



Revista Brasileira de Direito Processual Penal

ISSN: 2525-510X

Instituto Brasileiro de Direito Processual Penal

Avella, Marcela del Pilar Roa; Sanabria-Moyano, Jesús Eduardo; Dinas-Hurtado, Katherin  
Uso del algoritmo COMPAS en el proceso penal y los riesgos a los derechos humanos<sup>1</sup>  
Revista Brasileira de Direito Processual Penal, vol. 8, núm. 1, 2022, Enero-Abril, pp. 275-310  
Instituto Brasileiro de Direito Processual Penal

DOI: <https://doi.org/10.22197/rbdpp.v8i1.615>

Disponible en: <https://www.redalyc.org/articulo.oa?id=673971913008>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org



Sistema de Información Científica Redalyc  
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso  
abierto


# Uso del algoritmo COMPAS en el proceso penal y los riesgos a los derechos humanos<sup>1</sup>

*Use of the COMPAS algorithm in the criminal process and the risks to human rights*

**Marcela del Pilar Roa Avella<sup>2</sup>**

Universidad Militar Nueva Granada – Bogotá, Colombia


marcela.roa@unimilitar.edu.co

 <https://orcid.org/0000-0001-6307-3627>

**Jesús Eduardo Sanabria-Moyano<sup>3</sup>**

Universidad Militar Nueva Granada – Bogotá, Colombia

jesus.sanabria@unimilitar.edu.co

 <https://orcid.org/0000-0001-9715-5305>

- 
- <sup>1</sup> Este artículo es producto del proyecto INV DER 3159 “Inteligencia Artificial: retos y riesgos de los Derechos Humanos en el Sistema Penal”, de la línea de investigación en derecho penal y justicia militar, del Grupo de Derecho Público, que se adelanta en el Centro de Investigaciones Jurídicas, Políticas y Sociales de la Facultad de Derecho de la Universidad Militar Nueva Granada, financiado por la Vicerrectoría de Investigaciones de la Universidad Militar Nueva Granada, dentro de la convocatoria Proyectos de Investigación Científica vigencia 2020.
  - <sup>2</sup> Profesora e investigadora del grupo “Derecho Público” y de la línea “Derecho Penal y Justicia Militar” del Centro de Investigaciones Jurídicas, Políticas y Sociales de la Facultad de Derecho, Bogotá de la Universidad Militar Nueva Granada. Abogado por la Universidad del Rosario (Colombia). Especialista en Ciencias Penales y Ciencias Penales de la Universidad Externado (Colombia). Magister en derecho penal de la Universidad de Santo Tomás en convenio con la Universidad de Salamanca (Colombia).
  - <sup>3</sup> Profesor e investigador del grupo “Derecho Público” y de la línea “Derecho Internacional, Derechos Humanos y Derechos Humanos Internacionales” del Centro de Investigaciones Jurídicas, Políticas y Sociales de la Facultad de Derecho, Bogotá de la Universidad Militar Nueva Granada. Abogado por la Universidad Militar Nueva Granada (Colombia). Especialista en Derechos Humanos y Derecho Internacional Humanitario aplicado a los Conflictos Armados de la Facultad de Derechos Humanos y Derecho Internacional Humanitario del Ejército Nacional (Colombia). Magister en Derecho Público Militar de la Universidad Militar Nueva Granada (Colombia).

---

**RESUMEN:** En la actividad policial, administración de justicia y procesos penales, se han empezado a utilizar diversos algoritmos predictivos de riesgo, los cuales aportan celeridad en la toma de decisiones. Han surgido controversias acerca de su aplicación, debido a ciertas variables subjetivas ponderadas, lo que implicaría una posible vulneración para los derechos humanos. En el presente artículo se describen los principales algoritmos predictivos de riesgo, haciendo énfasis en el algoritmo COMPAS y se determinan los impactos a los derechos humanos derivados de su uso. El método utilizado es deductivo-inductivo, partiendo de un análisis documental para describir el funcionamiento de los algoritmos y un proceso inductivo para determinar si se garantizan los derechos humanos con su utilización. Los resultados indican que en efecto sí hay riesgos cuando se utilizan estos sistemas inteligentes predictivos, los cuales son derivados de la falta de transparencia de la arquitectura del algoritmo y de la presencia de sesgos. Lo anterior impediría identificar el proceso por el cual se llega a la calificación de nivel de riesgo, de un lado, y de otro, una aplicación objetiva de la evaluación; lo que pondría en riesgo derechos como el debido proceso, la libertad personal y no discriminación, entre otros.

**PALABRAS CLAVE:** Inteligencia artificial; justicia penal; algoritmos; derechos humanos.

**ABSTRACT:** *Diverse risk assessment algorithms used in policing, justice administration and criminal proceedings have contributed to promptness in decision making. Controversies have emerged regarding their enforcement, due to certain subjective weighted variables that may lead to a possible breach of human rights. This article describes the main risk assessment*

---

<sup>4</sup> Joven investigadora de la Universidad Militar Nueva Granada del Grupo Derecho público y línea de derecho penal y justicia militar. Abogada por la Universidad Militar Nueva Granada (Colombia).

*algorithms, emphasizing the COMPAS algorithm and determining its impact on human rights. The method used is deductive-inductive, on the basis of a documentary analysis that describes the operation of algorithms, and an inductive process used to determine if human rights are guaranteed while it is used. Results indicate that, in fact, risks exist at the time of using intelligent predictive systems, derived from a lack of transparency in the design of the algorithm and bias. The above would not allow the identification of the process used to classify the risk level and the unbiased application of the evaluation, jeopardising rights like due process, freedom and non-discrimination, among others.*

**KEYWORDS:** Artificial intelligence; criminal justice; algorithms; human rights.

---

## INTRODUCCIÓN

En el desarrollo de la denominada era digital, se ha evidenciado la creciente presencia de herramientas derivadas de la inteligencia artificial (en adelante IA) en el derecho, lo cual se puede ver materializado en la administración de justicia, en la actividad policial y en los procesos penales, en donde se han empezado a utilizar diversos algoritmos predictivos de riesgo, como lo son el PSA, el LSI-R, HART, CAS, y COMPAS. Según Linares son potencialmente aplicables a los procesos de toma de decisiones relacionados con la valoración del riesgo de aquellos individuos que se ven envueltos en el proceso judicial<sup>5</sup>, y su función es determinar el grado de riesgo de reincidencia del individuo en diferentes etapas del proceso penal; no obstante algunos algoritmos también son usados en actividades policiales.

A pesar de que los mencionados algoritmos son herramientas que aportan celeridad a la justicia y precisión en la toma de decisiones, han surgido controversias acerca de su aplicación, debido a que ponderan criterios tales como el comportamiento social, nivel educativo, entre otros; lo que implicaría una posible vulneración para los derechos humanos

---

<sup>5</sup> 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.

como son, los derechos a la igualdad, no discriminación, debido proceso, libertad personal, honra, dignidad, entre otros.

En el mismo orden de ideas, se ha establecido que hay “algoritmos secretos” de predicción de niveles de riesgo, respecto de los cuales no es posible conocer la forma exacta cómo funciona el algoritmo ya que son de propiedad exclusiva de su creador; y ello pondría en riesgo, derechos tales como el de contradicción, de defensa o el derecho a la información.

Así mismo ProPublica<sup>6</sup> realizó un estudio sobre el algoritmo COMPAS en el que advirtió la probabilidad de una discriminación algorítmica hacia los individuos afrodescendientes, lo que conllevaría a un posible riesgo de vulneraciones de los derechos humanos antes señalados.

Se ha determinado que los algoritmos predictivos de riesgo contienen posibles sesgos al ser aplicados. En ese sentido encontramos el caso hito: *State Wisconsin vs Loomis* en donde fue usado el algoritmo COMPAS y se evidenció inseguridad jurídica, sesgos por razones de raza, discriminación por género, entre otros.<sup>7</sup> Así mismo Rizer & Watney establecieron que los algoritmos no son completamente neutrales, no son herramientas objetivas debido a que contienen una conservación inadecuada de datos.<sup>8</sup> Si bien es cierto que se evidencia una preocupación por el uso de estas herramientas, no existe aún una regulación internacional, o suficientes estudios acerca de los eventuales derechos vulnerados con el uso de estos algoritmos predictivos; es así como en el presente artículo se aportará este nuevo conocimiento, mediante un análisis crítico de los principales algoritmos predictivos y los eventuales impactos para los derechos humanos.

Al observarse la problemática anteriormente mencionada surge la pregunta problema de investigación a resolver en el presente artículo: ¿cuáles son los impactos a los derechos humanos derivados de la utilización

<sup>6</sup> ProPública es una sala de prensa independiente sin fines de lucro que produce periodismo de investigación con fuerza moral, <https://www.propublica.org/>

<sup>7</sup> 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. 10.

<sup>8</sup> RIZER, Arthur; WATNEY Caleb. Artificial Intelligence Can Make Our Jail System More Efficient, Equitable and Just. *Texas Review of Law & Politics*, Austin, v. 23, n. 1, 2018, p. 15. (traducción libre).

de algoritmos predictivos de riesgo en la administración de justicia penal, y en especial el algoritmo COMPAS? Por tanto, el objetivo general consiste en establecer los posibles impactos a los derechos humanos por la utilización de algoritmos predictivos de riesgo en la administración de justicia penal.

