

UNIVERSIDAD NACIONAL DE LA PLATA

FACULTAD DE CIENCIAS JURÍDICAS Y SOCIALES



UNIVERSIDAD
NACIONAL
DE LA PLATA

**¿CONSTRUYENDO PUENTES O BARRERAS? EL PAPEL DE LA
INTELIGENCIA ARTIFICIAL COMO HERRAMIENTA EN EL
ACCESO A LA JUSTICIA.**

SEMINARIO DE GRADO: “PRUEBA PENAL Y ERROR JUDICIAL. EL PROBLEMA DE LAS CONDENAS ERRÓNEAS”.

SECRETARIA DE INVESTIGACIÓN CIENTÍFICA

DIRECCIÓN SEMINARIOS

PRIMER CUATRIMESTRE 2024

DIRECTORA: HUENCHIMAN, VALERIA MARINA

COORDINADOR: DE ANOTONI, ROMAN

ALUMNA: NOUNOU, LAILA ABRIL

N° LEGAJO: 135063/1

E-MAIL: lailaabril20@gmail.com

FECHA DE ENTREGA: 11/02/2024

SUMARIO

El presente estudio analiza y examina los efectos adversos que puedan llegar a ser derivados de la implementación de la IA en las diversas fases del procedimiento penal. Esto enfocado desde el lado de que la inteligencia artificial tiende a, en cierta forma, perpetuar los sesgos humanos de desigualdad y discriminación, contribuyendo así a dejarse de lado humano y fomentando la vulneración para los derechos humanos. Esta situación la veremos enfocada desde un lado más agravado cuando se emplea en el ámbito de la justicia penal, ya que puede generar algoritmos predictivos de riesgo, inexactos o erróneos que resultan en la privación de libertad de individuos inocentes y refuerzan estereotipos/sesgos al determinar las penas. Podremos verlo reflejado, entre otros, a lo largo del trabajo, por ejemplo, en detenciones erróneas en USA por sesgos de algoritmos, como por el algoritmo *COMPAS*, lo que ha puesto en riesgo derechos como el debido proceso, la libertad personal y no discriminación, entre otros.

PALABRAS: Falacias y sesgos en investigaciones penales. Sesgos algorítmicos. Sesgos cognitivos. La inteligencia artificial como herramienta actual.

OBJETIVO

Es evidente que hay una preocupación por el empleo de esta herramienta, y aún no se cuenta con una regulación a nivel internacional ni con suficientes investigaciones sobre los posibles derechos que podrían verse afectados con la utilización de la IA, de los algoritmos predictivos. Por lo tanto, este trabajo tiene como objetivo proporcionar este entendimiento a través de una observación crítica de los principales algoritmos y sus potenciales repercusiones en el proceso penal y su correlato con los derechos humanos. Por ende, uno de los propósitos principales es identificar los efectos potenciales sobre los derechos humanos derivados del uso de algoritmos predictivos de riesgo en el sistema de justicia penal.

INTERROGANTES

¿Qué consecuencias tienen para los derechos humanos el uso de la inteligencia artificial y, en particular, de los algoritmos predictivos de riesgo, en la gestión de la justicia penal?

La IA está plagada de sesgos, ¿Cómo influirían estos a la hora de condenar a las personas? ¿Existe la posibilidad de que estos sistemas automatizados, cuando se les

suministra información sesgada en términos de justicia, luego terminen reproduciendo y amplificando esos sesgos?

En un futuro ¿se debería garantizar acceso abierto a la arquitectura de determinados algoritmos predictivos de la IA, para conocer el curso o mecanismo a través del cual se llega al resultado del nivel de riesgo?

¿Podríamos regular en cierta forma, la llegada de la IA en la administración de justicia penal?

INDICE

I LA INTELIGENCIA ARTIFICIAL	4
II CUESTIONAMIENTOS	9
1) Los sesgos que trae aparejado la inteligencia artificial	9
2) Sesgos algoritmos	9
3) La IA, los sesgos, su impacto en el proceso penal y frente a los DD HH	10
III CASO SOBRE DETENCIÓN ERRONEA EN USA POR LA UTILIZACIÓN DE SESGOS ALGORÍTMICOS	11
→ <i>STATE OF WISCONSIN V. ERIC. L. LOOMIS</i>	11
IV NORMATIVA	13
V CONCLUSIÓN	15
VI BIBLIOGRAFIA	16

RESUMEN EJECUTIVO

¿Qué pasaría si los jueces, en un futuro, tendrían una especie de “colaborador” o “asistente” que les brinda consejos a la hora de condenar a diferentes personas por la comisión de un delito? Es decir, que aconseje cuanto debería durar determinada pena, si la persona tendría posibilidades de reincidir en un futuro o no, entre otras cosas. Y vuelvo a retrucar la pregunta del inicio ¿Qué pasaría si ese polémico “ayudante” se tratase de un algoritmo predictivo?

