

**Evaluación de Riesgo Criminal Mediante Inteligencia Artificial: Análisis de
la Posible Implementación un Sistema como COMPAS en Colombia y sus**

Implicaciones Legales y Éticas

*Artificial Intelligence in Criminal Risk Assessment: An Analysis of the
Potential Adoption of COMPAS like Systems in Colombia and Their Legal and
Ethical Implications*



Presentado por:

Evelyn Camila Herrera Guana

Colegio Mayor de Nuestra Señora del Rosario - Universidad del Rosario

Facultad de Jurisprudencia

Tutor: Samuel Augusto Escobar

Bogotá D.C., Colombia

27 de abril de 2025

Índice

1. Resumen / Abstract	3
2. Introducción	6
3. Capítulo 1: La inteligencia artificial y su incursión en el sistema penal	8
3.1. Concepto y evolución de la inteligencia artificial.....	8
3.2. Fundamentos de la IA aplicada al derecho penal.....	11
3.3. El sistema COMPAS: origen, lógica y funcionamiento.....	12
3.4. Riesgos constitucionales y éticos de los algoritmos en el proceso penal.....	14
4. Capítulo 2: El caso State v. Loomis y el debate constitucional en EE. UU	16
4.1. Contexto y Antecedentes.....	16
4.2. Decisión de la Corte Suprema de Wisconsin.....	17
4.3. Reacciones doctrinales y fallas estructurales de COMPAS.....	20
4.4. Estándares Internacionales y riesgos para sistemas jurídicos latinoamericanos.....	24
5. Capítulo 3: Análisis jurídico de la posible implementación de COMPAS en Colombia	25
5.1 El principio del juez natural en la Constitución Colombiana.....	25
5.2 Análisis de la Sentencia T-323 de 2024.....	28
5.3 Compatibilidad normativa y jurisprudencial con la IA judicial.....	31
6. Capítulo 4: Desafíos éticos, culturales y tecnológicos en el contexto colombiano	33
6.1 Sesgos algorítmicos y diversidad sociocultural.....	33
6.2 Riesgos para la autonomía judicial y la justicia diferencial.....	36
7. Conclusiones	39
8. Bibliografía	44

1. Resumen

La presente monografía analiza si es legal y éticamente viable que el poder judicial aplique un sistema de IA como COMPAS (Gestión de Delincuentes Correccionales para Perfiles de Sanciones Alternativas) en Colombia, actualmente empleado en EE. UU. para estimar el riesgo de reincidencia criminal. A través de una metodología dogmática y cualitativa, analiza los riesgos de automatizar decisiones judiciales penales, específicamente aquellas relacionadas con el estado de derecho, el juez natural y la transparencia en la administración de justicia.

Primero, se discute cómo funciona COMPAS y cómo está sesgado, es inexplicable, y tiene un sesgo discriminatorio, como se ha demostrado en varios estudios. Luego aborda el caso del Estado contra Loomis, decidido por la Corte Suprema de Wisconsin, donde se permite el uso de la herramienta a los tribunales para apoyar sentencias impugnadas en apelación, a pesar de las dudas sobre si se llega al juicio con la evidencia disponible y el principio de contradicción.

Luego proporciona una visión general de la ley constitucional colombiana, a través del lente de la Sentencia T-323 de 2024 de la Corte Constitucional, que establece estándares estrictos para permitir la confianza legítima en la inteligencia artificial en el ámbito judicial, tales como transparencia, explicabilidad, control humano significativo y salvaguarda para bases de datos personales. Por último, se refiere a los desafíos éticos y culturales-tecnológicos que la aplicación de un modelo algorítmico de evaluación de riesgo criminal en Colombia enfrentarían, basados en la diversidad sociocultural y marcadamente desigual del país.

Concluye que, como resultado de su construcción, este modelo actual de COMPAS es incompatible con el modelo constitucional colombiano de justicia, dado que

su uso pondría en riesgo principios fundamentales del estado social de derecho. Solo si se asegura el respeto por los derechos fundamentales, la independencia del poder judicial y un control democrático efectivo sobre las herramientas tecnológicas aplicadas, el uso de tecnologías predictivas en la impartición de justicia puede considerarse legítimo.

Abstract

This monograph examines the legal and ethical viability of implementing an artificial intelligence system such as COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) within the Colombian judiciary. COMPAS, currently utilized in the United States to assess the risk of criminal recidivism, serves as the focal point for an analysis conducted through a dogmatic and qualitative methodology, addressing the risks inherent in the automation of criminal judicial decision-making — particularly with respect to the rule of law, the principle of the natural judge, and transparency in the administration of justice.

The study first outlines the operational mechanics of the COMPAS system, highlighting its opacity, demonstrable biases, and discriminatory effects, as evidenced by multiple empirical studies. It subsequently analyzes the case of *State v. Loomis*, adjudicated by the Wisconsin Supreme Court, wherein the judiciary authorized the use of COMPAS to substantiate sentencing decisions subject to appellate review, notwithstanding serious concerns regarding evidentiary access and the adversarial principle.

The monograph further presents an overview of Colombian constitutional law, particularly through the framework established by Judgment T-323 of 2024 of the Constitutional Court, which sets forth stringent requirements for the legitimate reliance upon artificial intelligence in judicial proceedings, including mandates of transparency,

explainability, meaningful human oversight, and the safeguarding of personal data repositories.

Finally, the study addresses the ethical and cultural-technological challenges that would arise from the application of an algorithmic criminal risk assessment model within Colombia, emphasizing the profound sociocultural diversity and systemic inequality characterizing the national context.

The analysis concludes that, given its current structure, the COMPAS model is fundamentally incompatible with the Colombian constitutional framework of justice, as its use would compromise essential principles underpinning the social rule of law. The deployment of predictive technologies in the judiciary can only be deemed legitimate insofar as fundamental rights are respected, judicial independence is preserved, and effective democratic oversight of technological tools is maintained.

2. Introducción

La Inteligencia Artificial (IA) ha experimentado un crecimiento considerable en las últimas décadas y ha impactado todos los ámbitos de la vida social. Por supuesto, el Derecho no es una excepción. En particular, cuando se utilizan sistemas basados en algoritmos para impartir justicia, surgen profundas preguntas sobre la legitimidad y la transparencia, y, por ende, sobre la protección de los derechos fundamentales.

Estas consideraciones son particularmente importantes en relación con el derecho penal, donde las sentencias judiciales afectan directamente derechos constitucionales relacionados con la negación de la libertad personal, la falta de debido proceso y la presunción de inocencia. Un ejemplo típico de este tipo es el sistema COMPAS (Correctional Offender Management Profiling for Alternative Sanctions). Es un método para predecir la probabilidad de que las personas que cometen delitos continúen participando en conductas delictivas, y ha sido utilizado en varios estados de Estados Unidos. Aunque su propósito original es mejorar la eficiencia del sistema judicial, diversos estudios han cuestionado seriamente su transparencia, neutralidad y exactitud, señalando que este tipo de herramientas pueden reproducir y perpetuar desigualdades estructurales, especialmente por sesgos raciales y socioeconómicos (Angwin et al., 2016; Chouldechova, 2017).

No pasará mucho tiempo antes de que se necesiten reflexiones muy serias en Colombia sobre si debería comenzar a contemplar la incorporación de este tipo de tecnología en el sistema judicial. Por mucho que la tecnología de IA pueda ser capaz de perfeccionar lo que se está haciendo, su implementación en dominios que tocan derechos fundamentales demanda una evaluación vigilante. En este sentido, la Corte Constitucional ha especificado, por ejemplo, en la Sentencia T-323 de 2014, ciertos principios claros que

deben sustentar todos los usos judiciales de herramientas tecnológicas, tales como la transparencia, la explicabilidad, el control humano y la estricta adhesión al procedimiento.

Este documento se basa en la investigación pragmática y cualitativa y tiene como propósito determinar si un sistema como COMPAS podría operar en Colombia. Por esa razón, comenzará con un análisis doctrinal, jurisprudencial y normativo de fuentes nacionales e internacionales; y será cuidadoso en atender la advertencia de la criminología crítica latinoamericana sobre los peligros de adoptar tecnologías legales que se desarrollaron en circunstancias culturales y sociales distintas a las nuestras.

Por lo tanto, la pregunta que guía esta investigación es saber lo siguiente: ¿Debería implementarse en Colombia un software como COMPAS, considerando el principio del juez natural y los derechos fundamentales de los acusados?. Para este fin, en primer lugar, recopilamos la historia del sistema COMPAS, rastreando su desaprobación pública; luego, se realiza una revisión del caso estadounidense de *State v. Loomis* (2016); en tercer lugar, se examina su compatibilidad con el orden constitucional colombiano; finalmente, abordamos cuestiones éticas, culturales y tecnológicas que su eventual instalación en nuestro país podría presentar.

Sin embargo, este texto no busca de ninguna manera subestimar los avances tecnológicos. Se realiza un extenso análisis para sostener que cada progreso en el ámbito legal no debería contradecir ni desestabilizar aspectos fundamentales de nuestra estructura social. En realidad, la auténtica mejora en la justicia no se encuentra en un simple aumento de procedimientos, sino que emerge como resultado del fortalecimiento de los derechos humanos y la instauración de un sistema judicial que sea más responsable, inclusivo y democrático.