Para dar cumplimiento a este objetivo se hará uso de una metodología cualitativa de tipo básico jurídico y multidisciplinaria, porque integra sistemas de análisis de las ciencias computacionales, que permitirán un alcance descriptivo donde se identificarán y analizarán los elementos de los algoritmos el PSA, el LSI-R, HART, CAS, NDAS y COMPAS, que son utilizados para la cuantificación del riesgo de actividades delictivas; por lo cual, el método utilizado será de tipo deductivo-inductivo, debido a que primero se realizará el análisis documental de fuentes primarias, secundarias y terciarias, que describen el funcionamiento de los algoritmos, para deducir en ellos los elementos sobre los cuales podrían generarse conflictos frente a los derechos humanos. Para lo anterior será necesario identificar aquellos que se encuentran inmersos en el objeto y fin de la utilización de los algoritmos, aplicando entonces un proceso de inducción que logre contrastar el respeto y garantía del pleno goce y ejercicio de estos derechos humanos, y las garantías procesales en la utilización de los algoritmos.

Por último, los resultados se presentarán de la siguiente forma: primero se describirán los algoritmos predictivos de riesgo y las variables que son usadas en su aplicación, seguidamente se describirá el algoritmo COMPAS en el marco del proceso penal y finalmente, se determinarán cuáles son los impactos de los derechos humanos relacionados con el uso de aquellos algoritmos y en especial del algoritmo COMPAS.

Los resultados apuntan a la hipótesis de que en efecto sí hay riesgos hacia los derechos humanos cuando se utilizan algoritmos predictivos de niveles de riesgo, en el proceso penal.

## **1. ALGORITMOS PREDICTIVOS DE RIESGO EN EL MARCO DEL PROCESO PENAL**

La inteligencia artificial y en especial los algoritmos basados en aprendizaje automático se encuentran presentes en diferentes áreas del

conocimiento. Estos algoritmos analizan grandes cantidades de datos y realizan predicciones según el tipo de necesidad requerida, siendo utilizados principalmente en herramientas del sector financiero, en sistemas de georreferenciación, así como en el ámbito de la comunicación, marketing y redes sociales. Sin embargo, la predicción también está siendo utilizada en el ámbito de la administración de justicia en asuntos tan importantes como la criminalidad, a través de los llamados algoritmos predictivos de riesgo; también “en las últimas décadas, gran parte de la comunidad de investigación de AI & Law se ha centrado en desarrollar modelos computacionales integrales de argumentación legal (CMLA)”<sup>9</sup>.

Volviendo a los algoritmos predictivos de riesgo, se trata de patrones matemáticos basados en la automatización del proceso de predicción del comportamiento de un individuo a través de un puntaje (score), 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,<sup>10</sup> en un proceso judicial.

A su vez, los mencionados algoritmos apoyan las actividades policiales a través de información automatizada en tiempo real, por ejemplo, acerca de las áreas en las que deben desplegarse operativos policiales.

En el marco del proceso penal, están presentes varios algoritmos predictivos de riesgo como lo son el PSA (Public Safety Assessment), el cual es 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.

Otro algoritmo predictivo de riesgo es el LSI-R (Level of Service Inventory-Revised), el cual se usa en los permisos de salida y la libertad condicional de los procesados; aquí, la ponderación utiliza criterios como los antecedentes penales, lugar de residencia, educación, empleo, ocio, familia, problemas de alcohol o drogas, actitudes emocionales y personales.

---

<sup>9</sup> ASHLEY, Kevin. Computational Models of Legal Argument. In Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age, 2017. p. 127. (traducción libre)

<sup>10</sup> 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.

Por otro lado, se encuentra el algoritmo HART (Harm Assessment Risk Tool), 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; este algoritmo ayuda a determinar si el individuo que tiene un riesgo bajo puede ser enviado por los agentes de policía a un programa de rehabilitación denominado Checkpoint.

Además, existe otro algoritmo denominado CAS (Crime Anticipation System), su función es predecir los delitos de alto impacto, por ejemplo, el hurto, que puede llegar a ocurrir en zonas específicas. Las variables que usa CAS están relacionadas con aquellos delitos que se han materializado en dicho lugar con anterioridad, además de tener en cuenta datos “socioeconómicos y demográficos sobre género, etnia [...], edad e ingresos/bienestar, código postal en la ciudad; a la vez usa datos como el historial penal y la distancia de los posibles sospechosos”.<sup>11</sup> Claramente se trata de una herramienta que utiliza las ideas de la criminología ambiental, y del defensible space, según las cuales existen lugares y personas predispuestas para el hecho criminal, que pueden ser determinadas analizando las relaciones entorno-individuo.

También se encontró el algoritmo NDAS (National Data Analytics Solution), el cual es utilizado para predecir delitos violentos, el nivel de riesgo del individuo de cometer un delito con arma blanca o con arma de fuego y quién puede ser víctima del delito. Los criterios que usa el algoritmo son registros de individuos detenidos y de los delitos cometidos. Una vez son identificados los posibles sospechosos se les ofrecen servicios sociales o de salud con el fin de evitar la ocurrencia del delito.

Los diversos algoritmos mencionados, ponderan diferentes variables y son utilizados primordialmente con el objetivo de anticipar la decisión criminal o la reincidencia, e inclusive apoyan la toma de decisiones relacionadas con la libertad, rehabilitación o servicios sociales. Inmerso en estos algoritmos, el que ha tenido un mayor análisis en su aplicación es el COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), el cual “es entendido como una herramienta estructurada que valora el riesgo de reincidencia del procesado y las

---

<sup>11</sup> MUTSAERS, Paul. Changing chronotopes of policing: timespace, race, and juvenile justice in amsterdam, p. 11 (traducción libre).



necesidades criminológicas del sujeto”,<sup>12</sup> dicha herramienta será objeto de análisis en el siguiente capítulo.

A pesar de no ser materia puntual del presente escrito, debe mencionarse también la cada vez más estrecha relación entre neurociencia y derecho penal, la cual ha impulsado el desarrollo de herramientas cada vez más sofisticadas, que permitan como lo advierte de Souza, acercarnos más a la verdad, a la historia que se pretende describir en el proceso penal. En ese escenario, herramientas como el examen neurológico llamado P-300, el cual pretende “reunir elementos de conocimiento cognitivo que el imputado tiene sobre el delito [...] a través del análisis de su actividad eléctrica cerebral frente a ciertos estímulos”, generan dudas tales como la señalada por de Souza, en torno a si “¿son estos métodos aceptables desde la perspectiva de un proceso equitativo y equitativo, basado inequívocamente en la protección de derechos, libertades y garantías?”<sup>13</sup>.

## 2. ALGORITMO COMPAS EN EL MARCO DEL PROCESO PENAL

El algoritmo COMPAS fue desarrollado por la empresa privada Northpointe que 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”.<sup>14</sup>

En la misma línea, “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

<sup>12</sup> MARTÍNEZ, Lucía. Peligrosidad, algoritmos y due process: el caso state vs loomis. *Revista de Derecho Penal y Criminología (uned)*, n. 20, 2018, p. 485.

<sup>13</sup> AIRES DE SOUSA. Susana. Neurociências e Direito Penal: em busca da “verdade” perdida (na mente)?, 2017, p. 4-6.

<sup>14</sup> NORTHPOINTE. Practitioner’s Guide to COMPAS Core, p. 1, (traducción libre).

dinámico (factores criminógenos) como riesgo estático (factores históricos)”.<sup>15</sup>

En efecto, el mencionado algoritmo está siendo utilizado en el proceso penal para valorar el grado de riesgo de reincidencia del sujeto, “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”;<sup>16</sup> 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.

Ahora bien, después de realizada dicha evaluación y ponderación de factores, se obtiene un puntaje (score) dentro de una escala del 1 al 10, en la que 1 es el riesgo más bajo y 10 es el riesgo más alto.

La Guía del Usuario del COMPAS de 2015 realizada por Northpointe, determinó los tipos de escalas de cada nivel de riesgo:

Tipos y escalas principales del algoritmo Compas	
Escala	Tipo de escala
Riesgo de reincidencia violenta	1
Riesgo de reincidencia general	1
Riesgo de liberación previa al juicio	1
Participación criminal	1
Historial de incumplimiento	1
Historia de violencia	1
Violencia actual	1
Asociados criminales / compañeros	4
Abuso de sustancias	2
Problemas financieros / Pobreza	3
Problemas vocacionales / educativos	3

<sup>15</sup> NORTHPOINTE. Practitioner's Guide to COMPAS Core, p. 1, (traducción libre).

<sup>16</sup> 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.

Tipos y escalas principales del algoritmo Compas	
Escala	Tipo de escala
Pensamiento criminal	3
Criminalidad familiar	3
Problemas del entorno social	3
Ocio y recreación	3
Inestabilidad residencial	3
Problemas de adaptación social	3
Fracaso de socialización	3
Oportunidad criminal	3
Personalidad criminal	3
Aislamiento social	3

**GRÁFICO 1.** Elaboración propia, información tomada de NORTHPOINTE, p. 8-9, (traducción libre).

Algunas de las variables utilizadas por el algoritmo para establecer el grado de riesgo de reincidencia están basadas en información privada y posiblemente subjetiva. Por ejemplo, las escalas de ocio o fracaso de socialización, lo cual podría considerarse discriminatorio al tenerse en cuenta criterios ajenos a un riesgo real de reincidencia.