Esto es la llegada de la Inteligencia Artificial al mundo y que se ve reflejado en diversos contextos de la sociedad, pero cada vez vemos más trascendencia en el sector jurídico, y acá es donde surge la preocupación y donde voy a enfocarme a lo largo de este trabajo de investigación.

Si bien el uso de estos “algoritmos” o de la IA en sí, aportaría un mayor celeridad a los procesos en la administración de justicia y procesos penales, tiene varias controversias que se desarrollarán a lo largo de este trabajo, principalmente estaría implicando una vulneración a los derechos humanos, y a las diferentes garantías del proceso penal, como el debido proceso, derecho a la defensa, etc.

El tipo de controversia o problemática que trae consigo el uso de la IA o algoritmos, es la presencia de “sesgos”. Que a su vez, vamos a ver que, estos llamados sesgos no vienen intrínsecamente con la IA, sino que, como sabemos, la IA se vale de la recopilación de información, referencias y datos, y como tal, estos no están libres de la intervención humana, es decir, de la influencia social, sus implicaciones culturales, etc.

Por tanto, lo que se va dar cuenta en este estudio es que, la tecnología no es neutral, sino que está en manos de los humanos, y la IA puede terminar reforzando estos sesgos, es decir los va a replicar, y refiriéndonos centralmente en lo que es su uso en la administración de justicia, y más específicamente en materia penal en sí, se podría llegar a potenciar aún más las desigualdades, etc.

Este aprendizaje y reproducción de sesgos humanos por parte de la IA, lo lleva claramente a las predicciones erróneas, y se ve reflejado en casos conocidos de detenciones erróneas en USA por sesgos algorítmicos, como por ejemplo el sistema COMPAS. En el presente trabajo, se comentará sobre el caso hito de Eric Loomis.

El trabajo finalizará con mi conclusión final en base a la investigación que pude realizar, -más allá de la poca regulación normativa que tiene y que, como daré cuenta en la

conclusión final, una de las “soluciones” que veo viables es justamente la creación de una normativa específica para ello-.

I – LA INTELIGENCIA ARTIFICIAL

Antes de abordar de lleno los sesgos de la IA, y su uso como herramienta en el mundo jurídico, debería describir la noción básica de la IA.

Existen discrepancias y ninguna definición es absoluta acerca de lo que es la Inteligencia artificial. A.M Turing (1950) pionero en el desarrollo de la inteligencia artificial, se preguntaba “¿pueden las maquinas pensar?” Y señalaba que, la inteligencia artificial podría decirse que describe la posibilidad de que las máquinas, en alguna medida, «piensen», o más bien imiten el pensamiento humano a base de aprender y utilizar las generalizaciones que las personas usamos para tomar nuestras decisiones habituales¹.

Por lo tanto vemos, como volvemos a ver la idea de que, en la IA existe la intervención humana directa o indirectamente, es decir, al “imitar” esos pensamientos humanos por medio de la utilización de los algoritmos predictivos.

Citando al jurista Fenoll (2018) interpreta que las palabras inteligencia artificial y algoritmos van de la mano. Un algoritmo, es el conjunto de instrucciones informáticas que posee una máquina para realizar determinadas tareas mediante el almacenamiento de todas las opciones posibles de decisión según los datos proporcionados.²

Y ahora citando a un escritor y experto en IA, Lasse Rouhiainen, establece una noción sobre lo que podría ser y en la cual la mayoría de los autores coinciden, y es que, «la IA es la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano».³

La inteligencia artificial, especialmente los algoritmos de aprendizaje automático, se aplican en diversas áreas del saber. Estos algoritmos examinan vastas cantidades de información y hacen predicciones adaptadas a las necesidades específicas, siendo utilizados en diferentes ámbitos, aquí hablaremos y nos enfocaremos en el ámbito de la administración de justicia.

¹ A. M. Turing, «Computing machinery and intelligence», Mind, 1950, 49, p. 433.