3. Capítulo 1: La inteligencia artificial y su incursión en el sistema penal.

3.1. Concepto y evolución de la inteligencia artificial

La inteligencia artificial (IA) ha sido objeto de diversas definiciones a lo largo de su evolución. Según la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO, 2021), la IA se refiere a “los sistemas tecnológicos que, para un conjunto dado de objetivos definidos por humanos, pueden percibir su entorno, procesar información, razonar y actuar para alcanzar esos objetivos”. Esta definición resalta la capacidad de la IA para interactuar con el mundo de manera autónoma, procesando datos y tomando decisiones de acuerdo con ciertos parámetros predefinidos por los humanos. Por otro lado, el Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia (MinTIC, 2020) la describe como “la simulación de procesos de inteligencia humana, como el aprendizaje (adquisición de información y reglas para su uso), el razonamiento (uso de reglas para llegar a conclusiones aproximadas o definitivas) y la autocorrección”. Ambas definiciones subrayan una característica esencial de la IA: su capacidad para aprender, adaptarse y corregirse con el tiempo, lo que la convierte en una herramienta cada vez más útil y poderosa en la sociedad contemporánea.

La historia de la inteligencia artificial se remonta a mediados del siglo XX, cuando los pioneros en computación comenzaron a imaginar la posibilidad de replicar las capacidades cognitivas humanas en una máquina. Los primeros intentos se basaron en modelos simbólicos, en los cuales las máquinas seguían reglas programadas explícitamente por los seres humanos. Un ejemplo temprano de estos esfuerzos fueron los sistemas expertos desarrollados en la década de 1970, que se utilizaron para resolver problemas complejos en áreas como el diagnóstico médico y la ingeniería. Según Russell y Norvig (2020), los sistemas expertos eran "sistemas computacionales diseñados para

imitar el proceso de toma de decisiones de un experto humano en un campo específico". Aunque estos sistemas fueron innovadores en su momento, tenían la limitación de depender completamente de reglas predefinidas y no podían adaptarse a nuevas situaciones sin intervención humana directa.

Un avance crucial en la IA llegó con el desarrollo del *machine learning* o aprendizaje automático, un campo introducido formalmente por Arthur Samuel en 1959. Samuel (1959) definió el *machine learning* como el "campo de estudio que da a las computadoras la capacidad de aprender sin ser programadas explícitamente". En lugar de depender de reglas predefinidas, los sistemas de *machine learning* permiten que las máquinas aprendan patrones y relaciones a partir de grandes volúmenes de datos. Este avance significó un cambio paradigmático, ya que las máquinas ya no solo realizaban tareas predefinidas, sino que podían aprender y mejorar su desempeño con el tiempo.

Con el crecimiento del *machine learning*, surgió el campo del *deep learning* o aprendizaje profundo, que ha transformado radicalmente la capacidad de las máquinas para realizar tareas complejas. El *deep learning* se basa en el uso de redes neuronales artificiales, que imitan las estructuras neuronales del cerebro humano para procesar información. LeCun, Bengio y Hinton (2015) destacaron que el *deep learning* "ha permitido avances notables en el reconocimiento de imágenes, el procesamiento de lenguaje natural y otras áreas que requieren el procesamiento de datos no estructurados". Gracias al aprendizaje profundo, las máquinas pueden realizar tareas como la identificación de objetos en imágenes, la traducción automática de idiomas y la predicción del comportamiento humano con una precisión sorprendente. Este avance ha permitido la creación de sistemas de IA capaces de realizar tareas complejas de manera autónoma, lo

que ha abierto nuevas posibilidades en campos tan diversos como la salud, la educación, la industria y la justicia.

A pesar de estos avances, la IA no es un campo exento de desafíos. Como señala Barocas y Selbst (2016), la IA no es neutral; sus resultados están profundamente influenciados por los datos con los que se entrenan los algoritmos y por las decisiones tomadas por los programadores. Este fenómeno se conoce como "sesgo algorítmico", y se refiere a la tendencia de los sistemas de IA a replicar y amplificar los prejuicios preexistentes en la sociedad. La investigación de ProPublica (Angwin et al., 2016) demostró que el sistema COMPAS, utilizado en el sistema de justicia penal estadounidense para predecir el riesgo de reincidencia, presenta una tasa desproporcionada de falsos positivos entre las personas afroamericanas, clasificándolas erróneamente como de alto riesgo en comparación con los individuos blancos. Este tipo de sesgo algorítmico plantea serias preocupaciones sobre la equidad y la justicia de los sistemas de IA, especialmente en contextos sensibles como la justicia penal.

Según Chouldechova (2017), los sesgos en los algoritmos pueden surgir no solo debido a los datos sesgados, sino también debido a los diseños inapropiados de los sistemas. En este sentido, los sistemas de IA pueden estar predispuestos a tomar decisiones que perpetúan las desigualdades existentes en la sociedad, lo que plantea serias preocupaciones sobre la justicia y la equidad. Además, Wexler (2018) advierte que estos sesgos, al estar ocultos bajo una fachada de "neutralidad tecnológica", pueden generar un efecto de legitimación de decisiones discriminatorias, lo que hace aún más difícil identificar y corregir los problemas.

En conclusión, la inteligencia artificial ha recorrido un largo camino desde sus primeras formulaciones en la década de 1950. Si bien los avances en *machine learning* y

deep learning han permitido a las máquinas realizar tareas que antes eran impensables, también han surgido importantes desafíos éticos y técnicos. La falta de transparencia en los algoritmos, los sesgos preexistentes en los datos y las decisiones de diseño de los programadores son factores que deben abordarse para garantizar que la IA se utilice de manera justa y responsable. Como se evidencia en los estudios sobre COMPAS, la implementación de la IA en sectores como la justicia penal debe ser cuidadosamente supervisada para evitar la perpetuación de desigualdades y prejuicios preexistentes. Solo mediante un enfoque ético y responsable se podrá aprovechar todo el potencial de la IA sin comprometer los principios fundamentales de justicia y equidad.

3.2. Fundamentos de la inteligencia artificial aplicada al Derecho penal

En el ámbito jurídico, la aplicación de la IA plantea desafíos particulares. Si bien estas tecnologías pueden optimizar procesos, también amenazan principios esenciales del Estado de Derecho, como la autonomía judicial, el debido proceso y la igualdad ante la ley (Constitución Política de Colombia, arts. 13 y 29). La inteligencia artificial ha emergido como una herramienta con alto impacto en múltiples sectores, incluyendo la justicia penal.

Aplicada al Derecho penal, la IA se ha comenzado a utilizar en herramientas de evaluación de riesgo criminal, que buscan predecir la probabilidad de reincidencia de los acusados y asistir al juez en decisiones sobre medidas preventivas, libertad condicional o penas. Estas herramientas procesan grandes cantidades de datos —como antecedentes penales, condiciones socioeconómicas y factores personales— y generan una “puntuación de riesgo” que puede influir considerablemente en la decisión final. Tal como advierte Ferrer Mac-Gregor (2021), más allá de su utilidad, lo que debe preocupar al Derecho es que estas decisiones tecnológicas no sustituyan el juicio, ni desdibujen los principios

constitucionales que garantizan un proceso justo y humano. Ahora bien, la evaluación de riesgo criminal busca estimar la probabilidad de que un individuo cometa nuevas conductas delictivas. Según Baird (2009), estas evaluaciones históricamente han sido realizadas mediante análisis profesionales, utilizando entrevistas, historiales clínicos, antecedentes penales y observaciones socioambientales.

En efecto, el uso de IA en la función jurisdiccional, sin los debidos controles, puede comprometer principios como el juez natural, la presunción de inocencia y el debido proceso, pilares del orden constitucional colombiano. Así lo ha advertido la Corte Constitucional en su Sentencia T-323 de 2024, al establecer que toda herramienta tecnológica utilizada en procesos judiciales debe observar los principios de transparencia, explicabilidad, proporcionalidad, y estar siempre bajo la supervisión de un juez (Corte Constitucional, 2024, ap. 69).

3.3. El sistema COMPAS: lógica, funcionamiento y críticas jurídicas

El sistema COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*) fue desarrollado a finales de los años noventa por la empresa Northpointe (hoy Equivant) en respuesta a la necesidad de los tribunales de contar con herramientas objetivas que midieran el riesgo de reincidencia. Inspirado en el modelo de “riesgo-necesidad-responsividad” de la criminología penológica, COMPAS integra tres tipos de instrumentos:

- **Riesgo:** evalúa la probabilidad de reincidencia general o violenta.
- **Necesidad:** identifica necesidades criminógenas (por ejemplo, abuso de sustancias).
- **Responsividad:** sugiere las modalidades de intervención más adecuadas según el perfil del individuo (Brennan, Dieterich & Ehret, 2009).

Para generar estas evaluaciones, el sistema utiliza más de 130 variables recabadas mediante un cuestionario semiestructurado y registros oficiales: historial delictivo, empleo, nivel educativo, relaciones personales y contexto comunitario. Un algoritmo propietario —no accesible al público— procesa estos datos y asigna una “puntuación de riesgo” que clasifica al acusado en categorías de bajo, medio o alto riesgo de reincidencia.

Sin embargo, su implementación ha sido objeto de críticas jurídicas severas. La primera preocupación se refiere a la opacidad del algoritmo, ya que COMPAS funciona como un sistema cerrado cuyo código fuente es propiedad de una empresa privada. Esto impide su auditoría por las partes procesales, el juez o los defensores, lo que vulnera el derecho a la defensa y el principio de contradicción (González, 2022).

A esto se suma la evidencia práctica de sesgos raciales. El estudio de Angwin et al. (2016) publicado por ProPublica mostró que COMPAS tiende a predecir incorrectamente que personas afroamericanas reincidirán en mayor proporción que personas blancas con antecedentes similares. Esta situación genera preocupación respecto a la igualdad ante la ley y a la discriminación indirecta. De igual manera, Chouldechova (2017) comprobó que la herramienta presenta tasas más altas de falsos positivos para personas negras, lo que puede derivar en decisiones judiciales arbitrarias y desproporcionadas.

En Colombia, donde existen profundas desigualdades sociales, económicas y culturales, la introducción de una tecnología como COMPAS sin un análisis contextual podría reproducir y agravar los sesgos estructurales ya presentes en el sistema penal. Como advierte Vallejos (2020), la importación de tecnologías jurídicas del norte global, sin una adaptación crítica a las realidades del sur, constituye una forma de “colonialismo tecnológico” que pone en riesgo los principios de justicia material y equidad.