En el mismo orden de ideas en la Guía se mencionó una puntuación para cada tipo de escala, de la siguiente forma:

Puntos para las escalas utilizadas en el algoritmo Compas			
Tipo 1	Bajo (1-4)	Medio (5-7)	Alto (8-10)
Tipo 2	Improbable (1-2)	Probable (3-4)	Muy probable (5-10)
Tipo 3	Improbable (1-5)	Probable (6-7)	Muy probable (8-10)
Tipo 4	Improbable (1-4)	Probable (5-7)	Muy probable (8-10)

**GRÁFICO 2.** Elaboración propia, información tomada de NORTHPOINTE, p. 8-9, (traducción libre).

Cada tipo de escala tendrá una puntuación de 1 a 4 y de igual forma un puntaje de 1 a 10; lo que permitiría determinar 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<sup>17</sup>.

En el mismo orden de ideas, en la guía ya mencionada, 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.

De otro lado, se hará mención de algunas preguntas que integran la encuesta Core las cuales alimentan el algoritmo:

Nombre de la escala	Tipo de pregunta	Tipo de respuesta
No comparecencia	¿Cuántas veces el individuo no se presentó a comparecer ante el tribunal?	___ 0
		___ 1
		___ 2
		___ 3
		___ 4
		___ 5+
Comportamiento cognitivo	¿Cuántos de sus amigos/conocidos consumen drogas con regularidad?	___ Ninguno
		___ Pocos
		___ La mitad
		___ La mayoría
Oportunidad criminal	¿Hay mucha delincuencia en su vecindario?	___ Sí
		___ No
Personalidad criminal	Me meto en problemas porque hago las cosas sin pensar	___ Muy en desacuerdo
		___ En desacuerdo
		___ No estoy seguro
		___ De acuerdo
		___ Muy de acuerdo
Pensamiento criminal	Algunas personas deben ser tratadas con rudeza o golpeadas solo para enviarles un mensaje claro	___ Muy en desacuerdo
		___ En desacuerdo
		___ No estoy seguro
		___ De acuerdo
		___ Muy de acuerdo
Delincuencia familiar	¿Fue arrestado alguna vez su padre (o figura paterna que lo crio principalmente)?	___ Sí
		___ No

<sup>17</sup> NORTHPOINTE. Practitioner's Guide to COMPAS Core, p. 8, (traducción libre).

Nombre de la escala	Tipo de pregunta	Tipo de respuesta
Historial de violencia	¿Cuántos arrestos previos por delitos graves de agresión (no asesinato, sexo o violencia doméstica) ha tenido?	<input type="checkbox"/> 0
		<input type="checkbox"/> 1
		<input type="checkbox"/> 2
		<input type="checkbox"/> 3+
		<input type="checkbox"/> Nunca
Ocio y recreación	¿Con qué frecuencia sintió que no tenía nada que hacer en su tiempo libre?	<input type="checkbox"/> Varias veces al mes
		<input type="checkbox"/> Varias veces a la semana
		<input type="checkbox"/> Diariamente
		<input type="checkbox"/> A
Educación	¿Cuáles fueron sus calificaciones habituales en la escuela secundaria?	<input type="checkbox"/> B
		<input type="checkbox"/> C
		<input type="checkbox"/> D
		<input type="checkbox"/> E / F
		<input type="checkbox"/> No asistió
Historia de enfermedad mental	¿Ha intentado suicidarse alguna vez?	<input type="checkbox"/> Sí
		<input type="checkbox"/> No
Salud mental	¿Alguna vez ha visto cosas u oído voces que no estaban realmente presentes?	<input type="checkbox"/> Sí
		<input type="checkbox"/> No

**GRÁFICO 3.** Elaboración propia, información tomada de NORTHPOINTE, p. 5, s.s, (traducción libre).

Dentro de las escalas y preguntas recién mencionadas, se evidencia que algunas pueden ser polémicas, al estar relacionadas con asuntos de salud como trastornos mentales; información que debería ser utilizada para brindar el acompañamiento psico-social adecuado, pero no deberían impactar negativamente la calificación de riesgo de reincidencia. Adicionalmente, cabe la pena preguntarse, si los eventuales sesgos de los algoritmos predictivos de riesgo, podrían incentivar lo que Silva ha llamado la teoría de la inocuización, de acuerdo con la cual “resulta posible individualizar a un número relativamente pequeño de delincuentes (high risk offenders), 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 inocuización 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”<sup>18</sup>.

<sup>18</sup> SILVA, Jesús. La expansión del derecho penal aspectos de la política criminal en las sociedades postindustriales, 1999, p. 144.

La tendencia a esa inocuización, es preocupante, en tanto podría llevar a un círculo vicioso que aliente la segregación de sectores tradicionalmente discriminados, poniendo en riesgo garantías tradicionales de la persecución penal, tales como la presunción de inocencia, o el propio derecho penal de acto, para regresar a ideas fines al derecho penal de autor.

## **2.1 CASO HITO EN LA APLICACIÓN DEL ALGORITMO COMPAS: STATE WISCONSIN VS LOOMIS**

Algunos estados y condados de Estados Unidos han implementado el uso del algoritmo COMPAS. El caso de Wisconsin State vs Loomis puede catalogarse como hito, ya que al parecer fue la primera vez que un tribunal se pronunció sobre la admisibilidad del uso de herramientas de IA dentro del proceso penal.

En este caso Loomis fue acusado por cinco delitos, ya que estuvo involucrado en un tiroteo, fue imputado por cargos de posesión de arma de fuego, posesión de un rifle de cañón corto, intento de huida de un agente de tráfico, peligro a la seguridad pública y conducción de un vehículo motorizado ajeno sin el consentimiento del propietario. El procesado no admitió su participación en el tiroteo, pero sí aceptó que condujo el vehículo y que intentó huir del oficial de tránsito; finalmente llegó a un acuerdo con la Fiscalía en donde solamente aceptó los cargos de puesta en peligro de la seguridad pública y conducción de un vehículo ajeno sin autorización.

Seguidamente la Fiscalía aportó un informe con el resultado del algoritmo COMPAS, en el que se estableció que el Señor Loomis presentaba un alto grado de riesgo de reincidencia. El juez de primera instancia teniendo en cuenta entre otros, el mencionado informe, impuso una pena de 6 años de prisión además de una pena de 5 años de libertad vigilada.

El apoderado del Señor Loomis apeló la decisión ante la Corte Suprema ya que vulneraba el “derecho al debido proceso por tres razones: (1) viola el derecho del acusado a ser sentenciado con base en información precisa, en parte debido a la naturaleza patentada de COMPAS, lo que impide evaluar su exactitud; (2) viola el derecho del acusado a una

sentencia individualizada; y (3) utiliza indebidamente evaluaciones de género en las sentencias”<sup>19</sup>.

Sobre el primer argumento, la defensa de Loomis, decidió utilizar como ejemplo el precedente del caso *Gardner vs. Florida*, en donde el “Tribunal de Circuito impuso la pena de muerte basándose en parte en la información de un informe de investigación anterior a la sentencia, tal informe no se reveló a los abogados de las partes [...] y se concluyó que se le negó el debido proceso cuando se impuso la pena de muerte, al menos en parte, sobre la base de información que no tuvo oportunidad de negar o explicar [...] no obstante, la corte de apelaciones de Wisconsin determinó que fue un ejercicio erróneo por parte de la Corte de circuito el negar categóricamente la divulgación del informe de investigación a un acusado”<sup>20</sup>.

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”<sup>21</sup>.

No obstante, lo anterior, la Corte estableció que la analogía con el caso *Gardner and Skaff* es imperfecta, ya que Loomis no podría revisar y cuestionar cómo el algoritmo COMPAS calculó el riesgo, al menos podría revisar y cuestionar las puntuaciones de riesgo resultantes establecidas en el informe adjunto; a diferencia del caso *Gardner and Skaff*, en la que el acusado no tuvo oportunidad de refutar información utilizada por el Tribunal al momento de dictar sentencia<sup>22</sup>.

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,

---

<sup>19</sup> SUPREME COURT OF WISCONSIN. Appeal from an order of the circuit court, p. 13. (traducción libre).

<sup>20</sup> SUPREME COURT OF WISCONSIN. Appeal from an order of the circuit court, p. 20-21. (traducción libre).

<sup>21</sup> SUPREME COURT OF WISCONSIN. Appeal from an order of the circuit court, p. 21. (traducción libre).

<sup>22</sup> SUPREME COURT OF WISCONSIN. Appeal from an order of the circuit court, p. 22. (traducción libre).

explica que las puntuaciones se basan en gran medida en información estática (antecedentes penales), con un uso limitado de algunas variables dinámicas”<sup>23</sup>.

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 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 delinquentes de alto riesgo, no un individuo de alto riesgo en particular<sup>24</sup>.

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<sup>25</sup>.

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<sup>26</sup>.

El caso Loomis ha sido paradigmático ya que ha puesto de presente que las garantías del proceso penal y los derechos de los procesados podrían

---

<sup>23</sup> SUPREME COURT OF WISCONSIN. Appeal from an order of the circuit court, p. 22. (traducción libre).

<sup>24</sup> SUPREME COURT OF WISCONSIN. Appeal from an order of the circuit court, p. 30-31. (traducción libre).

<sup>25</sup> SUPREME COURT OF WISCONSIN. Appeal from an order of the circuit court, p. 32. (traducción libre).