² Fenoll, J.N. (2018). Cap. I La impronta de la inteligencia artificial. En inteligencia artificial y proceso judicial (pág.21). Editorial Marcial Pons.

³ 1 Lasse P.R. (2018). Cap. I introducción a la inteligencia artificial. En Inteligencia artificial, 101 cosas que debes saber sobre nuestro futuro. (pág.17) Barcelona, España. Editorial Plantea, S.A.

En este último ámbito, la predicción está comenzando a ser muy utilizada en algunas partes del mundo. Se los llama “algoritmos predictivos de riesgo”, estos son patrones matemáticos basados en la automatización del proceso de predicción del comportamiento de un individuo a través de un puntaje, el cual determina su nivel de riesgo y son utilizados en los procesos de toma de decisiones, por ejemplo, para la aplicación de medidas cautelares, o también tras la condena en un proceso judicial.⁴

En el marco del proceso penal, están presentes varios algoritmos predictivos de riesgo, como lo son el algoritmo PSA (usado para determinar si el procesado puede permanecer en libertad hasta la etapa del juicio, teniendo en cuenta el nivel de riesgo de reincidencia o de no comparecencia a la audiencia de juicio), el algoritmo LSI-R (se usa en los permisos de salida y la libertad condicional de los procesados; acá, la ponderación usa criterios como los antecedentes penales, lugar de residencia, educación, empleo, , familia, problemas de alcohol o drogas, actitudes emocionales y personales, etc), el algoritmo HART (su función es predecir el nivel de riesgo de los individuos de cometer delitos en un lapso de dos años, usando criterios como edad, sexo e historial de delitos).

Estos diferentes algoritmos mencionados consideran diversas variables y luego se emplean principalmente para prever decisiones criminales o posibles reincidencias.

Entre todos estos mencionados anteriormente, nos podemos detener un poco más detenidamente, para conocer de qué se trata en el llamado algoritmo *COMPAS (Correctional Offender Management Profiling for Alternative Sanctions/Perfiles de administración de delincuentes correccionales para sanciones alternativas)*. Entendido como “una herramienta estructurada que valora el riesgo de reincidencia del procesado y las necesidades criminológicas del sujeto”.⁵

¿Se podrán predecir los delitos antes de que se produzcan gracias a la IA? Desde 1998 en EE.UU esta pregunta está presente. El algoritmo COMPAS fue desarrollado por la empresa privada Northpointe y lo describe como “un instrumento de evaluación de riesgo de cuarta generación, el cual permite el monitoreo de los individuos a lo largo del tiempo a partir de la información de factores de riesgo dinámicos, diseñados para ayudar en la intervención correccional con el fin de disminuir la probabilidad de que los infractores reincidan”.⁶

⁴ MIRÓ, Fernando. Inteligencia artificial y justicia penal: más allá de los resultados causados por robot. Revista de Derecho Penal y Criminología, n. 20, 2018, p. 107.

⁵ MARTÍNEZ, Lucía. Peligrosidad, algoritmos y due process: el caso state vs loomis. Revista de Derecho Penal y Criminología (uned), n. 20, 2018, p. 485.

⁶ Academia Lab. (2024). COMPAS (software). *Enciclopedia*

El algoritmo COMPAS tiene dos modelos principales de riesgo: riesgo de reincidencia general y riesgo de reincidencia violenta, así mismo este algoritmo cuenta con escalas que miden tanto el riesgo dinámico (factores criminógenos) como riesgo estático (factores históricos).

De hecho, el algoritmo mencionado se está empleando en el sistema penal para evaluar el nivel de riesgo de reincidencia del individuo, “sobre la base de un análisis complejo que implica el uso de la información obtenida de una encuesta de 137 preguntas, dividida en varias secciones diferentes, así como de la información correspondiente a los antecedentes penales individuales”;⁷ además se realiza una entrevista al procesado con el fin de obtener toda la información necesaria para realizar la ponderación y determinar un puntaje del nivel de riesgo de reincidencia.

Luego que se realiza esa evaluación, se obtiene un puntaje de riesgo del 1 al 10. Esto a su vez tiene los tipos de escala de cada nivel de riesgo. Algunas de las variables utilizadas por el algoritmo para establecer el grado de riesgo de reincidencia están basadas en información privada y subjetiva en diferentes escalas, como “criminalidad familiar”, “inestabilidad familiar”, “problemas financieros” “abuso de sustancias”, etc. Por ejemplo, utilizan también escalas de ocio o fracaso de socialización, lo cual podría considerarse discriminatorio al tenerse en cuenta criterios ajenos a un riesgo real de reincidencia.