3.4. Riesgos constitucionales y éticos de los algoritmos en el proceso penal

Desde una perspectiva constitucional, la implementación de sistemas algorítmicos en el proceso penal plantea riesgos significativos, que van más allá de la eficiencia técnica de los mismos. Uno de los principales riesgos es la posible vulneración del principio del juez natural, un principio fundamental consagrado en el artículo 29 de la Constitución Política de Colombia, que garantiza que toda persona sea juzgada por un juez imparcial, previamente establecido por la ley. Este principio exige que las decisiones judiciales se basen en la interpretación de normas jurídicas, y no en evaluaciones automáticas, por lo que delegar parte de la toma de decisiones a un algoritmo opaco podría afectar seriamente la legitimidad del proceso judicial. Como señala López (2019), el uso de herramientas algorítmicas opacas puede erosionar la figura del juez, transformándolo en un mero validador de lo que la tecnología determina, sin la posibilidad de comprender y justificar los razonamientos subyacentes de esa decisión. En este contexto, el juez pierde su función de garante de los derechos fundamentales del acusado y se convierte en un ejecutor de una lógica tecnológica que no siempre es accesible ni comprensible para las partes involucradas.

Además, el principio de explicabilidad, esencial para garantizar la transparencia y la justicia en el proceso judicial, se ve amenazado por la opacidad inherente a muchos sistemas algorítmicos. El Consejo de Europa (2020) ha advertido que ningún sistema automatizado debería ser adoptado sin que permita la trazabilidad de sus procesos. Esto significa que es imprescindible que los algoritmos sean comprensibles y auditables, de modo que los jueces, abogados y defensores puedan entender cómo se llegó a una determinada conclusión. De lo contrario, el derecho de defensa se ve gravemente afectado, ya que el acusado se encuentra ante decisiones que no puede impugnar ni

discutir adecuadamente. La falta de explicabilidad puede, en última instancia, minar la confianza en el sistema judicial, pues los involucrados en el proceso no pueden estar seguros de que las decisiones se basen en principios legales coherentes y transparentes.

El Alto Comisionado de la ONU (2021) también ha señalado que la falta de transparencia tecnológica puede derivar en violaciones sistemáticas de derechos humanos, ya que, al ser opacos los mecanismos de toma de decisiones, los afectados no tienen la capacidad de cuestionar ni defender sus derechos frente a posibles errores, sesgos o discriminación inherentes al sistema. Esta preocupación se amplía cuando se considera que los algoritmos, al ser utilizados en el ámbito penal, tienen un impacto directo en la libertad y los derechos fundamentales de las personas, lo que exige un mayor escrutinio y control en su implementación.

Desde una perspectiva ética, la introducción de algoritmos en el proceso penal también plantea el riesgo de “automatizar la desigualdad”. Eubanks (2018) describe este fenómeno como el proceso mediante el cual las tecnologías, al ser implementadas sin un análisis crítico y consciente de las desigualdades preexistentes, terminan reforzando los sesgos institucionales ya presentes en la sociedad. En los sistemas de justicia penal, donde históricamente han existido desigualdades raciales, económicas y sociales, el uso de algoritmos puede perpetuar o incluso agravar estas disparidades, al tomar decisiones basadas en datos que ya están contaminados por sesgos previos. Como advierte Eubanks (2018), lo que se presenta como una herramienta neutral y objetiva puede ocultar tras de sí mecanismos de discriminación, especialmente en contextos donde ya existe una desigualdad estructural.

En países como Colombia, donde las desigualdades sociales y económicas están profundamente arraigadas, resulta indispensable cuestionarse no solo si la herramienta

tecnológica funciona, sino también a quién beneficia y a quién perjudica. ¿Es justo que un algoritmo diseñado para predecir el riesgo de reincidencia en personas acusadas de un crimen pueda estar basado en datos que reflejan prejuicios raciales, económicos o de clase? La crítica ética se centra en la necesidad de evaluar el impacto real de estas tecnologías, más allá de su aparente eficiencia. La automatización de decisiones en el ámbito penal debe considerar los principios fundamentales de la justicia, la equidad y la no discriminación, para evitar que, bajo la apariencia de imparcialidad tecnológica, se perpetúen o incluso se incrementen las injusticias existentes.

Por lo tanto, el uso de algoritmos en el proceso penal debe ser considerado cuidadosamente, no solo desde el punto de vista técnico y de eficiencia, sino también en términos de su compatibilidad con los derechos fundamentales consagrados en la Constitución y los estándares internacionales en materia de derechos humanos. La transparencia, la explicabilidad y la supervisión ética son aspectos cruciales para evitar que la implementación de estas tecnologías en el ámbito judicial implique un retroceso en los avances hacia una justicia verdaderamente equitativa y accesible para todos.

4. Capítulo 2: El caso *State v. Loomis* y el debate constitucional en EE.

UU.

4.1. Contexto y Antecedentes

En abril de 2012, Brent Loomis fue acusado en Wisconsin de cinco delitos relacionados con un tiroteo. Los cargos incluyeron posesión de arma de fuego, tenencia de un rifle de cañón corto, intento de eludir a un agente de tránsito, poner en peligro la seguridad pública y manejar un vehículo motorizado sin el consentimiento del propietario. Aunque el acusado negó su participación en el tiroteo, admitió haber conducido el vehículo y haber intentado escapar del oficial. Eventualmente, llegó a un

acuerdo con la Fiscalía, aceptando únicamente los cargos de poner en riesgo la seguridad pública y conducir un vehículo ajeno sin autorización. Tras su condena en el tribunal de primera instancia, el juez ordenó una evaluación con COMPAS para decidir la modalidad y las condiciones de su libertad condicional. El sistema arrojó una puntuación de alto riesgo, y en consecuencia se imponiendo así una sentencia de seis años de prisión junto con cinco años de libertad vigilada (Avella, Sanabria-Moyano, & Dinas-Hurtado, 2022, p. 284).

La defensa de Loomis impugnó esta decisión ante la Corte de Apelaciones de Wisconsin, sosteniendo que el uso del algoritmo COMPAS vulneraba el derecho al debido proceso por tres motivos principales: *“(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”* (Avella, Sanabria-Moyano, & Dinas-Hurtado, 2022, p. 285).

La Corte de Apelaciones mantuvo la decisión, y finalmente la Corte Suprema de Wisconsin aceptó el caso, convirtiendo a *Loomis* en el primer precedente relevante sobre admisión de pruebas algorítmicas en el sistema penal estadounidense.

4.2. Decisión de la Corte Suprema de Wisconsin y riesgos para el debido proceso y derecho de defensa

El primero de junio de 2016, la Corte Suprema de Wisconsin, en el caso *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016), confirmó en una votación dividida, 4 votos a favor y 3 en contra, que la puntuación de riesgo generada por COMPAS podía ser admitida como prueba complementaria en la fase de determinación de la libertad condicional, sin que ello implicara la sustitución del juicio humano. El tribunal enfatizó que COMPAS

debía entenderse como una herramienta informativa destinada a “apoyar, no reemplazar, la discreción judicial” (State v. Loomis, 2016, p. 755). Este primer punto puso de manifiesto que, aunque el algoritmo carece de transparencia total, su uso no viola automáticamente el debido proceso siempre que el juez mantenga la última palabra y motive sus decisiones con base en las normas aplicables.

En segundo lugar, la Corte impuso condiciones específicas para su empleo. Primero, el juez debía advertir de manera explícita a la defensa y al acusado que la puntuación de COMPAS no era vinculante ni definitiva, sino meramente indicativa de un nivel de riesgo estimado (García & Sherman, 2018). Segundo, se prohibió fundamentar la privación o concesión de la libertad condicional únicamente en ese puntaje: la decisión final debía articularse también con otras pruebas como informes sociales, antecedentes personales y testimonios. Con ello, la mayoría buscó equilibrar la potencial utilidad del algoritmo con las garantías procesales que exige el principio de contradicción y la carga probatoria propia del derecho penal (State v. Loomis, 2016).

No obstante, la mayoría rechazó exigir la divulgación del código fuente de COMPAS o el detalle pormenorizado de su lógica interna, al considerarlo un secreto comercial protegido. Según el fallo, la divulgación de esos elementos pondría en riesgo la innovación y competitividad de la empresa desarrolladora, sin que existiera un mandato constitucional que obligara a revelar algoritmos propietarios (García & Sherman, 2018). Esta decisión dejó abierta la tensión entre el derecho a la información de las partes procesales y el interés económico de los proveedores tecnológicos.

La opinión disidente, firmada por los jueces Abrahamson, Roggensack y Ziegler, planteó un fuerte reparo al admitir pruebas algorítmicas sesgadas. Para ellos, permitir que un “código negro” influya en decisiones de libertad era contrario al debido proceso, pues

la defensa carece de medios para cuestionar la validez estadística, la relevancia de las variables o la posible presencia de sesgos raciales (Harcourt, 2015). Asimismo, advirtieron que, sin acceso al algoritmo, no existe forma de verificar si COMPAS respeta principios de igualdad ante la ley o si infringe la presunción de inocencia al cargar en el acusado el peso de demostrar que el sistema está equivocado.

Adicionalmente, la Corte abordó tres críticas específicas planteadas por la defensa de Loomis. En primer lugar, aunque se alegó que el uso de COMPAS vulneraba el derecho a ser juzgado con base en información precisa —dado que el carácter patentado del algoritmo impedía evaluar su exactitud—, el Tribunal consideró que la situación era distinta al precedente *Gardner v. Florida*, ya que Loomis pudo acceder a los resultados del informe de riesgo, aunque no a su lógica interna (Avella, Sanabria-Moyano, & Dinas-Hurtado, 2022, p. 289).