<sup>26</sup> SUPREME COURT OF WISCONSIN. Appeal from an order of the circuit court, p. 35-36. (traducción libre).



estar amenazados, por la utilización de los informes de nivel de riesgo derivados del algoritmo COMPAS. Lo anterior, por cuanto los acusados no tienen información suficiente para contradecir tales resultados, al no ser posible establecer el peso, la valoración y ponderación exacta que el algoritmo realiza en relación con las diferentes variables. Lo anterior impactaría el derecho al debido proceso, el principio de contradicción e inclusive la presunción de inocencia.

## 2.2 PROBLEMÁTICAS JURÍDICAS DERIVADAS DE LA APLICACIÓN DEL ALGORITMO COMPAS

Una de las problemáticas surgidas con uso del algoritmo COMPAS es la posibilidad de la presencia de sesgos, en ese sentido Joanna Bryson ha establecido que “la IA tiene el potencial de reforzar los sesgos existentes, porque a diferencia de los humanos, los algoritmos pueden no estar equipados para contrarrestar conscientemente los sesgos aprendidos”<sup>27</sup>. La investigadora ha insistido, en que no se trata de que la IA tenga sesgos per se, sino que ella está aprendiendo y replicando los sesgos presentes en los seres humanos<sup>28</sup>.

Se ha indicado que, si estos sistemas automatizados son alimentados con información “de justicia sesgada, acabarán reproduciendo y potenciando esos sesgos. Si en un país como Estados Unidos las personas de ascendencia africana tienen muchas más probabilidades de acabar en la prisión sin fianza, y entrenamos a una red neuronal con esos datos, el algoritmo acabará replicando ese tipo de sesgos”.<sup>29</sup> Lo anterior quiere decir, que el sesgo de un tipo de datos utilizados por el programador del algoritmo puede verse reflejado en la herramienta de inteligencia artificial lo cual conllevaría a una injusticia algorítmica, al replicarse en esta los mencionados sesgos.

<sup>27</sup> THE GUARDIAN, AI programs exhibit racial and gender biases, research reveals, párr. 6, (traducción libre).

<sup>28</sup> THE GUARDIAN, AI programs exhibit racial and gender biases, research reveals, párr. 6 (traducción libre).

<sup>29</sup> CASACUBERTA, David. Sesgo en bucle: alimentando la injusticia algorítmica, párr. 3.

Quizás la más importante de las críticas ha sido realizada por ProPublica,<sup>30</sup> a través de un informe en el que argumentó que el COMPAS tiene un sesgo racista, lo cual demostró entre otras en el siguiente gráfico:

	Blancos	Afroamericanos
Etiquetado como de mayor riesgo, pero no reincidió	23.5%	44.9%
Etiquetado como de menor riesgo, pero reincidió	28.0%	47.7%

**GRÁFICO 4.** *Elaboración propia, información tomada de Propublica, párr. 62, (traducción libre).*

Este gráfico señala que la calificación de los acusados blancos se inclina hacia el menor riesgo, a diferencia de los acusados afrodescendientes, respecto de los cuales la tendencia se inclina hacia el mayor riesgo. Además se evidenció el denominado falso positivo, entendido como la calificación que da el algoritmo sobre si el procesado va a reincidir cuando en realidad no lo hace, lo cual se vio reflejado con un 44.9% en los procesados afroamericanos a diferencia de los procesados blancos con un 23.5%, así mismo está el denominado falso negativo, que opera de forma contraria, determinando que el procesado no va a reincidir pero efectivamente sí reincide y es así como se reflejó un resultado de un 47.7% en los procesados afroamericanos a diferencia del 28.0% de los procesados blancos, lo que conllevaría a un error discriminatorio en la arquitectura que posee COMPAS.

No obstante, la raza no es el único sesgo que se ha evidenciado en el algoritmo, pues al parecer se presenta un sesgo adicional dirigido hacia grupos minoritarios, ya que “las poblaciones que históricamente han sido perseguidas por las fuerzas del orden público de manera desproporcionada, especialmente las comunidades con ingresos bajos y las minorías; ahora están en riesgo de recibir puntuaciones más altas de reincidencia. Como resultado, el algoritmo podría amplificar y perpetuar

<sup>30</sup> ProPublica es una sala de prensa independiente sin fines de lucro que produce periodismo de investigación con fuerza moral, <https://www.propublica.org/>

los sesgos ya existentes y generar aún más datos afectados por esos sesgos para alimentar ese ciclo vicioso”.<sup>31</sup>

En el mismo orden de ideas el juez Hillman de Estados Unidos, “estableció que varios estados han empezado a experimentar el uso de la IA, como una herramienta para predecir el riesgo de reincidencia de los acusados [...] y para considerar dicha evaluación al momento de dictar la sentencia. Según Hillman, al menos un acusado ha presentado una apelación sobre la puntuación determinada por el algoritmo, sin embargo, dichas apelaciones no han tenido éxito alguno”.<sup>32</sup>

De igual forma para Hillman, “estos experimentos pueden presagiar el crecimiento de dicho uso en la toma de decisiones críticas al dictar sentencia y reflejar un nivel de comodidad cada vez mayor con el uso de la IA entre los jueces”.<sup>33</sup> Lo anterior podría conllevar a que el juez no use esta herramienta solo como apoyo en la toma de decisión (como elemento de convicción), sino que fundamente su decisión exclusivamente en el puntaje establecido en el algoritmo, con lo que se replicaría en la decisión una injusticia algorítmica, y se perdería el elemento de valoración y raciocinio humano. En ese sentido, el juez ha afirmado que “el uso de tecnología predictiva de inteligencia artificial en la sentencia es un desarrollo preocupante, que los tribunales [...] deben enfrentar con escepticismo y un escrutinio minucioso”.<sup>34</sup>

Debe considerarse que “el puntaje de evaluación de riesgo generado por el algoritmo se presenta al tribunal como una presunta determinación fáctica”<sup>35</sup>. Lo que implica que el procesado no cuenta con las garantías procesales suficientes, ya que por un lado dicha tecnología predictiva se incorporaría al proceso como una prueba que no podría

---

<sup>31</sup> HAO, Karen. La IA que evalúa a delincuentes perpetúa la injusticia hacia las minorías, párr. 11.

<sup>32</sup> HILLMAN, Judge. The Use of Artificial Intelligence in Gauging the Risk of Recidivism, párr. 1, (traducción libre).

<sup>33</sup> HILLMAN, Judge. The Use of Artificial Intelligence in Gauging the Risk of Recidivism, párr. 1, (traducción libre).

<sup>34</sup> HILLMAN, Judge. The Use of Artificial Intelligence in Gauging the Risk of Recidivism, párr. 1, (traducción libre).

<sup>35</sup> HILLMAN, Judge. The Use of Artificial Intelligence in Gauging the Risk of Recidivism, párr. 6, (traducción libre).

ser controvertida, “sin una oportunidad concomitante de probar [o desvirtuar] los datos, los supuestos e incluso los prejuicios que subyacen a la conclusión”,<sup>36</sup> debido a que el procesado no podría acceder a la información que tuvo en cuenta el algoritmo para llegar al informe final, ya que este proceso se encuentra protegido por el denominado secreto de empresa; entonces la pregunta que surge es ¿de qué forma se podrían ejercer los derechos de contracción y de defensa si no se puede conocer cómo funciona el algoritmo?

En consecuencia de lo anterior, puede afirmarse que estos sesgos algorítmicos se materializan en los sistemas de aprendizaje automático, específicamente el algoritmo COMPAS, y al ser incorporados en los procesos judiciales como un supuesto fáctico incontrovertible, se estarían violando derechos humanos como la igualdad, no discriminación, debido proceso, libertad personal, honra y dignidad.

Adicionalmente a lo mencionado, tampoco se materializaría el modelo garantista expuesto por Ferrajoli, de acuerdo con el cual los ordenamientos jurídicos penales, al alinearse con tal modelo deben garantizar los principios de legalidad y presunción de inocencia<sup>37</sup>, entre otros.

### 2.3. INCERTIDUMBRE FRENTE A LOS *DERECHOS HUMANOS* CON EL USO DEL ALGORITMO *COMPAS*.

Ya hemos mencionado que la creciente utilización de algoritmos predictivos al interior de la administración de justicia, en especial en materia penal, suscita dudas acerca de las eventuales violaciones de derechos humanos que podrían surgir de tal uso. Se ha evidenciado una gran preocupación por parte de varios sectores de la sociedad civil, que han sugerido cuál debería ser el adecuado uso de los algoritmos predictivos de riesgo, en un marco de protección y respeto de los derechos humanos y garantías judiciales. Si bien es cierto que tales algoritmos aportan celeridad en la toma de decisión judicial, también se han visto comprometidos

---

<sup>36</sup> HILLMAN, Judge. The Use of Artificial Intelligence in Gauging the Risk of Recidivism, párr. 6, (traducción libre).

<sup>37</sup> FERRAJOLI. Derecho y razón, 1995 p. 33.

algunos derechos humanos; lo anterior se hizo evidente por ejemplo en el caso Loomis anteriormente citado.

