3-319-47175-4_20
- Kolenov, S. (2018). AI-oncologist IBM Watson was convicted of medical errors. Network edition“Hightech. plus”. *Alta Tecnología+*. Disponible en: <https://hightech.plus/2018/07/27/ii-onkologa-ibm-watson-ulichili-vo-vrachebnih-oshibkah->
- Kong, B., Supancic, J., Ramanan, D., & Fowlkes, C. (2017). Cross-domain forensic shoeprint matching. Available at: https://www.ics.uci.edu/~fowlkes/papers/KongSRF_BMVC_2017.pdf.
- Lagioia, F., Rovatti, R., & Sartor, G. (2023). Algorithmic fairness through group parities? The case of COMPAS-SAPMOC. *AI & society*, 38(2), 459-478.
- Ligeti, K. (2019). Artificial Intelligence and Criminal Justice. *AIDP-IAPL International Congress of Penal Law*, 1-16.
- Manning, C. (2020). Artificial Intelligence Definitions. Stanford University, Human-Centered Artificial Intelligence. Disponible en: <https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf>
- Marr, B. (2018). The key definitions of artificial intelligence (AI) that explain its importance. *Forbes*. Disponible en:

<https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/>

- Naciones Unidas (2015). Objetivos y metas de desarrollo sostenible. Naciones Unidas. [30/06/2024]. Disponible en: <https://www.un.org/sustainabledevelopment/es/sustainable-development-goals/>
- Nevejans, N. (2016). European civil law rules in robotics. Directorate-General for Internal Policies, Policy Department C: Citizens' Rights and Constitutional Affairs, European Parliament. Disponible en: [https://www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU\(2016\)571379_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU(2016)571379_EN.pdf)
- OCDE. (2019). OCDE AI Principles overview. OCDE AI Policy Observatory Portal. Disponible en: <https://oecd.ai/en/ai-principles>
- Parlamento Europeo (2020). Regulación de la inteligencia artificial en la UE: la propuesta del Parlamento. Parlamento Europeo. [11/06/2024]. Disponible en: <https://www.europarl.europa.eu/topics/es/article/20201015STO89417/regulacion-de-la-inteligencia-artificial-en-la-ue-la-propuesta-del-parlamento>
- Parlamento Europeo (2024). La Eurocámara aprueba una ley histórica para regular la inteligencia artificial. Parlamento Europeo. [11/06/2024]. Disponible en: <https://www.europarl.europa.eu/news/es/press-room/20240308IPR19015/la-eurocamara-aprueba-una-ley-historica-para-regular-la-inteligencia-artificial>
- Pega. (2018). What Consumers Really Think about AI: A global study. Pega. [20/05/2024]. Disponible en: <https://www.pega.com/ai-survey>
- Portal de Administración Electrónica (2023). Ley de Inteligencia Artificial: el Consejo y el Parlamento Europeo llegan a un acuerdo sobre las primeras normas para regular la IA en el mundo. Disponible en: https://administracionelectronica.gob.es/pae_Home/pae_Actualidad/pae_Noticias/Anio2023/Diciembre/Noticia-2023-12-11-Ley-de-inteligencia-artificial-el-Consejo-y-el-Parlamento-Europeo-llegan-a-un-acuerdo-sobre-las-primeras-normas-en-el-mundo-para-regular-la-Inteligencia-Artificial.html
- Portillo, R. A., Montaña, A., & López, J. D. (2022). Modelo predictivo para determinar los retiros de uniformados en el grado de patrullero de la policía nacional de Colombia. Los Libertadores, Fundación Universitaria.

- Puente, M., & Chang, C. (2019). LAPD changing controversial program that uses data to predict where crimes will occur. Los Angeles Times. [27/05/2024]. Disponible en: <https://www.latimes.com/california/story/2019-10-15/lapd-predictive-policing-changes#:~:text=Los%20Angeles%20Police%20Department%20leaders%20announced%20changes%20Tuesday,department%E2%80%99s%20predictive-policing%20program%2C%20called%20PredPol%2C%20helped%20reduce%20crime.>
- Raponi, S., Oligeri, G., & Ali, I. M. (2022). Sound of guns: digital forensics of gun audio samples meets artificial intelligence. *Multimedia tools and applications*, 81(21), 30387-30412.
- Rigano, C. (2019). Using artificial intelligence to address criminal justice needs. *National Institute of Justice Journal*, 280, 1-10.
- Tan, L., & Wang, N. (2010). Future internet: The internet of things. *International Conference on Advanced Computer Theory and Engineering*, 5, 5-376.
- Ulnicane, I. (2022). *Artificial Intelligence in the European Union: Policy, ethics and regulation*. Taylor & Francis.
- UNESCO (2022). *Recommendation on the Ethics of Artificial Intelligence*. [26/06/2024]. Disponible en: <https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence>
- Veale, M., Delacroix, S., Olhede, S., Blacklaws, C., & Adams Bhatti, S. (2019). *Algorithms in the criminal justice system*. The Law Society.
- *Wisconsin v. Loomis*, 881 N.W.2d 749 (2016). Disponible en: <https://cases.justia.com/wisconsin/supreme-court/2016-2015ap000157-cr.pdf?ts=1468415026>
- Zarina I, K., Ildar R, B., & Elina L, S. (2019). Artificial Intelligence and Problems of Ensuring Cyber Security. *International Journal of Cyber Criminology*, 13(2).
- Završnik, A. (2021). Algorithmic justice: Algorithms and big data in criminal justice settings. *European Journal of Criminology*, 18(5), 623-642.