En segundo lugar, respecto a la alegada afectación al derecho a una sentencia individualizada, la Corte reconoció que COMPAS trabaja sobre estadísticas grupales, pero precisó que las evaluaciones de riesgo deben ser consideradas solo como una herramienta adicional, sin desplazar el análisis particular de cada caso (Avella, Sanabria-Moyano, & Dinas-Hurtado, 2022, p. 290).

Finalmente, en cuanto al argumento de que el software introducía un sesgo de género, el fallo concluyó que incorporar la variable de género no resultaba discriminatorio per se, sino que respondía a criterios de precisión estadística, en atención a las diferencias empíricas en las tasas de reincidencia entre hombres y mujeres (Avella, Sanabria-Moyano, & Dinas-Hurtado, 2022, p. 290).

No obstante, se dejó en evidencia que el uso de herramientas algorítmicas como COMPAS plantea riesgos latentes para el debido proceso, el principio de contradicción y

la presunción de inocencia, en tanto limita la capacidad de los procesados para impugnar de manera efectiva la información que influye en su condena.

A pesar de las cautelas impuestas, investigaciones posteriores mostraron que muchos jueces conferían un peso excesivo a la “objetividad” de COMPAS, tendiendo a considerar la puntuación como un factor determinante en la práctica, más allá de las advertencias formales (Angwin et al., 2016). Esta dinámica reveló la fuerza persuasiva de los algoritmos en la toma de decisiones y la dificultad de contener su influencia mediante simples advertencias procesales.

En última instancia, *State v. Loomis* reafirmó que los sistemas predictivos pueden integrarse al arsenal probatorio, siempre que existan salvaguardas procesales — advertencias, prohibiciones de uso exclusivo y, al menos en teoría, información mínima sobre su funcionamiento—. Sin embargo, al preservar la confidencialidad del código fuente, la Corte dejó sin resolver la síntesis entre eficiencia algorítmica y garantías constitucionales, abriendo la puerta a posteriores debates legislativos y regulatorios sobre la transparencia y la rendición de cuentas de la IA en justicia (Wexler, 2018).

4.3. Críticas doctrinales y fallas estructurales de COMPAS

La imposibilidad de acceder al algoritmo subyacente en COMPAS ha sido, de manera unánime, uno de los aspectos más vehementemente criticados por la doctrina jurídica. Este fenómeno ha sido calificado por diversos autores como una manifestación clara de lo que se conoce como una "caja negra" (Harcourt, 2015), lo que significa que el proceso mediante el cual se generan las decisiones del sistema es completamente opaco y no puede ser examinado ni cuestionado de manera efectiva por las partes involucradas en el proceso judicial. Según Harcourt, este nivel de opacidad atenta directamente contra los principios fundamentales del debido proceso, particularmente contra la posibilidad de

cuestionar la prueba, analizar su fundamento lógico y someterla a un proceso de contradicción. El hecho de que no sea posible comprender cómo un algoritmo llega a una conclusión sobre el riesgo de una persona condenada plantea serias dudas sobre la legitimidad y la justicia de las decisiones tomadas con su asistencia.

El principio de contradicción, que implica el derecho de las partes a ser escuchadas y a contradecir la prueba en su contra, es esencial en cualquier sistema judicial que se base en los principios de transparencia y equidad. Sin embargo, como bien señala López (2019), la opacidad tecnológica de COMPAS impide que la defensa ejerza este derecho de manera plena. Al no contar con acceso a los parámetros que el algoritmo utiliza para calificar a los acusados, la defensa se ve privada de los medios necesarios para cuestionar las conclusiones del sistema. Este vacío de información sobre el funcionamiento interno del algoritmo socava la capacidad de los abogados para presentar argumentos sólidos en defensa de sus clientes y, por ende, afecta la calidad del juicio y la protección de los derechos fundamentales del acusado.

En este sentido, la doctrina también ha señalado que la utilización de COMPAS comporta una inversión indebida de la carga probatoria. Tradicionalmente, en cualquier procedimiento judicial, es la parte acusadora la que tiene la carga de probar la culpabilidad del acusado más allá de toda duda razonable. No obstante, con el uso de COMPAS, se invierte esta lógica: es la defensa quien se ve obligada a demostrar que el algoritmo está errado o que la puntuación generada por el sistema no refleja la realidad del caso. Este fenómeno, como argumentan Gómez y Rojas (2019), resulta materialmente imposible de llevar a cabo si no se conocen los parámetros y la metodología interna del algoritmo. La defensa se encuentra, por lo tanto, en una posición sumamente desventajosa, ya que no dispone de los elementos necesarios para refutar o impugnar una prueba que tiene un

carácter altamente técnico y que, por ende, resulta completamente inaccesible para la mayoría de los abogados. Este giro en la carga probatoria no solo vulnera el principio de presunción de inocencia, sino que también desplaza la carga de la prueba en detrimento del procesado, contraviniendo uno de los pilares fundamentales del derecho penal.

Por otro lado, las críticas también se centran en el impacto de los sesgos raciales presentes en el algoritmo de COMPAS, que han sido revelados mediante diversos estudios empíricos y análisis estadísticos. El reporte de ProPublica (Angwin et al., 2016) expuso que COMPAS presenta una tasa significativamente más alta de falsos positivos entre las personas afroamericanas, clasificándolas como de "alto riesgo" más frecuentemente que a las personas blancas, incluso cuando ambas se encuentran en condiciones similares. Este sesgo racial no solo es un error técnico del algoritmo, sino que tiene repercusiones graves sobre la vida de los acusados, afectando la probabilidad de que reciban sentencias más severas y perjudicando su posibilidad de reintegrarse a la sociedad. La investigación de Angwin et al. (2016) señaló que las personas afroamericanas son más propensas a ser clasificadas como de alto riesgo, lo que incrementa las probabilidades de encarcelamiento, aunque el riesgo real que presentan pueda ser inferior al de otros grupos raciales. Este hallazgo subraya cómo los algoritmos pueden perpetuar, e incluso acentuar, los sesgos raciales ya existentes en el sistema judicial, en lugar de ofrecer una solución objetiva y justa.

Por su parte, Chouldechova (2017) corroboró estos hallazgos desde un enfoque estadístico, señalando que los sesgos raciales no solo son una anomalía aislada, sino que están profundamente arraigados en los datos con los que se alimentan estos sistemas. En otras palabras, los algoritmos como COMPAS reflejan las desigualdades y los prejuicios de la sociedad, ya que están diseñados para aprender de patrones históricos, los cuales

están permeados por discriminación y exclusión. Este tipo de sesgos plantea un riesgo grave, ya que, al ser invisibles en el proceso de toma de decisiones, pueden resultar en una "legitimación" de decisiones discriminatorias que, de otro modo, serían cuestionadas de manera abierta. Wexler (2018) también abordó este punto, advirtiendo que estos sesgos raciales, al estar ocultos bajo la apariencia de una neutralidad tecnológica, generan un efecto de validación y refuerzo de las decisiones discriminatorias. La opacidad de COMPAS no solo impide detectar estos sesgos, sino que los esconde bajo una capa de "objetividad" que podría resultar aún más perjudicial al darles una falsa legitimidad.

Desde una perspectiva crítica, estos sesgos no son meros errores aislados, sino que reflejan una falla estructural más profunda en el diseño y la implementación de COMPAS. Como bien señala la doctrina, la dependencia de datos históricos sesgados y la falta de un marco normativo adecuado para auditar y corregir estos algoritmos implica que COMPAS no puede ofrecer una solución justa o imparcial. En lugar de proporcionar una herramienta que respete el principio de igualdad ante la ley, COMPAS perpetúa y magnifica las desigualdades existentes, especialmente aquellas relacionadas con la raza, el género y el estatus socioeconómico. Este es, sin lugar a dudas, uno de los mayores problemas del sistema y una de las razones por las cuales su utilización en el contexto judicial debe ser cuestionada y revisada con urgencia.

En resumen, las críticas doctrinales hacia COMPAS señalan que la opacidad del algoritmo, la inversión de la carga probatoria y los sesgos raciales inherentes al sistema representan fallas estructurales profundas que vulneran principios esenciales del derecho penal y del debido proceso. Estos problemas no solo afectan la validez de las decisiones tomadas con base en COMPAS, sino que también socavan la confianza en el sistema judicial, al exponerlo a la posibilidad de decisiones arbitrarias y discriminatorias. La falta

de transparencia y la presencia de sesgos en los algoritmos subraya la necesidad de revisar con urgencia su utilización en el ámbito judicial, para garantizar que la justicia se administre de manera imparcial, transparente y respetuosa de los derechos fundamentales de todas las personas, sin importar su raza, género o estatus socioeconómico.

4.4. Estándares Internacionales y riesgos para sistemas jurídicos latinoamericanos

El caso *State v. Loomis* debe ser analizado no solo desde una perspectiva interna del sistema jurídico estadounidense, sino también a la luz de los estándares internacionales de derechos humanos que resultan vinculantes o al menos orientadores para los sistemas de justicia del continente. El artículo 8.1 de la Convención Americana sobre Derechos Humanos consagra el derecho a un juicio justo, público, con garantías y dentro de un plazo razonable, ante un juez competente, independiente e imparcial. La Corte Interamericana de Derechos Humanos ha interpretado que este estándar incluye el acceso a la totalidad de las pruebas, su contradicción efectiva y la posibilidad de comprender y cuestionar los elementos sobre los que se construye una decisión judicial (Corte IDH, *Maldonado Ordoñez v. Guatemala*, 2016).