De acuerdo con la Declaración Universal de Derechos Humanos (en adelante DUDH) “todos los seres humanos nacen libres e iguales en dignidad y derechos”<sup>38</sup>, lo cual se alinea con lo establecido en el Pacto Internacional de Derechos Civiles y Políticos (en adelante PIDCP) que establece en su artículo 14 “[...] todas las personas son iguales ante los tribunales y cortes de justicia [...]”<sup>39</sup>

En igual sentido los Principios y directrices de las Naciones Unidas sobre el acceso a asistencia jurídica en los sistemas de justicia penal, específicamente en el principio No 6, se ha señalado 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”.<sup>40</sup>

Ahora bien, algunos de los tipos de escalas que son aplicadas en el algoritmo COMPAS que fueron mencionadas en el capítulo anterior, podrían desconocer ese derecho a la igualdad y no discriminación, al valorarse criterios tales como la situación financiera, el desempleo o la pobreza. Lo anterior llevaría a la nefasta conclusión de que, si un procesado cuenta con un nivel económico superior a otro, el primero tendría un puntaje de escala de riesgo de reincidencia más bajo, lo que evidenciaría un tratamiento desigual de los procesados al ponderar sus recursos económicos.

En relación con la discriminación, la Organización de Naciones Unidas (ONU) “ha indicado que se trata de una forma de intolerancia por motivos de raza, color, linaje u origen nacional o étnico y además las víctimas pueden sufrir formas múltiples o agravadas de discriminación por otros motivos conexos, como el sexo, el idioma, la religión, las opiniones

<sup>38</sup> NACIONES UNIDAS. La Declaración Universal de Derechos Humanos, p. 5.

<sup>39</sup> NACIONES UNIDAS. Pacto Internacional de Derechos Civiles y Políticos, p. 5.

<sup>40</sup> NACIONES UNIDAS. Principios y directrices de las Naciones Unidas sobre el acceso a la asistencia jurídica en los sistemas de justicia penal, p. 10.

políticas o de otra índole, el origen social, la situación económica y el nacimiento [...]”<sup>41</sup>; además la DUDH en su artículo 7 estableció que todas las personas son iguales ante la ley y tienen, sin distinción, derecho a igual protección de la ley [...]”<sup>42</sup>.

Al respecto, ProPublica hizo evidente el sesgo racial del COMPAS, el cual si bien no utiliza la raza como una variable en la encuesta, si se inclina hacia el sesgo racista, en los denominados falsos positivos, esto es, las personas afrodescendientes son etiquetadas equivocadamente como de alto riesgo, con mayor frecuencia que las personas blancas, quienes son con más frecuencia etiquetadas como falsos negativos, con lo que es claro, que la tendencia se inclina de manera negativa hacia las personas afrodescendientes, evidenciando una discriminación por parte del mencionado algoritmo.

En el mismo sentido, Dieterich, Mendoza y Brennan en estudio realizado en 2016, determinaron que la raza sí afecta los puntajes que son obtenidos en el algoritmo COMPAS, “el porcentaje de acierto a la hora de determinar el riesgo de reincidencia, aunque similar, era superior en el caso de personas blancas (un 67%), frente al 63,8% para las personas negras. Sin embargo, lo relevante de este estudio es que las personas negras que COMPAS puntuó con un alto riesgo fueron incorrectamente catalogadas en un 45% de los casos, dado que no reincidieron, mientras que los blancos lo fueron en un 23,5% de los casos. Por el contrario, los blancos que obtuvieron un porcentaje de bajo riesgo de reincidencia posteriormente reincidieron en un 47,7% de los casos, frente a un 28% de las personas negras. Esta estadística demuestra que COMPAS sí parte de un sesgo racial”.<sup>43</sup>

Por ende, de acuerdo con el mencionado estudio y el informe Propublica entre otros, se pudo determinar con certeza que el derecho a la no discriminación está siendo vulnerado con la aplicación de la evaluación de riesgo de reincidencia. En el mismo sentido Arnold, Dobbie & Yang,

---

<sup>41</sup> NACIONES UNIDAS. Protocolo Facultativo de la Convención contra la Tortura y Otros Tratos o Penas Crueles, p. 11.

<sup>42</sup> NACIONES UNIDAS. La Declaración Universal de Derechos Humanos, p. 16.

<sup>43</sup> DIETERICH, W. et al, como se citó en HERNÁNDEZ, María. Inteligencia artificial y derecho penal. *Revista Actualidad Jurídica Iberoamericana*, n. 10, 2019 p. 825.

intentaron identificar prejuicios raciales en las decisiones de fianza que son tomadas por los jueces, y de acuerdo con los resultados, se encontró que se están cometiendo errores de predicción con sesgo racial, y ello podría llevar a los jueces de fianza a exagerar el peligro relativo de liberar acusados negros versus acusados blancos e igualmente consolidar los estereotipos anti-negros, lo cual exacerba aquellas diferencias entre estos grupos raciales.<sup>44</sup>

En este punto, es interesante un segundo estudio realizado por ProPublica, “donde a través de una solicitud de registros públicos se obtuvieron dos años de puntajes del COMPAS de la Oficina del Sheriff del Condado de Broward en Florida, es así como recibieron datos de las 18,610 personas que fueron calificadas en 2013 y 2014, y construyeron un perfil de los antecedentes penales de cada persona, tanto antes como después de haber sido calificados”.<sup>45</sup> Según dicho estudio para determinar la raza, el algoritmo COMPAS “utilizó las clasificaciones raciales de la Oficina del Sheriff del Condado de Broward, que identifica a los acusados como negros, blancos, hispanos, asiáticos y nativos americanos, adicionalmente en un primer momento ProPublica dedujo que a veces los nombres de las personas o las fechas de nacimiento se ingresaron incorrectamente en algunos registros, lo que condujo a coincidencias incorrectas entre la puntuación COMPAS de una persona y sus antecedentes penales”.<sup>46</sup>

Seguidamente, se analizó la reincidencia entre los acusados blancos y los acusados afrodescendientes detectando que el sesgo racial influyó porcentualmente en una puntuación de 45% más alta del grado de reincidencia.<sup>47</sup> Se concluyó entonces que los acusados afrodescendientes

<sup>44</sup> ARNOL, David, DOBBIE Will & YANG Crystal. Racial Bias in Bail Decisions. *The Quarterly Journal of Economics*, v. 133, n. 4, p. 3, (traducción libre).

<sup>45</sup> LARSON, Jeff, MATTU Surya, KIRCHNER, Lauren & ANGWIN, Julia. How We Analyzed the COMPAS Recidivism Algorithm, párr. 12, (traducción libre).

<sup>46</sup> LARSON, Jeff, MATTU Surya, KIRCHNER, Lauren & ANGWIN, Julia. How We Analyzed the COMPAS Recidivism Algorithm, párr. 12, (traducción libre).

<sup>47</sup> LARSON, Jeff, MATTU Surya, KIRCHNER, Lauren & ANGWIN, Julia. How We Analyzed the COMPAS Recidivism Algorithm, párr. 12, (traducción libre).

tienen el doble de probabilidad de ser calificados en riesgo de reincidir en la escala de tipo violenta, lo cual conlleva un sesgo algorítmico al ponderar la raza como un factor determinante.

Continuando con otros derechos en posible riesgo, el derecho al debido proceso es entendido como “un derecho humano, civil y político, que consiste en el conjunto de garantías procesales que protegen al individuo sometido a un procedimiento judicial”.<sup>48</sup> Al respecto en el ya señalado caso *Wisconsin vs Loomis*, uno de los principales argumentos de la defensa, fue precisamente la violación de este derecho, derivada de la imposibilidad de controvertir la operación realizada por el algoritmo, para establecer el nivel de riesgo.

Es claro que el derecho al debido proceso debe verse reflejado en todas las decisiones judiciales, aun en aquellas que son tomadas con la ayuda de algoritmos predictivos, garantizando que el procesado pueda conocer, analizar y contradecir todos los elementos probatorios en los que se base la decisión judicial, en el mismo orden de ideas la ONU ha establecido que este derecho es “[...]el conjunto de requisitos que deben observarse a fin de que las personas puedan defenderse adecuadamente ante cualquier acto del Estado que pueda afectar sus derechos” [...].<sup>49</sup>

Es así como los individuos que se vean involucrados en decisiones tomadas con ayuda de algoritmos predictivos de riesgo, tienen derecho a contradecir la decisión cuando no se encuentren de acuerdo con la misma. Esto implicaría garantizar acceso abierto a la arquitectura del algoritmo, para conocer el proceso a través del cual se llega al resultado del nivel de riesgo.

En el caso en concreto de *Loomis*, la defensa señaló que este derecho fue vulnerado ya que el procesado en un principio no pudo conocer la forma en que funcionó el algoritmo, lo cual impidió el adecuado ejercicio de la contradicción, y también el derecho a recurrir la sentencia, ya que la opacidad del algoritmo impide refutar la ponderación y valoración previa al resultado arrojado por el COMPAS.

---

<sup>48</sup> ANAYA, Alejandro. *Glosario de términos básicos sobre derechos humanos*, p. 46.

<sup>49</sup> NACIONES UNIDAS, *Protocolo facultativo de la convención contra la tortura y otros tratos o penas crueles, inhumanos o degradantes*, p. 5.



Hasta el momento y de acuerdo con el precedente del caso Loomis, la protección al derecho de empresa termina por impactar negativamente otros derechos y garantías del procesado, por lo cual, si la toma de decisiones judiciales fuese soportada entre otros, por algoritmos privados que el acusado no puede examinar o controvertir [...] el derecho al debido proceso o de defensa estaría siendo violado, conllevando a la falta de confianza pública en la integridad del sistema judicial,<sup>50</sup> por ello sería interesante determinar cómo debe hacerse la ponderación de estos derechos involucrados.