Además existe en este algoritmo una puntuación para cada tipo de las escalas antes mencionadas, entre otras. Se determinará una puntuación del nivel de reincidencia del procesado. Northpointe en su guía establece que hay dos tipos de escalas, la de necesidades y la de riesgo, la primera no es predictiva, sino que permite describir de forma sencilla al procesado.

También se clasifican dichas escalas en: reincidencia general, reincidencia violenta y la mala conducta previa al juicio; no obstante, la guía carece de especificidad en cuanto a la descripción del funcionamiento del algoritmo, por lo cual el COMPAS se comporta como un algoritmo de caja negra, lo cual implica que su arquitectura no permite conocer su funcionamiento ni saber cómo influyen verdaderamente esas escalas en la determinación del nivel de riesgo de reincidencia, por ende puede afirmarse que hay opacidad en el algoritmo.

⁷ ROMEO, Carlos. Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad. Revista de Derecho, Empresa y Sociedad (REDS), n. 13, 2018, p. 49

Cabe la pena preguntarse, si los eventuales sesgos de los algoritmos predictivos de riesgo, podrían incentivar lo que Silva (1999) ha llamado la teoría de la *inocuidación*, de acuerdo con la cual “resulta posible individualizar a un número relativamente pequeño de delincuentes, respecto de los cuales cabe determinar que han sido responsables de la mayor parte de hechos delictivos y predecir, sobre la base de criterios estadísticos, que lo seguirán siendo. De este modo, se entiende que la inocuidación de los mismos, esto es, su retención en prisión el máximo tiempo posible conseguiría una radical reducción del número de hechos delictivos y, con ello, importantes beneficios al menor coste.⁸

La inclinación hacia esa inocuidación es inquietante, ya que podría generar un ciclo perjudicial que fomente la exclusión de grupos históricamente discriminados, poniendo en peligro garantías fundamentales del sistema judicial, como la presunción de inocencia, o incluso el principio del derecho penal de acto, retrocediendo hacia concepciones propias del derecho penal de autor.

Y por último en esta pequeña introducción a lo que es la inteligencia artificial voy a comentar sobre las 2 técnicas de aprendizaje de la IA.

Por un lado el aprendizaje *automático* y por el otro el aprendizaje *profundo*.

Según L. Rouhiainen entiende al aprendizaje profundo como: *Un subcampo del aprendizaje automático que se utiliza para resolver problemas muy complejos y que normalmente implican grandes cantidades de datos. El aprendizaje profundo se produce mediante el uso de redes neuronales, que se organizan en capas para reconocer relaciones y patrones complejos en los datos. Su aplicación requiere un enorme conjunto de información y una potente capacidad de procesamiento. Actualmente, se utiliza en el reconocimiento de voz, el procesamiento del lenguaje natural, la visión artificial y la identificación de vehículos en los sistemas de asistencia al conductor.*⁹

Y Sandoval entiende al aprendizaje automático como: *Una rama de la inteligencia artificial que genera algoritmos que tienen la capacidad de aprender. El desarrollador no tiene que imaginar todas las alternativas posibles, solo alimenta al algoritmo con un volumen gigante de datos para el que el mismo aprenda y sepa que hacer en cada caso.*

⁸ SILVA, Jesús. La expansión del derecho penal aspectos de la política criminal en las sociedades postindustriales, 1999, p. 144.

⁹ Lasse P.R. (2018). Cap. I introducción a la inteligencia artificial. En Inteligencia artificial, 101 cosas que debes saber sobre nuestro futuro. (pág.22) Barcelona, España. Editorial Plantea, S.A.

A su vez, explica que: existen dos tipos de aprendizajes: el *supervisado* (cuando entrenamos un algoritmo dándole las preguntas –características- y las respuestas –etiquetas- y luego el algoritmo realiza una predicción conociendo las características y etiquetas), y el *no supervisado* (solo le damos las características al algoritmo, nunca las etiquetas. Agrupa los datos que le dimos según sus características. El algoritmo solo sabe que como los datos comparten ciertas características, de esa forma asume que pueda que pertenezcan al mismo grupo).¹⁰

Finalmente, concluyo esta parte del trabajo, dando cuenta que, en resumen, la inteligencia artificial examina datos en busca de patrones únicos, lo que le permite diferenciar entre diferentes categorías. Por lo tanto, los datos son esenciales en la IA, ya que son la base de información que determina si las conclusiones son correctas o incorrectas. Sin embargo, los datos suelen ser una representación limitada de la realidad, lo que puede llevar a confusiones en la inteligencia artificial. Además, es crucial que los datos estén libres de sesgos, aunque esto es difícil de lograr debido a la influencia humana en su creación.