En esa línea, el uso de algoritmos cerrados como COMPAS, cuyo código fuente y lógica de funcionamiento no pueden ser auditados por las partes procesales, entra en tensión con este estándar. La Relatoría Especial de la ONU sobre la independencia de magistrados y abogados advirtió en su informe de 2022 que ninguna herramienta algorítmica puede emplearse en decisiones que afecten derechos fundamentales si no cumple con los principios de explicabilidad, auditabilidad, supervisión humana y trazabilidad (Naciones Unidas, 2022). De lo contrario, se compromete la legitimidad del proceso y se priva a la defensa de herramientas efectivas para ejercer su derecho.

A este problema estructural se suma el riesgo de asumir que la IA en este contexto constituye una “objetividad” técnica como garantía de imparcialidad. Barocas y Selbst (2016) han advertido que, los algoritmos no son entes neutros: están diseñados por personas, alimentados con datos históricamente sesgados y ejecutados en sistemas institucionales atravesados por dinámicas de poder y desigualdad. La aparente neutralidad matemática de herramientas como COMPAS puede ocultar y legitimar prácticas discriminatorias, dificultando su identificación y corrección.

5. Capítulo 3: Análisis jurídico de la posible implementación de COMPAS en Colombia

5.1. El principio del juez natural en la Constitución Colombiana

El principio del juez natural constituye uno de los pilares esenciales del Estado social y democrático de derecho consagrado en la Constitución Política de Colombia. Se desprende directamente del artículo 29, que consagra el derecho al debido proceso, garantizando que "nadie podrá ser juzgado sino conforme a leyes preexistentes, ante juez o tribunal competente, con observancia plena de la forma propia de cada juicio" (Constitución Política de Colombia, 1991, art. 29). Este mandato implica la existencia de un juez predeterminado por la ley, independiente e imparcial, cuya competencia debe estar definida conforme a criterios generales, abstractos y objetivos.

La Corte Constitucional ha reiterado en múltiples oportunidades que el principio del juez natural no solo protege contra la creación de tribunales ad hoc, sino que también garantiza la imparcialidad, la independencia judicial y el respeto pleno de las garantías procesales. En la Sentencia C-373 de 2002, la Corte señaló que “el juez natural es aquel previamente establecido por la ley para conocer de un determinado asunto, evitando así

designaciones caprichosas o hechas para casos concretos” (Corte Constitucional, Sentencia C-373 de 2002, M.P. Rodrigo Escobar Gil).

Asimismo, en la Sentencia T-949 de 2003, el Tribunal Constitucional enfatizó que el principio del juez natural implica no solo la existencia formal de un juez competente, sino también el respeto de las garantías propias del proceso, como el derecho de defensa, la contradicción de la prueba y el acceso a una administración de justicia imparcial y efectiva (Corte Constitucional, Sentencia T-949 de 2003, M.P. Rodrigo Escobar Gil).

Este desarrollo jurisprudencial se ha visto reforzado en decisiones posteriores. En la Sentencia C-252 de 2001, se sostuvo que la designación previa del juez competente protege contra eventuales manipulaciones del proceso por parte del Estado, asegurando que todos los ciudadanos puedan prever razonablemente qué autoridad resolverá sus controversias (Corte Constitucional, Sentencia C-252 de 2001, M.P. Alfredo Beltrán Sierra).

En el mismo sentido, la Sentencia T-063 de 2017 precisó que el juez natural debe reunir condiciones de competencia funcional, subjetiva y material, y que su alteración arbitraria puede configurar una causal de nulidad (Corte Constitucional, Sentencia T-063 de 2017, M.P. Luis Guillermo Guerrero Pérez). Finalmente, en la Sentencia SU-159 de 2021, el Alto Tribunal reafirmó que la función jurisdiccional no puede ser delegada, ni siquiera parcialmente, en terceros que no estén sometidos al control constitucional y legal (Corte Constitucional, Sentencia SU-159 de 2021, M.P. Cristina Pardo Schlesinger).

Desde el derecho internacional de los derechos humanos, el principio del juez natural también ha sido consagrado en instrumentos como la Convención Americana sobre Derechos Humanos, cuyo artículo 8.1 garantiza el derecho a ser oído “por un juez o tribunal competente, independiente e imparcial, establecido con anterioridad por la ley”

(Convención Americana sobre Derechos Humanos, 1969, art. 8.1). La Corte Interamericana de Derechos Humanos ha desarrollado este estándar en casos como *Apitz Barbera y otros (“Corte Primera de lo Contencioso Administrativo”) vs. Venezuela*, donde indicó que el juez natural “constituye una garantía contra la arbitrariedad y una condición para la confianza pública en la administración de justicia” (Corte IDH, 2008, párr. 55).

A partir de este sólido marco constitucional e internacional, resulta evidente que la eventual implementación de sistemas automatizados como COMPAS en Colombia debe ser analizada con extrema cautela. La utilización de algoritmos cerrados y opacos, cuya lógica de funcionamiento no puede ser auditada por las partes ni controlada judicialmente, comprometería gravemente el principio del juez natural, en tanto introduciría un “tercer decisor” ajeno al aparato judicial.

Este riesgo ha sido advertido expresamente por la Corte Constitucional en la Sentencia T-323 de 2024, donde se alertó que la función judicial no puede ser sustituida por la validación automática de recomendaciones algorítmicas, pues ello vulneraría principios como la autonomía judicial, la motivación racional de las decisiones y la protección efectiva de los derechos fundamentales (Corte Constitucional, Sentencia T-323 de 2024, ap. 58).

En definitiva, la implementación de sistemas como COMPAS, sin las debidas salvaguardas de transparencia, auditabilidad y supervisión humana significativa, resultaría incompatible con el principio del juez natural y, en consecuencia, con las bases mismas del Estado constitucional de derecho en Colombia.

5.2. Análisis de la Sentencia T-323 de 2024

La Sentencia T-323 de 2024 de la Corte Constitucional representa un precedente estructural para el tratamiento jurídico de la inteligencia artificial (IA) en Colombia, especialmente en relación con su aplicación en contextos judiciales. En este fallo, el Alto Tribunal analiza con profundidad la compatibilidad de las herramientas algorítmicas con los derechos fundamentales, y traza una línea jurisprudencial clara que delimita su uso a condiciones estrictas de explicabilidad, transparencia, y proporcionalidad.

El proceso de tutela tuvo origen en una decisión judicial proferida en el marco de un proceso civil, en el que se debatía la imposición de una medida cautelar. En el análisis del caso, se evidenció que el juez de segunda instancia hizo uso de la herramienta ChatGPT 3.5 como fuente auxiliar para sustentar su decisión sin informar a las partes ni permitirles conocer los criterios técnicos que sustentaban dicho resultado. Esta herramienta fue desarrollada por OpenAI, una empresa estadounidense dedicada a la investigación y desarrollo de inteligencia artificial, fundada en el año 2015. Según la propia compañía, “los modelos de la serie GPT de OpenAI son sistemas de IA rápidos, versátiles y rentables, diseñados para comprender el contexto, generar contenido y razonar sobre textos, imágenes y otros formatos de entrada” (OpenAI, 2023). La parte accionante alegó que el uso de esta herramienta vulneraba su derecho al debido proceso, en especial por desconocer el principio de contradicción de la prueba, la motivación judicial y el acceso efectivo a la administración de justicia.

Si bien no se trató de una IA jurídica especializada, el hecho de que el juez copiara parte de la motivación directamente de un modelo de lenguaje no supervisado ni diseñado para funciones judiciales, generó una fuerte preocupación constitucional. La Corte señaló que se encontraron dudas relevantes sobre si quien emitió la decisión fue verdaderamente un juez de la República o una inteligencia artificial, y si la decisión fue debidamente

motivada o producto de alucinaciones y sesgos generados por la IA (Corte Constitucional, 2024, ap. 58).

Este contexto materializó una problemática que ya había sido advertida en pronunciamientos anteriores. En la Sentencia T-119 de 2013, la Corte había indicado que “el uso de instrumentos técnicos por parte del juez debe estar subordinado al juicio jurídico, no puede sustituirlo ni desplazarlo” (Corte Constitucional, 2013). También, en la T-088 de 2018, la Corte insistió en que la autonomía judicial no puede ser formal, sino sustantiva: el juez debe mantener su capacidad crítica y decisoria incluso cuando acude a expertos o a herramientas técnicas.

Ahora bien, en la T-323 de 2024, la Corte Constitucional estableció que el uso de herramientas de IA no es inconstitucional en sí mismo, pero sí lo es cuando se omite garantizar los derechos fundamentales de las partes y el carácter deliberativo, público y contradictorio del proceso. En esa dirección, el fallo no estableció un catálogo cerrado de principios, pero desarrolló un conjunto de criterios orientadores constitucionales que, en conjunto, delimitan las condiciones mínimas para el uso legítimo de estas tecnologías en el ejercicio de la función judicial.

Uno de los criterios fundamentales es el de explicabilidad, entendido como la obligación de asegurar que tanto el juez como las partes procesales comprendan el funcionamiento del sistema de IA, la lógica detrás de sus recomendaciones y los datos en que se basa. Este principio se conecta directamente con el derecho a la defensa y a la contradicción de los elementos probatorios. La Corte también destacó el principio de transparencia, al señalar que no es admisible el uso de sistemas cerrados o de “caja negra” cuya lógica interna esté protegida por secretos comerciales, pues esto impediría ejercer

un control adecuado sobre su validez, exactitud o sesgos (Corte Constitucional, 2024, ap. 370–373).

A estos principios se suman el deber de protección de la privacidad, en especial frente al uso de datos personales y sensibles en entornos procesales, y la exigencia de supervisión humana significativa, que implica que el juez conserve la titularidad plena del acto decisorio. Las herramientas de IA, advierte la Corte, no pueden convertirse en sustitutos del juicio jurídico ni limitar la motivación racional, autónoma y pública que debe caracterizar toda providencia judicial. De igual modo, se resaltó el principio de responsabilidad judicial, según el cual la autoridad judicial no puede escudarse en decisiones automatizadas para justificar su fallo, sino que debe responder por su razonamiento, aún si emplea apoyos técnicos.