De otro lado, se evidenció el derecho a la explicación y el derecho al acceso de la información con transparencia, los cuales se encuentran relacionados con el de contradicción, en este sentido la Comisión Europea presentó una propuesta a través del denominado “libro blanco”, en el cual se planteó que la IA al ser aplicada debe tener varios criterios en cuenta, tales como: “acción y supervisión humana, solidez técnica y seguridad, gestión de la privacidad y de los datos, transparencia, diversidad, no discriminación y equidad, bienestar social y medioambiental y rendición de cuentas”.<sup>51</sup>

Ahora bien, la falta de conocimiento sobre las operaciones que ejecuta internamente una herramienta basada en aprendizaje automático, implica que “la falta de transparencia (opacidad de la IA) hace difícil detectar y demostrar los posibles incumplimientos de la legislación, especialmente las que protegen los derechos fundamentales que imputan responsabilidades y permiten reclamar una indemnización [...]”<sup>52</sup>.

El factor de transparencia también se relaciona con la creación de los algoritmos, y lo que se ha calificado como el diseño ético de los mismos; es así como Rizer & Watney determinaron que “los algoritmos no son completamente neutrales ni herramientas objetivas, ya que pueden estar sesgadas debido a la conservación inadecuada de datos y

<sup>50</sup> RIZER, Arthur & WATNEY, Caleb. Artificial intelligence can make our jail system more efficient, equitable and just. *Texas Review of Law & Politics*, Austin, v. 23, n. 1, 2018, p. 28, (traducción libre).

<sup>51</sup> COMISIÓN EUROPEA. Libro Blanco. Sobre la Inteligencia Artificial - Un Enfoque Europeo Orientado a la Excelencia y la Confianza, p. 11.

<sup>52</sup> RIZER, Arthur & WATNEY, Caleb. Artificial intelligence can make our jail system more efficient, equitable and just. *Texas Review of Law & Politics*, Austin, v. 23, n. 1, 2018, p. 25, (traducción libre).

la selección de las variables que los algoritmos buscarán optimizar. Si los desarrolladores humanos no tienen cuidado, o son intencionalmente maliciosos en la creación de estos algoritmos, estos sesgos pueden tener efectos sociales nocivos”.<sup>53</sup>

En este orden de ideas, los algoritmos predictivos que se comporten como una caja negra y los sesgos que los pueden integrar, pondrían en riesgo derechos y garantías procesales, el cual se vería agravado cuando los tribunales realizan una ponderación a favor del derecho a la propiedad intelectual y el secreto de empresa.

Otro derecho que se ha señalado como vulnerado en el caso Loomis ha sido la garantía a un juez o tribunal imparcial, el cual se encuentra regulado en el artículo 10 en la DUDH donde se establece que “toda persona tiene derecho a ser oída públicamente y con justicia por un tribunal independiente e imparcial [...]”.<sup>54</sup>

De la misma forma en el artículo 14 del PIDCP se señala que “[...] toda persona tendrá derecho a ser oída públicamente y con las debidas garantías por un tribunal competente, independiente e imparcial, establecido por la ley, en la substanciación de cualquier acusación de carácter penal [...]”,<sup>55</sup> ahora bien, en el caso en concreto el juez fundamentó en gran parte su decisión por el puntaje establecido por el COMPAS; si bien el algoritmo es una herramienta de apoyo para el juez en la toma de la decisión, ya se ha concluido que el resultado arrojado por el algoritmo puede estar viciado con diferentes sesgos, y en ese orden de ideas el juez o tribunal perderían la imparcialidad, como quiera que esta herramienta de convicción puede entregar resultados sesgados por la raza, el nivel económico, la salud mental, entre otros.

Es evidente que la decisión tomada sobre el algoritmo COMPAS impacta el derecho a la libertad personal, respecto del cual el PIDCP señala en su artículo 9, que todo individuo tiene derecho a la libertad y a la seguridad personal, por ende, nadie podrá ser sometido a detención

---

<sup>53</sup> RIZER, Arthur & WATNEY, Caleb. Artificial intelligence can make our jail system more efficient, equitable and just. *Texas Review of Law & Politics*, Austin, v. 23, n. 1, 2018, p. 25, (traducción libre).

<sup>54</sup> NACIONES UNIDAS. La Declaración Universal de Derechos Humanos, p. 22.

<sup>55</sup> NACIONES UNIDAS. Pacto Internacional de Derechos Civiles y Políticos, p. 5.

o prisión arbitraria <sup>56</sup>. En el caso Loomis la herramienta determinó un alto grado de nivel de reincidencia por lo cual el juez decidió privar de la libertad al procesado, teniendo en cuenta el informe arrojado por COMPAS como uno de los criterios para determinar la pena privativa de libertad en la sentencia; por lo tanto es evidente que el derecho a la libertad personal del procesado podría no verse garantizado cuando el algoritmo usa algunos criterios subjetivos tales como el ocio o fracasos de socialización, por ejemplo.

Otro derecho que es conveniente mencionar, es la dignidad humana, reconocido en el artículo 7 de la DUDH: “todos son iguales ante la ley y tienen, sin distinción, derecho a igual protección de la ley [...]” <sup>57</sup>, en la misma línea el PIDCP establece en su preámbulo que todos los derechos allí mencionados emanan de la dignidad inherente al ser humano<sup>58</sup>.

En el caso Loomis este derecho se vio quebrantado al tenerse en cuenta algunas variables subjetivas que escapan al control del procesado, tales como vivir en un barrio peligroso, o no haber conseguido trabajo por un largo tiempo, además de la raza, la cual, aunque no es tenida en cuenta en las preguntas, presuntamente sí influye en el nivel de riesgo de reincidencia del procesado, lo que conlleva a una discriminación algorítmica que materializa una afectación a la dignidad humana e igualdad ante la ley.

Por otra parte, la DUDH en el artículo 12, precisa que nadie será objeto de ataques contra su honra o reputación, y que la ley protegerá a las personas que puedan llegar a verse afectadas con tales actos<sup>59</sup>. Si bien el derecho a la honra no fue mencionado por la defensa de Loomis, este puede verse afectado con la utilización del COMPAS, ya que la calificación de riesgo establece una etiqueta algorítmica. Esa categorización del individuo trae consigo las denominadas ceremonias de degradación a las que se refería Garfinkel, señalando que aquella consiste en una degradación del status, que implica una transformación de la identidad que pasa a

---

<sup>56</sup> NACIONES UNIDAS. Pacto Internacional de Derechos Civiles y Políticos, p. 4.

<sup>57</sup> NACIONES UNIDAS. La Declaración Universal de Derechos Humanos, p. 3.

<sup>58</sup> NACIONES UNIDAS. Pacto Internacional de Derechos Civiles y Políticos, p. 1.

<sup>59</sup> NACIONES UNIDAS. La Declaración Universal de Derechos Humanos, p. 26.

ser percibida como inferior.<sup>60</sup> Así pues, el resultado de la evaluación de riesgo crea un perfil del procesado que lo etiqueta como peligroso sin que se haya materializado el posible delito, con lo que se estaría violando el derecho a la honra y reputación.

Se puede concluir que los derechos humanos están enfrentándose a vulneraciones inminentes con los algoritmos predictivos de riesgo, por razones que ya fueron descritas; sin embargo, estas herramientas son innovadoras y aportan grandes utilidades al sistema penal, por ello, es importante hacer una reflexión sobre la forma adecuada en la que deben construirse y utilizarse, con el fin de garantizar el respeto a las garantías procesales y los derechos que puedan llegar a verse afectados.

### **3. REGULACIÓN Y ESTUDIOS EXISTENTES ACERCA DE LOS ALGORITMOS PREDICTIVOS DE RIESGO**

Una de las consecuencias derivadas de las controversias jurídicas anteriormente planteadas, ha sido la iniciativa de una regulación del uso de los algoritmos predictivos de riesgo al interior de la administración de justicia.

La principal regulación actual se encuentra en el Reglamento General de Protección de Datos para ciudadanos de la Unión Europea (en adelante UE), en relación con el adecuado tratamiento de protección de datos personales y la libre circulación de estos, tanto de las organizaciones europeas como de aquellas externas que recogen datos de los residentes de la UE <sup>61</sup>.

El mencionado reglamento estableció que, para la elaboración de perfiles basados en datos se deben evitar efectos discriminatorios por razón de raza, opiniones políticas, religión o creencias, afiliación sindical, condición genética, estado de salud u orientación sexual, por lo cual el interesado podrá ejercer el derecho de oposición respecto del proceso de creación de perfiles cuando considere que se estén violando derechos o libertades fundamentales.

---

<sup>60</sup> GARFINKEL, Harold, como se citó en GARCÍA, German. *Criminología. Construcciones sociales e innovaciones teóricas*, 2011, p. 327.

<sup>61</sup> COMISIÓN EUROPEA. *El Reglamento general de protección de datos*, p. 1.