II- CUESTIONAMIENTOS

¿Qué son estos *sesgos* que vengo mencionando en lo que va del trabajo?

1) Los sesgos que trae aparejado la inteligencia artificial:

Los errores y sesgos humanos se generan de manera inconsciente cuando se utiliza la mente para procesar la información que proviene del exterior y simplificarla para la toma de decisiones. Los procedimientos o reglas heurística mentales de simplificación son útiles para su uso cotidiano, aunque pueden generar errores y sesgos.¹¹

2) Sesgos algorítmicos:

Vamos a enfocarnos en un tipo de sesgo en específico, y es el “sesgo algorítmico”. Como he mencionado en lo que va de este estudio, sabemos que los datos son cruciales en la IA, ya que estos constituyen la información fundamental que alimenta el sistema y sobre la cual el algoritmo evalúa la precisión de sus conclusiones. Sin embargo, el

¹⁰ 4 Sandoval, L.J. 2018. Algoritmos de aprendizaje automático para el análisis y predicción de datos. Revista ITCA-FEPADE, Revista Tecnológica N°11 2-4.

¹¹ Muñoz Aranguren A. (2011). La influencia de los sesgos cognitivos en las decisiones jurisdiccionales: el factor humano. Una aproximación. Barcelona, España. Revista Catalnes. Indret.

problema o cuestionamiento que vamos a abordar es cuando esas conclusiones están influenciadas por sesgos.

Lo subjetivo y la intervención humana no es algo que pueda no aparecer en la IA. Lo subjetivo cuando se desarrolla el sistema, y la intervención humana en lo que vimos en el apartado I en los tipos de aprendizaje de la IA, donde aparecen los datos aportados por los seres humanos, que claramente están sesgados, y muchas veces podrían contener desigualdades y datos discriminatorios, ya que tienen su propia perspectiva sobre el mundo basadas en prejuicios, evaluaciones basadas en sus vivencias, etc.

3) La IA, los sesgos, su impacto en el proceso penal y frente a los Derechos humanos:

Una de las deficiencias que posee la IA en el proceso penal, y que es muy preocupante, según mi punto de vista, sería, la incertidumbre que nace sobre los DD HH ante la IA como herramienta en un proceso.

El creciente empleo de algoritmos predictivos dentro del sistema judicial, particularmente en el ámbito penal, genera inquietudes sobre posibles violaciones de los derechos humanos.

De acuerdo a la Declaración Universal de Derechos Humanos "*todos los seres humanos nacen libres e iguales en dignidad y derechos*"¹². Y lo establecido en el Pacto Internacional de Derechos Civiles y Políticos en su artículo 14 "*(...) todas las personas son iguales ante los tribunales y cortes de justicia (...)*"¹³.

Luego en los Principios y directrices de las Naciones Unidas sobre el acceso a asistencia jurídica en los sistemas de justicia penal, se establece que: "los Estados deben garantizar la prestación de asistencia jurídica a todas las personas, independientemente de su edad, raza, color, sexo, idioma, religión o creencia, opinión política o de otra índole, origen nacional o social, patrimonio, ciudadanía o domicilio, nacimiento, instrucción o condición social, o de cualquier otra condición"¹⁴.

Luego de tener en claro lo mencionado anteriormente, recordemos lo que hablamos en el apartado n° I en cuanto al algoritmo COMPAS y sus tipos de escalas. Estaríamos en presencia de un claro desconocimiento al derecho a la igualdad y no discriminación, en tanto se valorarían criterios tales como la situación financiera, el

¹² NACIONES UNIDAS. La Declaración Universal de Derechos Humanos.

¹³ NACIONES UNIDAS. Pacto Internacional de Derechos Civiles y Políticos.

¹⁴ NACIONES UNIDAS. Principios y directrices de las Naciones Unidas sobre el acceso a la asistencia jurídica en los sistemas de justicia penal,

desempleo, la pobreza, entre otras. Es decir, lo podemos ver claro con el siguiente ejemplo: Si un acusado tiene un estatus económico más alto que otro, es probable que reciba un puntaje más bajo en la escala de riesgo de reincidencia, lo que mostraría una disparidad en el tratamiento de los acusados al considerar sus recursos financieros.