Asimismo, la Corte hizo referencia a estándares internacionales, como los establecidos por la UNESCO (2021) y por la Relatoría Especial de la ONU sobre la independencia de jueces y magistrados (2022), que recomiendan que las herramientas de IA solo pueden utilizarse bajo condiciones de auditoría, control humano y no discriminación.

La Corte concluyó que cualquier herramienta tecnológica que limite o sustituya las garantías del debido proceso, o que impida el ejercicio efectivo del derecho de defensa, es incompatible con el modelo de justicia constitucional colombiano. En palabras del fallo: “el acceso a la justicia debe preservarse como un espacio de deliberación, no como una instancia de validación algorítmica” (Corte Constitucional, 2024, ap. 83).

5.3. Compatibilidad normativa y jurisprudencial con la IA judicial

El análisis sobre la posible incorporación del sistema COMPAS en la justicia penal colombiana no puede agotarse en un enfoque técnico o de eficiencia, sino que debe

partir de una lectura sistemática del orden jurídico vigente. La pregunta no es únicamente si es posible usar inteligencia artificial en la administración de justicia, sino si dicho uso resulta compatible con la estructura normativa y jurisprudencial colombiana, especialmente con los principios del Estado social de derecho, el debido proceso, la autonomía judicial y el respeto a la dignidad humana.

Desde el punto de vista normativo, Colombia no cuenta con una legislación específica que regule de forma detallada el uso de inteligencia artificial en el ámbito judicial. Sin embargo, existen normas y principios generales plenamente aplicables. En primer lugar, la Constitución Política establece en su artículo 228 que “la administración de justicia es función pública” y que “las decisiones deben estar debidamente motivadas y fundadas en derecho”, lo cual implica que cualquier intervención de un sistema automatizado debe ser comprensible, jurídicamente coherente y auditable.

El Código General del Proceso (Ley 1564 de 2012) exige que toda actuación judicial se funde en prueba legal, accesible a las partes y susceptible de contradicción (arts. 9 y 10). Esto supone que los elementos aportados por una herramienta como COMPAS, para ser válidamente considerados por el juez, deben cumplir con los mismos requisitos de publicidad, contradicción y licitud que cualquier otro medio de prueba. Por su parte, la Ley 1581 de 2012 sobre protección de datos personales impone límites al tratamiento de datos sensibles, especialmente en contextos estatales, exigiendo consentimiento informado y medidas de seguridad reforzadas para proteger la privacidad de los titulares. Dado que COMPAS opera sobre datos socioeconómicos y conductuales, su uso sin consentimiento informado ni supervisión judicial efectiva podría vulnerar derechos como el *habeas data* y la autodeterminación informativa.

Desde la perspectiva jurisprudencial, la Corte Constitucional ha sido enfática en señalar que la tecnología no puede sustituir ni debilitar el juicio autónomo del juez. En la Sentencia SU-159 de 2021, el Alto Tribunal recalcó que “el operador jurídico no puede delegar en sistemas automáticos la función de decidir, ni actuar como un validador mecánico de recomendaciones tecnológicas”. Asimismo, en la Sentencia T-088 de 2018, se insistió en que las decisiones judiciales deben ser producto de un razonamiento jurídico completo, fundado en derecho y respetuoso del principio de contradicción. En la reciente Sentencia T-323 de 2024, la Corte fijó estándares constitucionales claros para el uso de herramientas algorítmicas en la justicia, destacando los principios de explicabilidad, transparencia, protección de datos, responsabilidad judicial y supervisión humana significativa.

A estos pronunciamientos se suman advertencias de otros órganos estatales. La Procuraduría General de la Nación, en su *Documento técnico sobre inteligencia artificial y justicia* (2023), sostuvo que la introducción de algoritmos en la justicia exige una regulación previa clara, criterios éticos, formación técnica a los jueces y mecanismos de supervisión institucional. El Consejo de Estado, por su parte, en sus recomendaciones de 2022, señaló que el uso de IA no puede reemplazar el juicio humano, ni comprometer el principio de igualdad ante la ley ni la motivación razonada de las decisiones.

En este contexto, resulta evidente que el sistema COMPAS —al menos en su versión actual— no se ajusta a los parámetros normativos ni jurisprudenciales colombianos. Su carácter cerrado, la falta de explicabilidad, la dependencia de datos sensibles y la imposibilidad de ser auditado por las partes contradicen principios constitucionales como la legalidad, la publicidad de la prueba, la autonomía judicial, el debido proceso y el derecho de defensa.

En conclusión, a la luz de los estándares constitucionales definidos por la Corte Constitucional en la Sentencia T-323 de 2024, la implementación de un sistema como COMPAS no sería compatible con el marco jurídico colombiano, salvo que se cumplan estrictamente los requisitos de acceso, control, auditabilidad, protección de derechos fundamentales y soberanía judicial. Como lo señala la propia sentencia: “el acceso a la justicia debe preservarse como un espacio de deliberación y razonamiento jurídico, no como una instancia de validación automática de recomendaciones generadas por máquinas” (Corte Constitucional, 2024, ap. 83).

6. Capítulo 4: Desafíos éticos, culturales y tecnológicos en el contexto colombiano

6.1. Sesgos algorítmicos y diversidad sociocultural en Colombia

La aplicación de sistemas de inteligencia artificial (IA) en la justicia penal colombiana implica un riesgo crítico de reproducción de sesgos estructurales en un entorno caracterizado por profundas desigualdades sociales, económicas y étnico-culturales. Colombia, como Estado pluriétnico y multicultural, reconoce en su Constitución Política que la diversidad cultural es un valor fundamental que debe ser protegido y promovido (Constitución Política de Colombia, 1991, art. 7). Esto implica que la administración de justicia debe adoptar un enfoque diferenciado, respetuoso de las particularidades culturales y sociales de cada individuo.

En contextos judiciales, los algoritmos predictivos como COMPAS representan un riesgo significativo de profundizar las brechas existentes. Estos sistemas son entrenados utilizando datos históricos que, inevitablemente, reflejan prácticas discriminatorias previas, sesgos institucionales y desigualdades sistémicas. ProPublica documentó que COMPAS generaba falsos positivos en personas afroamericanas en un 45

%, frente a un 23 % en personas blancas (Angwin et al., 2016). Del mismo modo, Chouldechova (2017) demostró que, incluso cuando los modelos cumplen métricas globales de precisión, pueden producir impactos desproporcionados que afectan el derecho a la igualdad y no discriminación.

Estos hallazgos adquieren una gravedad especial en el contexto colombiano. Según datos del Departamento Administrativo Nacional de Estadística (DANE, 2023), comunidades afrocolombianas, indígenas y campesinas presentan mayores índices de pobreza multidimensional, desplazamiento forzado, baja escolaridad y acceso limitado a servicios públicos. Variables como la zona de residencia, el nivel educativo o el tipo de ocupación, frecuentemente utilizadas como predictores en modelos algorítmicos, podrían estigmatizar de manera indirecta a estas poblaciones, consolidando un patrón de discriminación estructural.

Barocas y Selbst (2016) advierten que los sistemas de IA, al basarse en correlaciones estadísticas y no en juicios de valor jurídico, tienden a perpetuar sesgos históricos sin ofrecer mecanismos de corrección. Es decir, la “neutralidad” matemática de los algoritmos es ilusoria, pues dependen de los datos que reciben y de los marcos interpretativos bajo los cuales son programados.

En esta línea, la Corte Constitucional colombiana ha sido enfática al advertir que la administración de justicia debe reconocer los contextos culturales, históricos y territoriales de las personas que acceden a ella. En la Sentencia T-622 de 2016, donde reconoció al río Atrato como sujeto de derechos, el Tribunal insistió en la necesidad de adoptar un enfoque de protección especial hacia comunidades étnicas marginadas. Asimismo, en la Sentencia T-033 de 2022, la Corte reiteró que la igualdad real y efectiva

exige medidas diferenciadas que reconozcan las condiciones de vulnerabilidad y exclusión estructural que afectan a sectores históricamente discriminados.

La incorporación de IA en decisiones penales, sin salvaguardas adecuadas, podría entonces consolidar una “igualdad formal” que, lejos de remediar desigualdades, las perpetúe bajo una apariencia de objetividad técnica. Como advierte De Sousa Santos (2010), uno de los riesgos de los sistemas modernos es reciclar injusticias estructurales bajo nuevos lenguajes de legitimación, un fenómeno que la Criminología del Sur ha denunciado ampliamente.

En el plano internacional, la Relatoría Especial de Naciones Unidas sobre independencia de magistrados y abogados (ONU, 2022) ha enfatizado que cualquier herramienta algorítmica utilizada en procesos judiciales debe cumplir con los principios de explicabilidad, transparencia, auditabilidad y supervisión humana significativa. De lo contrario, se corre el riesgo de socavar no solo el debido proceso, sino también el derecho a ser juzgado de manera justa, conforme a los estándares internacionales de derechos humanos (Corte IDH, Caso Maldonado Ordoñez vs. Guatemala, 2016).

Finalmente, cabe resaltar que la UNESCO, en su Recomendación sobre la Ética de la Inteligencia Artificial (2021), instó a los Estados a realizar evaluaciones de impacto algorítmico antes de implementar sistemas de IA en contextos sensibles, incluyendo la administración de justicia. Estas evaluaciones deben considerar los efectos diferenciales en poblaciones vulnerables y establecer mecanismos efectivos de rendición de cuentas.