Por otro lado, en la legislatura del Estado de Idaho de Estados Unidos, existe una normativa sobre transparencia que exige la explicación acerca del cómo las herramientas llegaron a un resultado específico.<sup>62</sup>

Hasta el momento, no existe una regulación legal amplia sobre los algoritmos predictivos de riesgo, pero sí se ha evidenciado un interés de parte de la ciudadanía para que sean regulados; en ese sentido, un ciudadano presentó un proyecto de ley en el Estado de New York, en el que se solicita la creación de un grupo de trabajo, que proporcione recomendaciones sobre cómo la información de los sistemas de decisión automatizada, se podría compartir con el público y cómo se analizarían aquellos casos, en donde las personas puedan llegar a verse afectadas por dichas decisiones tomadas por los algoritmos.<sup>63</sup>

Se presentó otro proyecto de ley en el Estado de Massachusetts, en el cual se requiere la creación de una comisión para la toma de decisiones automatizadas de la agencia estatal con el fin de que no se sigan produciendo vulneraciones de los derechos humanos.<sup>64</sup>

De igual forma Amnesty International y Access Now se han pronunciado sobre la IA, a través de la Declaración de Toronto donde se estableció que se deben proteger los derechos a la igualdad y a la no discriminación en los sistemas de aprendizaje automático.<sup>65</sup>

En el mismo orden de ideas, existen organizaciones cuyo propósito es luchar en contra de las injusticias algorítmicas, y generar una cultura de diseño ético de algoritmos. Por ejemplo, el instituto AI now, el cual se dedica a comprender las implicaciones sociales de la IA, y en ese ejercicio, ha señalado la necesidad de analizar sus impactos respecto de derechos y libertades, a la vez que los sesgos y problemas de inclusión derivados de la falta de equidad respecto de poblaciones diversas.<sup>66</sup> En el mismo sentido, la activista Joy Buolamwini ha creado

<sup>62</sup> Legislatura de Idaho-Criminal Procedure, (traducción libre).

<sup>63</sup> A Local Law In Relation to Automated Decision Systems Used by Agencies, (traducción libre).

<sup>64</sup> An Act Establishing a Commission on Automated Decision-Making, Artificial Intelligence, Transparency, Fairness, and Individual Rights, (traducción libre).

<sup>65</sup> Toronto's Declaration on Machine Learning, (traducción libre).

<sup>66</sup> AI NOW.

la Liga de la Justicia algorítmica, y ha señalado que en el mundo actual la IA gobierna el acceso a la información, las oportunidades y la libertad, puntualizando que los sistemas inteligentes pueden perpetuar el racismo, la discapacidad, y en general muchas formas de discriminación.<sup>67</sup> Por ello propone la denominada "codificación inclusiva"<sup>68</sup>, para moderar los efectos de la subrepresentación de mujeres y personas de color en el sector de la tecnología<sup>69</sup>.

## CONCLUSIONES

Los algoritmos predictivos han irrumpido en diferentes áreas del conocimiento y el escenario jurídico no ha sido ajeno a tal realidad, lo cual se evidencia con las herramientas PSA, LSI-R, HART, CAS, NDAS, que están siendo aplicados en actividades policiales y también en el proceso penal, aportando celeridad y precisión en la toma de decisiones.

El COMPAS como otros sistemas inteligentes predictivos, ha sido objeto de críticas, ya que, en el proceso de determinación de riesgo de reincidencia, se ponderan algunos criterios subjetivos tales como la pobreza, la enfermedad mental, el ocio, el rendimiento escolar, la actividad criminal de amigos y familiares. Lo anterior evidencia una discriminación algorítmica derivada del etiquetamiento que implica un riesgo para los derechos humanos como la no discriminación, libertad personal, debido proceso, acceso a la información transparente, entre otros.

Así mismo, se logró determinar que existen impactos negativos para los derechos humanos inmersos en un proceso judicial por el uso de algoritmos predictivos de riesgo. La opacidad de este deriva de un funcionamiento cuyos datos de entrada y salida se conocen, pero no el proceso de ponderación de las variables, lo que implica que el algoritmo

---

<sup>67</sup> BUOLAMWINI, Joy. IBM Leads, More Should Follow: Racial Justice Requires Algorithmic Justice, párr. 1, (traducción libre).

<sup>68</sup> Para más información al respecto se puede consultar [https://www.ted.com/talks/joy\\_buolamwini\\_how\\_i\\_m\\_fighting\\_bias\\_in\\_algorithms?language=es#t-381312](https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms?language=es#t-381312)

<sup>69</sup> BUOLAMWINI, Joy. Artificial Intelligence Has a Problem with Gender and Racial Bias. Here's How To Solve It, párr. 1, (traducción libre).

se comporta como una caja negra que unido al secreto de empresa impide que el afectado acceda y controvierta estas operaciones; lo anterior impacta el adecuado ejercicio de los derechos a la explicación, a la información, a la contradicción, a un juez imparcial y derechos correlacionados.

Ahora bien, en el diseño, programación y entrenamiento de los algoritmos, los sesgos en los datos utilizados para alimentar la herramienta pueden reflejar prejuicios del diseñador del algoritmo y/o replicar los estereotipos presentes en la sociedad. Lo anterior genera una discriminación o afectación de derechos de minorías, derivados de la raza, género o estrato económico.

La escasa regulación en torno al diseño y utilización de la IA en ámbitos de la administración de justicia ha conllevado vulneraciones de los derechos humanos; principalmente porque los funcionarios judiciales no cuentan con las directrices, protocolos o normas vinculantes que señalen la adecuada aplicación de la IA en la toma de decisión. Como consecuencia de lo anterior, desde la sociedad civil, operadores de justicia, y en general los diversos actores del sistema judicial, se demanda la necesidad de una regulación precisa y completa. Actualmente existen muy pocas legislaciones que aborden el asunto, algunos documentos y exhortaciones no vinculantes como el Libro Blanco, entre otros, los cuales propenden por la defensa de los derechos que puedan llegar a verse afectados con ocasión de la aplicación de la IA.

Por último, es importante mencionar que el uso adecuado de los algoritmos predictivos de riesgo requiere un diseño ético de los mismos, de tal forma que se eviten los sesgos en los datos que lo alimentan tanto en su diseño, como en las fases de entrenamiento, validación y pruebas. Además, debe garantizarse la transparencia algorítmica, entendida como eliminación o moderación de la opacidad del algoritmo, a efectos de que sea posible conocer la forma exacta cómo funcionan las operaciones al interior de este. Las anteriores medidas permitirían aprovechar las utilidades de la IA en la administración de justicia, materializando plenamente las garantías y derechos procesales.

## BIBLIOGRAFÍA

ANAYA, Alejandro. *Glosario de términos básicos sobre derechos humanos*. Universidad Iberoamericana, Ciudad de México: Constanza Perez Reyes, 2006.

ANGWIN, Julia; LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren. Machine Bias. Disponible en: <<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing/>>. Acceso: 01 oct. 2020.

ARNOL VENTURES. *Public Safety Assessment FAQs* ("PSA 101"). 2019.

ARNOLD, David; DOBBIE, Will. YANG, Crystal S. Racial bias in bail decisions. *The Quarterly Journal of Economics*, v. 133, n. 4, p. 1885-1932, 2018. <http://dx.doi.org/10.3386/w23421>

ASHLEY, Kevin. Computational Models of Legal Argument. *Cambridge University Press*, p.127-168, (2017). <https://doi.org/10.1017/9781316761380.005>

AYUDA, Fernando. ¿Inteligencia Artificial y Derecho? Sí, pero ¿cómo?. *Revista Democracia Digital e Governo Eletrônico*, Florianópolis, v. 1, n. 18, p. 36-57, 2019, 36-57. 2019.

BUOLAMWINI, Joy. Artificial intelligence has a problem with gender and racial bias. Here's how to solve it. Disponible en: < <https://www.media.mit.edu/articles/artificial-intelligence-has-a-problem-with-gender-and-racial-bias-here-s-how-to-solve-it/> >. Acceso: 30 oct. 2020.

BUOLAMWINI, Joy. IBM Leads, More Should Follow: Racial Justice Requires Algorithmic Justice. Disponible en: < <https://medium.com/@Joy.Buolamwini/ibm-leads-more-should-follow-racial-justice-requires-algorithmic-justice-and-funding-da47e07e5b58/> >. Acceso: 30 oct. 2020.

CASABONA, Carlos María. Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad. *Revista de Derecho, Empresa y Sociedad*, n. 13, p. 39-55, 2018.

CASACUBERTA, David. Sesgo en bucle: alimentando la injusticia algorítmica. Disponible en: <<http://lab.cccb.org/es/sesgo-en-bucle-alimentando-la-injusticia-algoritmica/>>. Acceso: 01 oct. 2020.

COMISIÓN EUROPEA. *Libro Blanco. Sobre la Inteligencia Artificial: un enfoque europeo orientado a la excelencia y la confianza*. Bruselas: 2020

CORTE INTERAMERICANA DE DERECHOS HUMANOS. (s.f.). *Cuadernillo de jurisprudencia de la Corte Interamericana de Derechos Humanos n° 12: debido proceso*. s.f.



CORVALÁN, Juan Gutavo. (2018). Inteligencia artificial: retos, desafíos y oportunidades- pometea: la primera inteligencia artificial de Latinoamérica al servicio de la Justicia. *Revista de Investigações Constitucionais*, v. 5, n. 1, p. 295-316, 2018. <http://dx.doi.org/10.5380/rinc.v5i1.55334>

FERRAJOLI, Luigi. *Derecho y razón*. Madrid, Trotta, 1995.