Siguiendo con otro derecho que podría estar en riesgo y su impacto en el proceso penal, el derecho al debido proceso se define como "un derecho humano, civil y político, que engloba un conjunto de garantías procesales que protegen a la persona sometida a un procedimiento judicial"¹⁵. Más adelante, en el caso de *Wisconsin vs. Loomis*, veremos que uno de los principales argumentos de la defensa fue la supuesta violación de este derecho, basada en la incapacidad para impugnar el funcionamiento del algoritmo utilizado para determinar el nivel de riesgo.

En fin, la ausencia de comprensión acerca de las operaciones internas de una herramienta de aprendizaje automático dificulta la detección y comprobación de posibles violaciones legales. Esto es especialmente problemático en el caso de leyes que protegen los derechos fundamentales, ya que estas leyes atribuyen responsabilidades y permiten reclamar compensaciones. La transparencia también está vinculada a cómo se crean los algoritmos, así como a lo que se conoce como el desarrollo ético de los mismos.

Los algoritmos predictivos se comportan como una caja negra y los sesgos que los pueden integrar, ponen en riesgo derechos y garantías procesales.

III. CASO SOBRE DETENCIÓN ERRONEA EN USA POR LA UTILIZACIÓN DE SEGOS ALGORÍTMICOS

En este apartado desarrollare un caso hito de Estados Unidos, donde fue utilizado el algoritmo antes mencionado, el COMPAS.

Hablaremos del caso **State Wisconsin vs Loomis**¹⁶ y se cataloga como hito ya que se entiende que fue la primera vez que un tribunal se pronunció sobre la admisibilidad del uso de herramientas de IA dentro del proceso penal.

¹⁵ ANAYA, Alejandro. Glosario de términos básicos sobre derechos humano.

¹⁶ Harvard Law Review. (2017). *State v. Loomis*. Volume 130. 130 Harv. L. Rev. 1530.

¹⁷En este caso, Loomis enfrentó acusaciones por cinco delitos, relacionados con su participación en un tiroteo, incluyendo cargos por posesión de arma de fuego, intento de fuga de un oficial de tránsito, puesta en peligro de la seguridad pública y conducción de un vehículo sin autorización. Aunque negó su participación en el tiroteo, admitió haber conducido el vehículo y haber intentado escapar del oficial. Finalmente, llegó a un acuerdo con la Fiscalía en el que aceptó los cargos de puesta en peligro de la seguridad pública y conducción de un vehículo sin autorización.

La Fiscalía presentó un informe basado en el algoritmo COMPAS, el cual indicaba que Loomis tenía un alto riesgo de reincidencia. El juez de primera instancia, considerando este informe entre otros factores, impuso una sentencia de 6 años de prisión y 5 años de libertad condicional.

El abogado de Loomis apeló la decisión ante la Corte Suprema, argumentando que violaba el derecho al debido proceso por tres razones: (1) al infringir el derecho del acusado a ser sentenciado con información precisa, debido a la naturaleza patentada de COMPAS, que dificultaba evaluar su precisión; (2) al no proporcionar una sentencia individualizada; y (3) al utilizar evaluaciones de género de manera inapropiada en las sentencias.

Respecto de COMPAS, “al ser un instrumento patentado que cuenta con secreto comercial, no revela cómo se determinan las puntuaciones de riesgo o cómo se ponderan los factores. En ese sentido Loomis afirmó que no tuvo acceso a información que el Tribunal de Circuito consideró al momento de dictar la sentencia”.

Por otro lado, la Corte precisó que a pesar de que “el programa COMPAS no explica cómo usa la información para calcular los puntajes de riesgo, la Guía para profesionales de COMPAS de 2015 de Northpointe, explica que las puntuaciones se basan en gran medida en información estática (antecedentes penales), con un uso limitado de algunas variables dinámicas.

El segundo argumento de la defensa estableció que COMPAS vulneraba el derecho a una sentencia individualizada, ya que el software trabaja con base en estadísticas grupales, sobre este punto la Corte recalcó que “las evaluaciones de riesgo son junto otros factores de apoyo útiles para proporcionar la mayor cantidad de información posible para llegar a una sentencia individualizada”; [...] sin embargo las

¹⁷ Arienza, A. F. (2020). Comentario a la sentencia la Corte de apelaciones en el caso «State of Wisconsin v. Eric L. LOOMIS». Del Tribunal Supremo del Estado de Wisconsin. Derecho Digital e Innovación. Digital Law and Innovation Review, (5), 10.