Por tanto, en el contexto colombiano, antes de adoptar sistemas predictivos de IA en la justicia penal, resulta imprescindible realizar evaluaciones de impacto algorítmico con perspectiva de derechos humanos, asegurar la participación activa de las partes procesales en la validación de datos y variables utilizadas, adoptar medidas de corrección

de sesgos sistémicos, garantizar la auditabilidad plena de los algoritmos, y, sobre todo, preservar la deliberación crítica y autónoma del juez como centro de la decisión judicial.

Solo así se evitaría que la diversidad cultural y la pluralidad étnica reconocidas constitucionalmente se vean erosionadas por una tecnología que, en ausencia de controles, podría convertirse en un nuevo instrumento de exclusión.

6.2. Riesgos para la autonomía judicial y la justicia diferencial

La introducción de inteligencia artificial (IA) en el ejercicio de la función jurisdiccional plantea riesgos significativos para la autonomía e independencia del juez, pilares esenciales del Estado social y democrático de derecho que consagra la Constitución Política de Colombia en su artículo 1° y que se concretan en el derecho fundamental al debido proceso (art. 29). La autonomía judicial implica no solo la independencia respecto de presiones externas, sino también la obligación positiva de motivar las decisiones conforme a un razonamiento jurídico propio y verificable.

La Corte Constitucional, en su Sentencia T-323 de 2024, examinó de manera crítica el uso de ChatGPT 3.5 por parte de un juez de segunda instancia en un proceso de tutela. El Tribunal concluyó que esta práctica vulnera múltiples garantías constitucionales, al señalar que:

“El juez incumple sus deberes legales cuando, habiendo utilizado IA en una decisión judicial, (i) la emplea en funciones de raciocinio indelegables; (ii) no informa a las partes sobre su uso, comprometiendo el derecho de contradicción; (iii) omite verificar rigurosamente la fiabilidad de la información, poniendo en riesgo la imparcialidad; y/o (iv) no adopta cautelas para proteger derechos como el habeas data o la intimidad” (Corte Constitucional, Sentencia T-323 de 2024, ap. 298).

Esta posición refuerza la doctrina constitucional previamente trazada en decisiones como la Sentencia SU-159 de 2021, en la cual la Corte advirtió que el juez no puede convertirse en un "validador mecánico" de criterios externos ni abdicar de su deber de razonamiento crítico, y en la Sentencia T-088 de 2018, donde se reiteró que toda providencia debe ser fruto de un ejercicio autónomo, deliberativo y público de motivación (Corte Constitucional, Sentencia SU-159 de 2021; Sentencia T-088 de 2018).

El uso inadecuado de IA podría, en efecto, producir una delegación inconstitucional de la función jurisdiccional en sistemas automatizados que carecen de legitimidad democrática, no son sujetos de responsabilidad jurídica, ni están sometidos al control constitucional. Este riesgo resulta especialmente grave si se tiene en cuenta que, como ha reconocido la propia Corte Constitucional, el proceso judicial no es un trámite mecánico, sino un espacio deliberativo donde se garantiza la realización efectiva de los derechos fundamentales (Corte Constitucional, Sentencia C-037 de 1996).

Desde una perspectiva crítica, la Criminología del Sur ha llamado la atención sobre los peligros del "reciclaje jurídico" de tecnologías desarrolladas en contextos del norte global y trasladadas acríticamente a realidades del sur. Autores como Zaffaroni (2006), De Sousa Santos (2010) y Carvalho y Sozzo (2019) advierten que tales procesos tecnológicos, lejos de generar emancipación, tienden a profundizar las dinámicas de exclusión, colonialismo interno y discriminación estructural presentes en nuestras sociedades.

En ese sentido, la imposición de sistemas predictivos como COMPAS en Colombia —sin adaptación crítica ni control efectivo— podría homogeneizar criterios de riesgo que desconozcan las realidades sociohistóricas, culturales y territoriales de poblaciones vulnerables, en abierta contradicción con el mandato constitucional de

protección diferenciada (Corte Constitucional, Sentencia T-045 de 2010). La justicia penal debe reconocer las condiciones particulares de exclusión y vulnerabilidad de sectores como los pueblos indígenas, comunidades afrodescendientes y poblaciones rurales, so pena de reproducir, bajo la apariencia de neutralidad algorítmica, patrones de discriminación histórica.

Además, la Relatoría Especial de la ONU sobre independencia judicial (Naciones Unidas, 2022) ha subrayado que el uso de IA en decisiones que afectan derechos fundamentales exige una supervisión humana significativa, principios de transparencia, auditabilidad y no discriminación. De lo contrario, las herramientas algorítmicas se convertirían en “fuentes opacas de autoridad”, minando la legitimidad del sistema de justicia y el principio democrático de rendición de cuentas.

Para preservar la legitimidad de la función jurisdiccional, la incorporación de IA en la justicia penal colombiana debe quedar subordinada al raciocinio humano. En este sentido, es imprescindible garantizar que toda decisión sea producto de un juicio autónomo, deliberativo y motivado en derecho; exigir la máxima transparencia y explicabilidad en cualquier herramienta de apoyo utilizada; reconocer y corregir posibles sesgos estructurales que afecten el acceso a la justicia de poblaciones vulnerables; y, sobre todo, reafirmar que el juez es el único titular legítimo del poder de decidir, bajo el amparo de la Constitución y los tratados internacionales de derechos humanos.

En definitiva, cualquier incorporación de IA que desplace o debilite la autonomía judicial y la justicia diferencial no solo sería inconstitucional, sino que pondría en riesgo la esencia misma del Estado de derecho en Colombia.

7. Conclusiones

La inteligencia artificial aplicada a la evaluación del riesgo criminal representa uno de los desafíos más complejos y, a su vez, una de las áreas de mayor debate dentro del Derecho contemporáneo. En particular, la introducción de algoritmos predictivos en la administración de justicia plantea una profunda reflexión acerca de los riesgos y beneficios asociados con su implementación. Este avance tecnológico, que promete agilizar y mejorar la eficiencia en los procesos judiciales, no está exento de tensiones y contradicciones. La posibilidad de introducir tecnologías como COMPAS en la evaluación de los riesgos de los acusados genera, de manera inevitable, una fricción entre la necesidad de optimizar la administración de justicia y la obligación irrenunciable de garantizar la plena protección de los derechos fundamentales. En este sentido, la cuestión no se limita a un debate técnico sobre la eficacia de la inteligencia artificial, sino que plantea interrogantes sobre el modelo de justicia que se desea construir y los principios que deben prevalecer en un Estado social de derecho.

A lo largo de esta monografía, hemos abordado los distintos aspectos técnicos, éticos, y jurídicos relacionados con la implementación de COMPAS, mostrando cómo, aunque este tipo de herramientas prometen mejorar la eficiencia judicial, presentan serias limitaciones que pueden comprometer principios esenciales como la transparencia, la imparcialidad y la equidad. Los algoritmos, lejos de ser instrumentos neutrales, reflejan los sesgos históricos y sociales presentes en los datos con los que se entrenan. Esto puede resultar en una reproducción de las desigualdades estructurales ya existentes en la sociedad, lo cual representa una amenaza directa para el respeto de los derechos procesales de los acusados y para la consecución de una justicia verdadera y justa. En este contexto, resulta esencial una evaluación crítica que no solo examine las ventajas de la

inteligencia artificial, sino que también identifique los riesgos y desafíos que plantea su uso en el ámbito judicial.

En primer lugar, el análisis del caso *State v. Loomis* proporcionó una visión clara de los peligros asociados con la utilización de sistemas opacos y no auditables en el proceso judicial. La Corte Suprema de Wisconsin, al introducir COMPAS en su sistema judicial, permitió que un algoritmo determinara la evaluación de riesgo de los acusados, pero en un entorno donde dicho sistema no era completamente comprensible ni accesible para los jueces ni para los abogados. A pesar de las advertencias sobre el carácter "complementario" de la herramienta, la práctica demostró que la influencia del algoritmo fue considerablemente más allá de lo que se pretendía. Los resultados generados por COMPAS fueron utilizados para justificar decisiones que afectaron de manera directa la libertad y la condena de los acusados, lo que levantó preocupaciones serias sobre la imparcialidad del sistema judicial y la posible perpetuación de sesgos raciales. La situación descrita en el caso estadounidense es especialmente pertinente para Colombia, ya que nuestro país enfrenta una realidad de desigualdad estructural mucho más profunda, con altos índices de discriminación y exclusión social que se reflejan en los sistemas judiciales. La aplicación de tecnologías como COMPAS en este contexto podría resultar en un reforzamiento de los patrones de exclusión social y de discriminación racial, lo que sería incompatible con los principios constitucionales de igualdad y no discriminación.

Asimismo, el análisis de la Sentencia T-323 de 2024 de la Corte Constitucional dejó en claro que la administración de justicia en Colombia debe ejercerse de forma autónoma, deliberativa y pública. La Corte, al abordar la relación entre la inteligencia artificial y los derechos fundamentales, subrayó que la incorporación de nuevas tecnologías en el sistema judicial no puede ocurrir sin garantizar ciertos estándares

mínimos que aseguren la protección de los derechos de las personas. Entre estos estándares se incluyen la explicabilidad, la transparencia, la supervisión humana significativa y, especialmente, la protección de los datos personales, que en el contexto actual es un derecho clave de los ciudadanos. Estos principios no son opcionales ni meramente deseables, sino que se erigen como mandatos constitucionales que limitan el uso legítimo de tecnologías en la administración de justicia. En este sentido, las herramientas tecnológicas como COMPAS deben ser objeto de un escrutinio riguroso, con el fin de garantizar que no vulneren derechos fundamentales ni atenten contra la legitimidad del proceso judicial. La transparencia y la supervisión humana son esenciales para asegurar que las decisiones judiciales sigan siendo el resultado de un proceso deliberativo, basado en la reflexión crítica y no en la aplicación mecánica de algoritmos.