GARAY, Lucía. Peligrosidad, algoritmos y due process: el caso State v Loomis. *Revista de Derecho Penal y Criminología (uned)*, n. 20, p. 485-502, 2018. <http://dx.doi.org/10.5944/rdpc.20.2018.26484>

GARFINKEL, Harold. Conditions of successful degradation ceremonies. *American journal of Sociology*, v. 61, n. 5, p. 420-424, 1956. <http://dx.doi.org/10.1086/221800>

GIMÉNEZ, María. Inteligencia Artificial y Derecho Penal. *Actualidad Jurídica Iberoamericana*, n. 10, p. 792-843, 2019.

HAO, Karen. This is how AI bias really happens—and why it's so hard to fix. Disponible en: <<https://www.technologyreview.com/2019/02/04/137602/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/>> . Acceso: 01 oct. 2020.

HAO, Karen. La IA que evalúa a delincuentes perpetúa la injusticia hacia las minorías. Disponible en: <<https://www.technologyreview.es/s/10899/la-ia-que-evalua-delincuentes-perpetua-la-injusticia-hacia-las-minorias/>> . Acceso: 01 oct. 2020.

HILLMAN, Noel. The Use of Artificial Intelligence in Gauging the Risk of Recidivism. Disponible en: <[https://www.americanbar.org/groups/judicial/publications/judges\\_journal/2019/winter/the-use-artificial-intelligence-gauging-risk-recidivism/](https://www.americanbar.org/groups/judicial/publications/judges_journal/2019/winter/the-use-artificial-intelligence-gauging-risk-recidivism/)> . Acceso: 01 oct. 2020.

KEHL, Danielle; GUO, Priscilla; KESSLER, Samuel. Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing. *Berkman Klein Center for Internet & Society, Harvard Law School*, p. 2-36, 2017.

LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren; ANGWIN, Julia. How We Analyzed the COMPAS Recidivism Algorithm. Disponible en: <<https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm/>>. Aceso en: 21 sep. 2020.

LINARES, Fernando. Inteligencia Artificial y Justicia Penal: Más Allá de los Resultados Causados por Robots. *Revista de Derecho Penal y Criminología*, n. 20, p. 87-130, 2018. <http://dx.doi.org/10.5944/rdpc.20.2018.26446>

LIU, Han-Wei; LIN, Ching-Fu; CHEN, Yu-Jie. Beyond State v Loomis: artificial intelligence, government algorithmization and accountability. *International Journal of Law and Information Technology*, v. 27, n. 2, p. 122-141, 2019. <http://dx.doi.org/10.1093/ijlit/eaz001>

MUTSAERS, Paul. Changing Chronotopes of Policing: TimeSpace, Race, and Juvenile Justice in Amsterdam. *Environment and Planning D: Society & Space*, s.f.

NACIONES UNIDAS. *Informe de la Conferencia Mundial contra el Racismo, la Discriminación Racial, la Xenofobia y las Formas Conexas de Intolerancia*. Durban: 2001.

NACIONES UNIDAS. Protocolo Facultativo de la Convención contra la Tortura y Otros Tratos o Penas Crueles: *declaración provisional sobre el papel desempeñado por la revisión judicial y el debido proceso en la prevención de la tortura adoptado por el Subcomité para la Prevención de la Tortura y Otros Tratos o Penas Crueles, Inhumanos o Degradantes*. 2012.

NACIONES UNIDAS. *Principios y directrices de las Naciones Unidas sobre el acceso a la asistencia jurídica en los sistemas de justicia penal*. Viena: Oficina de las Naciones Unidas contra la droga y el delito, 2013.

NACIONES UNIDAS. Igualdad y no discriminación. Disponible en: <<https://www.un.org/ruleoflaw/es/thematic-areas/human-rights/equality-and-non-discrimination/>>. Acceso en: 21 sep. 2020.

NAVARRO, José Bonet. Algunas Reflexiones Sobre la Viabilidad de la Inteligencia Artificial en el Proceso Penal. *REVISTA INMEXIUS. Revista de Derecho Penal y Procesal Penal*. n. 40, p. 1-12, 2020.

NORTHPOINTE. *Measurement & treatment implications of compas core scales*. Northpointe Institute for Public Management, Inc, 2009.

NORTHPOINTE. *Practitioner's Guide to COMPAS Core*. Northpointe Inc, 2015.

OSOBA, Osonde, WELSER IV, William. *An intelligence in our image: The risks of bias and errors in artificial intelligence*. Rand Corporation, 2017.

RIZER, Arthur; WATNEY, Caleb. Artificial intelligence can make our jail system more efficient, equitable, and just. *Texas Review of Law & Politics*, Austin, v. 23, n. 1, p. 181-227, 2018. <http://dx.doi.org/10.2139/ssrn.3129576>

SILVA, Jesús. *La expansión del Derecho penal. Aspectos de la política criminal en las sociedades postindustriales*, Madrid, p. 144, 1999.

SOUSA, Susana. Neurociências e Direito Penal:em busca da “verdade” perdida (na mente)? *Nótulas á margem do “caso Ricla”*, p. 3-12, 2017.

WAKEFIELD, Jane. Así funciona “la mente” de Norman, el algoritmo psicópata del MIT que solo ve lo más tenebroso de la red. Disponible en: <<https://www.bbc.com/mundo/noticias-44482471>>. Acceso en: 19 sep. 2020.

### Authorship information

*Marcela del Pilar Roa Avella.* Profesora e investigadora del grupo “Derecho Público” y de la línea “Derecho Penal y Justicia Militar” del Centro de Investigaciones Jurídicas, Políticas y Sociales de la Facultad de Derecho, Bogotá de la Universidad Militar Nueva Granada. Abogado por la Universidad del Rosario (Colombia). Especialista en Ciencias Penales y Ciencias Penales de la Universidad Externado (Colombia). Magister en derecho penal de la Universidad de Santo Tomás en convenio con la Universidad de Salamanca (Colombia). [marcela.roa@unimilitar.edu.co](mailto:marcela.roa@unimilitar.edu.co)

*Jesús Eduardo Sanabria-Moyano.* Profesor e investigador del grupo “Derecho Público” y de la línea “Derecho Internacional, Derechos Humanos y Derechos Humanos Internacionales” del Centro de Investigaciones Jurídicas, Políticas y Sociales de la Facultad de Derecho, Bogotá de la Universidad Militar Nueva Granada. Abogado por la Universidad Militar Nueva Granada (Colombia). Especialista en Derechos Humanos y Derecho Internacional Humanitario aplicado a los Conflictos Armados de la Facultad de Derechos Humanos y Derecho Internacional Humanitario del Ejército Nacional (Colombia). Magister en Derecho Público Militar de la Universidad Militar Nueva Granada (Colombia). [jesus.sanabria@unimilitar.edu.co](mailto:jesus.sanabria@unimilitar.edu.co)

*Katherin Dinas-Hurtado.* Joven investigadora de la Universidad Militar Nueva Granada del Grupo Derecho público y línea de derecho penal y justicia militar. Abogada por la Universidad Militar Nueva Granada (Colombia). [u0304087@unimilitar.edu.co](mailto:u0304087@unimilitar.edu.co)

### **Additional information and author's declarations (scientific integrity)**

*Acknowledgement:* Reconocemos el trabajo del Ingeniero en Multimedia (Juan Sebastian Anaya Jaimes) quien brindo sus servicios de asesoría en los conceptos técnicos mencionados en el trabajo de investigación.

*Conflict of interest declaration:* the authors confirm that there are no conflicts of interest in conducting this research and writing this article.

*Declaration of authorship:* all and only researchers who comply with the authorship requirements of this article are listed as authors; all coauthors are fully responsible for this work in its entirety.

- *Marcela del Pilar Roa Avella:* conceptualization, methodology, data curation, investigation, writing – original draft, validation, writing – review and editing, final version approval.
- *Jesús Eduardo Sanabria-Moyano:* conceptualization, methodology, investigation, writing – original draft, validation, writing – review and editing, final version approval.
- *Katherin Dinas-Hurtado:* conceptualization, data curation, investigation, writing – original draft, writing – review and editing, final version approval.

*Declaration of originality:* the authors assure that the text here published has not been previously published in any other resource and that future republication will only take place with the express indication of the reference of this original publication; they also attest that there is no third party plagiarism or self-plagiarism.

### Editorial process dates

(<http://www.ibraspp.com.br/revista/index.php/RBDPP/about/editorialPolicies>)

- Submission: 13/07/2021
- Desk review and plagiarism check: 20/07/2021
- Review 1: 01/08/2021
- Review 2: 02/08/2021
- Review 3: 12/08/2021
- Preliminary editorial decision: 09/01/2022
- Correction round return 1: 29/01/2022
- Preliminary editorial decision 2: 13/02/2022
- Correction round return 2: 14/02/2022
- Final editorial decision: 20/03/2022

### Editorial team

- Editor-in-chief: 1 (VGV)
- Assistant-editor: 1 (DDE)
- Reviewers: 3

### HOW TO CITE (ABNT BRAZIL):

ROA AVELLA, Marcela del Pilar; SANABRIA-MOYANO, Jesús E.; DINAS-HURTADO, Katherin. Uso del algoritmo compas en el proceso penal y los riesgos a los derechos humanos. *Revista Brasileira de Direito Processual Penal*, vol. 8, n. 1, p. 275-310, jan./abr. 2022. <https://doi.org/10.22197/rbdpp.v8i1.615>



*License Creative Commons Attribution 4.0 International.*