-*State v. Loomis, 881 N.W.2d 749 (Wis. 2016). <https://www.courts.ca.gov/documents/BTB24-2L-3.pdf>

implicaciones del debido proceso, generan la necesidad de advertir a los Tribunales de Circuito que los puntajes de evaluación de riesgos de COMPAS se basan en datos grupales, que pueden identificar grupos de delincuentes de alto riesgo, no un individuo de alto riesgo en particular.

Por último, el defensor del Sr. Loomis estableció que el uso del software tomaba en cuenta el género, argumentando que el acusado no puede ser sentenciado basándose en su género; la Corte sostuvo que es necesario considerar el género en una evaluación de riesgo para lograr precisión estadística, además, 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.

Adicionalmente la Corte puntualizó que el uso del género por parte de COMPAS promueve la precisión que, en última instancia, redundaría en beneficio del sistema de justicia. Finalmente, la Corte concluyó que el acusado no demostró que la sentencia utilizó el género como un factor decisorio en la sentencia.

Con este breve resumen sobre el fallo de la Suprema Corte de Wisconsin damos cuenta de la inseguridad jurídica que genera, y la aparición de sesgos por razón de raza y discriminación por género.

IV NORMATIVA

Hasta el momento, no existe una regulación legal amplia sobre los algoritmos predictivos de riesgo ni la IA.

Conectándolo con las sugerencias que realizó el Tribunal S. de Wisconsin en el fallo anteriormente citado, podemos hablar de los principios que se establecieron en la Carta de Recomendación sobre la Ética de la Inteligencia Artificial.¹⁸

Se plantea la importancia de los principios de equidad y no discriminación en el contexto de la inteligencia artificial, instando a los agentes implicados a adoptar un enfoque inclusivo para prevenir la perpetuación de aplicaciones y resultados discriminatorios o sesgados a lo largo del ciclo de vida de los sistemas de IA. Esto se hace con el objetivo de asegurar la equidad en dichos sistemas.

¹⁸ 4 La Conferencia General de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). (2021) Recomendación sobre la Ética de la Inteligencia Artificial.

Además, se destaca la relevancia de los principios de transparencia y explicabilidad para garantizar el respeto, la protección y la promoción de los derechos humanos y los principios éticos en el desarrollo y uso de la inteligencia artificial. Se advierte que la falta de transparencia puede dificultar impugnar las decisiones basadas en los resultados de los sistemas de IA, lo que podría infringir el derecho a un juicio imparcial y a un recurso efectivo, y restringir los contextos legales en los que estos sistemas pueden emplearse.

Por otro lado, podemos mencionar al proyecto que aprobó el Consejo de Europa¹⁹ para regular a la IA, que tiene un resguardo de error.

La idea principal del proyecto “Reglamento de Inteligencia Artificial” es regular la IA en la medida que esta tenga la capacidad de dañar a la sociedad, siguiendo un enfoque basado en los riesgos: **a mayor riesgo, normas más estrictas.** Puede establecer un referente mundial para regular la IA en otras jurisdicciones, dado que es una de las primeras propuestas legislativas en el mundo.

La presidencia del Consejo y los negociadores del Parlamento Europeo en el acuerdo provisional sobre la propuesta relativa a normas armonizadas en materia de IA, tiene como principales elementos lo siguiente²⁰:

- normas sobre modelos de IA de uso general de gran impacto que pueden causar un riesgo sistémico en el futuro, así como sobre los sistemas de IA de alto riesgo;
- un sistema revisado de gobernanza con algunas competencias de ejecución a escala de la UE;
- ampliación de la lista de prohibiciones, pero con la posibilidad de utilizar la identificación biométrica remota por parte de las autoridades policiales en espacios públicos, con sujeción a salvaguardias;
- una mejor protección de los derechos mediante la obligación de que los implementadores de sistemas de IA de alto riesgo lleven a cabo una evaluación del impacto en los derechos fundamentales antes de poner en marcha un sistema de IA.

¹⁹ Reglamento de Inteligencia Artificial, orientación general del Consejo, 6.12.2022.

²⁰ Consejo de la UE. Comunicado de prensa. <https://www.consilium.europa.eu/es/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/>

V CONCLUSIÓN

A lo largo de este trabajo, se ha observado cómo los algoritmos predictivos han incursionado en diversos campos del conocimiento, incluido el ámbito legal y penal, o al menos es lo que se viene, o en muchas partes del mundo, como EE UU, se ha empezado a utilizar.