Por otro lado, el análisis ético y cultural que se llevó a cabo en el capítulo cuarto de esta monografía permitió identificar que el uso de sistemas como COMPAS en contextos como el colombiano no solo implica un riesgo técnico, sino también un riesgo de tipo epistemológico y político. La adopción irreflexiva de estas tecnologías, diseñadas en su mayoría en el norte global, sin un adecuado enfoque local, contribuye a reforzar patrones coloniales de criminalización que descontextualizan las realidades socioculturales y económicas de las poblaciones afectadas. La criminología del sur ha señalado que la aplicación de modelos tecnológicos desarrollados en otros contextos puede resultar en la perpetuación de injusticias, invisibilizando las particularidades locales de los grupos sociales más vulnerables. En Colombia, la marginalidad y la desigualdad social no son fenómenos aislados ni accidentales, sino que forman parte de la estructura misma del Estado, por lo que la implementación de cualquier sistema tecnológico debe ser cuidadosamente adaptada a esta realidad. En ausencia de una

adaptación contextual adecuada, tecnologías como COMPAS no solo corren el riesgo de ser ineficaces, sino que también pueden acentuar las brechas de exclusión y discriminación que ya existen en el país.

En virtud de lo anterior, esta investigación concluye que la implementación de sistemas como COMPAS en el sistema penal colombiano, en su forma actual, no es jurídicamente viable ni éticamente defendible. La opacidad inherente a estos sistemas, su falta de flexibilidad para adaptarse a la diversidad sociocultural del país y su potencial para afectar gravemente las garantías procesales fundamentales hacen inviable su incorporación en el sistema judicial colombiano sin una profunda transformación de sus fundamentos técnicos y operativos. No basta con modificar superficialmente los sistemas tecnológicos; es necesario replantear sus bases conceptuales y operativas para que puedan ser utilizados de manera responsable y respetuosa con los derechos humanos. Mientras estos sistemas no cuenten con la capacidad de ser completamente auditados, explicados y contextualizados, cualquier intento de integrarlos en el sistema de justicia penal debe ser rechazado, ya que sus riesgos superan ampliamente cualquier beneficio potencial.

En definitiva, el debate sobre la inteligencia artificial en la justicia penal no debe ser reducido a una mera cuestión de modernización tecnológica. Es, ante todo, un debate fundamental sobre el modelo de justicia que deseamos construir para el futuro. Este debate no solo tiene implicaciones técnicas, sino también filosóficas y políticas, pues cuestiona si queremos una justicia basada en la deliberación humana, el respeto a la dignidad de la persona y el control democrático, o si preferimos un sistema donde las decisiones más trascendentales sean delegadas a algoritmos sin conciencia ni sensibilidad, operando bajo lógicas estadísticas ajenas a los principios de justicia. En este contexto, la tecnología debe ser entendida como un medio para lograr una justicia más

eficaz, pero siempre subordinada a los valores que fundamentan nuestra democracia y nuestro orden constitucional.

Como futura abogada, considero que la innovación tecnológica debe ser bienvenida únicamente cuando sirva a la realización plena de los principios constitucionales y a la protección efectiva de los derechos humanos. La tecnología debe ser un instrumento al servicio del bien común, y no puede prevalecer sobre los valores y principios que definen nuestra sociedad democrática. En última instancia, la justicia no debe adaptarse a la tecnología, sino que la tecnología debe estar al servicio de la justicia, garantizando siempre el respeto a los derechos humanos, la dignidad de las personas y el Estado de derecho.

En conclusión, mientras no se logre establecer un marco normativo y tecnológico adecuado que garantice la existencia de algoritmos plenamente auditables, explicables y contextualizados, toda pretensión de introducir herramientas predictivas en la justicia penal debe ser rechazada. El verdadero progreso no consiste en acelerar la toma de decisiones, sino en profundizar en la legitimidad de las decisiones judiciales, asegurando siempre su respeto irrestricto por los derechos fundamentales y las garantías procesales. Solo a través de un compromiso ético y responsable con la protección de los derechos humanos y mediante un análisis crítico y contextualizado de las realidades sociales de Colombia, la inteligencia artificial podrá convertirse en una herramienta valiosa y responsable para la construcción de una justicia más humana, transparente y democrática.

8. Bibliografía

- Agencia Española de Protección de Datos. (2020). La inteligencia artificial en el sector público. <https://www.aepd.es/sites/default/files/2020-07/ai-sector-publico.pdf>
- Alto Comisionado de la ONU. (2021). *El impacto de la IA en los derechos humanos: un enfoque global y local*. Informe sobre el uso de IA en el sector público.
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Avella, M. D., Sanabria-Moyano, J. E., & Dinas-Hurtado, K. (2022). Uso del algoritmo COMPAS en el proceso penal y los riesgos a los derechos humanos. *Revista Brasileira de Direito Processual Penal*, 8(1), 275–310. <https://doi.org/10.22197/rbdpp.v8i1.615>
- Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, 104(3), 671–732. <https://doi.org/10.2139/ssrn.2477899>
- Brennan, T., Dieterich, W., & Ehret, B. (2009). Evaluating the predictive validity of the COMPAS Risk and Needs Assessment System. *Criminal Justice and Behavior*, 36(1), 21–40. <https://doi.org/10.1177/0093854808326545>
- Carvalho, S., & Sozzo, M. (2019). Criminologías del sur: Epistemologías insurgentes y críticas situadas. *Cuadernos de Criminología*, 13(1), 12–31.
- Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data*, 5(2), 153–163. <https://doi.org/10.1089/big.2016.0047>
- Consejo de Europa. (2020). *Recomendación sobre el uso de la inteligencia artificial en la justicia*.

Convención Americana sobre Derechos Humanos. (1969). Pacto de San José.
Organización de los Estados Americanos.

Corte Constitucional de Colombia. (2001). Sentencia C-252 de 2001. M.P. Jaime
Córdoba Triviño.

Corte Constitucional de Colombia. (2002). Sentencia C-373 de 2002. M.P. Rodrigo
Escobar Gil.

Corte Constitucional de Colombia. (2003). Sentencia T-949 de 2003. M.P. Rodrigo
Escobar Gil.

Corte Constitucional de Colombia. (2013). Sentencia T-119 de 2013. M.P. Jorge Iván
Palacio Palacio.

Corte Constitucional de Colombia. (2016). Sentencia T-622 de 2016. M.P. Antonio José
Lizarazo Ocampo.

Corte Constitucional de Colombia. (2017). Sentencia T-063 de 2017. M.P. Luis
Guillermo Guerrero Pérez.

Corte Constitucional de Colombia. (2018). Sentencia T-088 de 2018. M.P. Diana Fajardo
Rivera.

Corte Constitucional de Colombia. (2021). Sentencia SU-159 de 2021. M.P. Cristina
Pardo Schlesinger.

Corte Constitucional de Colombia. (2022). Sentencia T-033 de 2022. M.P. Natalia Ángel
Cabo.

Corte Constitucional de Colombia. (2024). Sentencia T-323 de 2024. M.P. Juan Carlos
Cortés González.

Corte Interamericana de Derechos Humanos. (2016). Caso Maldonado Ordoñez vs.
Guatemala. Sentencia de 3 de mayo de 2016. Serie C No. 311.

- De Sousa Santos, B. (2010). *Descolonizar el saber, reinventar el poder*. Siglo del Hombre Editores.
- Departamento Administrativo Nacional de Estadística (DANE). (2023). Informe de pobreza multidimensional 2023. <https://www.dane.gov.co>
- Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.
- García, V., & Sherman, L. (2018). Risk assessment and the state of criminal justice reform. *Harvard Law Review*, 131(6), 1146–1167.
- Gómez, J. C., & Rojas, M. A. (2019). El sesgo algorítmico en la justicia penal predictiva: una advertencia para Colombia. *Revista CES Derecho*, 10(2), 134–157. <https://doi.org/10.21615/cesder.10.2.6>
- Harcourt, B. E. (2015). Risk as a proxy for race. *Federal Sentencing Reporter*, 27(4), 237–243. <https://doi.org/10.1525/fsr.2015.27.4.237>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Ley 1564 de 2012. Por la cual se expide el Código General del Proceso y se dictan otras disposiciones. Diario Oficial No. 48.489, 12 de julio de 2012.
- Ley 1581 de 2012. Por la cual se dictan disposiciones generales para la protección de datos personales. Diario Oficial No. 48.587, 18 de octubre de 2012.
- López, F. (2019). *La opacidad algorítmica en el proceso penal y sus implicaciones para los derechos fundamentales*. *Revista de Derecho Penal*, 34(1), 45-67.
- Naciones Unidas. (2022). Informe de la Relatoría Especial sobre la independencia de magistrados y abogados (A/HRC/50/36). <https://digitallibrary.un.org/record/3972916>
- OpenAI. (2023). About GPT models. <https://platform.openai.com/docs/models/gpt-3-5>

- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). (2021). Recomendación sobre la ética de la inteligencia artificial. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- Procuraduría General de la Nación. (2023). Inteligencia artificial y justicia en Colombia: Recomendaciones preliminares. <https://www.procuraduria.gov.co>
- Russell, S., & Norvig, P. (2020). Inteligencia artificial: Un enfoque moderno (4.ª ed.). Pearson.
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3), 210-229.
- State v. Loomis, 881 N.W.2d 749 (Wisconsin Supreme Court, 2016).
- UNESCO. (2021). *Artificial Intelligence and the Future of Education*.
- Vallejos, D. (2020). Inteligencia artificial en la justicia penal: riesgos y oportunidades para América Latina. *Revista Iberoamericana de Derecho Procesal Constitucional*, 24(1), 25–45.
- Wexler, R. (2018). Life, liberty, and trade secrets: Intellectual property in the criminal justice system. *Stanford Law Review*, 70(5), 1343–1429.
- Zaffaroni, E. R. (2006). La cuestión criminal. Editorial Planeta.