En particular, vimos cómo el algoritmo predictivo COMPAS, usa en la evaluación del riesgo de reincidencia, criterios subjetivos como la situación económica, la salud mental, el tiempo libre, el rendimiento académico y la actividad delictiva de amigos y familiares. Esta inclusión de factores subjetivos evidencia una discriminación algorítmica, lo que plantea riesgos para los derechos humanos como la no discriminación, la libertad personal, el debido proceso y el acceso a la información transparente, entre otros.

Además, trate de narrar y evidenciar impactos negativos en los derechos humanos involucrados en procesos judiciales debido al uso de algoritmos predictivos de riesgo. La falta de transparencia en estos algoritmos, que operan como cajas negras debido a la opacidad en el proceso de ponderación de variables y al secreto empresarial, impide que los afectados accedan y cuestionen estas operaciones. Esta falta de transparencia afecta el ejercicio adecuado de derechos como el derecho a la información, a la contradicción y a un juicio imparcial, etc. Además, en el diseño, programación y entrenamiento de los algoritmos, los sesgos presentes en los datos utilizados pueden reflejar los prejuicios del diseñador del algoritmo y/o perpetuar los estereotipos presentes en la sociedad.

Y por último, concluyo en que la falta de regulación en relación con el diseño y la implementación de la inteligencia artificial en el ámbito judicial ha resultado en violaciones de los derechos humanos (o por lo menos es lo que se podría ocasionar si esto sigue avanzando cada vez más) especialmente porque los funcionarios judiciales carecen de protocolos o normativas vinculantes que establezcan cómo aplicar adecuadamente la inteligencia artificial en la toma de decisiones. Se requiere una regulación precisa y completa para abordar esta situación. Además, es fundamental destacar que el uso apropiado de los algoritmos predictivos de riesgo exige un diseño ético que evite sesgos en los datos utilizados, tanto en su creación como en las etapas de entrenamiento, validación y pruebas.

VI BIBLIOGRAFÍA

- A. M. Turing, «Computing machinery and intelligence», *Mind*, 1950, 49, p. 433.
- Academia Lab. (2024). COMPAS (software). Enciclopedia
- ANAYA, Alejandro. Glosario de términos básicos sobre derechos humano.
- Fenoll, J.N. (2018). Cap. I La impronta de la inteligencia artificial. En *inteligencia artificial y proceso judicial* (pág.21). Editorial Marcial Pons.
- Harvard Law Review. (2017). *State v. Loomis*. Volume 130. 130 Harv. L. Rev. 1530.
- La Conferencia General de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). (2021) Recomendación sobre la Ética de la Inteligencia Artificial.
- Lasse P.R. (2018). Cap. I introducción a la inteligencia artificial. En *Inteligencia artificial, 101 cosas que debes saber sobre nuestro futuro*. (pág.17) Barcelona, España. Editorial Plantea, S.A.
- MARTÍNEZ, Lucía. Peligrosidad, algoritmos y due process: el caso state vs loomis. *Revista de Derecho Penal y Criminología (uned)*, n. 20, 2018, p. 485.
- MIRÓ, Fernando. Inteligencia artificial y justicia penal: más allá de los resultados causados por robot. *Revista de Derecho Penal y Criminología*, n. 20, 2018, p. 107.
- Muñoz Aranguren A. (2011). La influencia de los sesgos cognitivos en las decisiones jurisdiccionales: el factor humano. Una aproximación. Barcelona, España. *Revista Catalnes. Indret*.
- NACIONES UNIDAS. La Declaración Universal de Derechos Humanos.
- NACIONES UNIDAS. Pacto Internacional de Derechos Civiles y Políticos.
- NACIONES UNIDAS. Principios y directrices de las Naciones Unidas sobre el acceso a la asistencia jurídica en los sistemas de justicia penal.
- Reglamento de Inteligencia Artificial, orientación general del Consejo, 6.12.2022.
- ROMEO, Carlos. Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad. *Revista de Derecho, Empresa y Sociedad (REDS)*, n. 13, 2018, p. 49
- Sandoval, L.J. 2018. Algoritmos de aprendizaje automático para el análisis y predicción de datos. *Revista ITCA-FEPADE, Revista Tecnológica N°11 2-4*.

SILVA, Jesús. La expansión del derecho penal aspectos de la política criminal en las sociedades postindustriales, 1999, p. 144.