

# **Inteligencia Artificial en el Sistema Jurídico Colombiano: Innovación, Desafíos Éticos y Protección de Derechos fundamentales**

**Juan Manuel Vélez Carvajal**

**Monografía para optar para el título de abogado**



**Universidad Autónoma Latinoamericana (UNAULA)**

**Facultad de Derecho**

**Medellín, Antioquia, Colombia**

**2025**

## ***RESUMEN***

La presente monografía analiza el impacto de la inteligencia artificial (IA) en el sistema jurídico colombiano, abordando sus implicaciones éticas, normativas y sociales. Se parte de una definición funcional de la IA y se identifican sus principales aplicaciones en la administración de justicia, incluyendo sistemas de apoyo a decisiones judiciales y automatización de trámites. Se examinan los riesgos que representa para los derechos fundamentales, cuando los sistemas toman decisiones injustas y los humanos dejamos de tener la última palabra. A nivel normativo, se constata un vacío legal que impide regular eficazmente el uso de estas tecnologías, aunque se resaltan avances como el Proyecto de Ley Estatutaria 154 de 2024 y otros esfuerzos legislativos recientes. La investigación adopta un enfoque documental, basado en el análisis de doctrina, jurisprudencia, proyectos normativos y estándares internacionales. Finalmente, se proponen sugerencias para crear normas claras y justas, que garanticen nuestros derechos y, al mismo tiempo, permitan que las nuevas tecnologías sigan avanzando.

**Palabras clave:** inteligencia artificial, derecho colombiano, derechos fundamentales, ética digital, regulación jurídica, sesgo algorítmico.

## ***ABSTRACT***

This monograph analyzes the impact of artificial intelligence (AI) on the Colombian legal system, addressing its ethical, regulatory, and social implications. It begins with a functional definition of AI and identifies its main applications in the administration of justice, including decision-support systems and the automation of legal procedures. The risks to fundamental rights are examined, particularly when systems make unfair decisions and humans no longer retain the final say. At the regulatory level, a legal gap is identified that hinders the effective governance of these technologies, although developments such as Statutory Bill 154 of 2024 and other recent legislative efforts are highlighted. The research adopts a documentary approach, based on the analysis of legal doctrine, case law, legislative proposals, and international standards. Finally, the study offers recommendations for creating clear and fair rules that safeguard fundamental rights while allowing emerging technologies to continue evolving.

**Keywords:** Artificial intelligence; Colombian law; fundamental rights; digital ethics; legal regulation; algorithmic bias.

## **Contenido**

Inteligencia Artificial en el Sistema Jurídico Colombiano: Innovación, Desafíos Éticos y Protección de Derechos fundamentales.....	1
.....	1
Introducción.....	4
Capítulo 1. Fundamentos de la inteligencia artificial y su vínculo con el derecho.....	5
1.1 Concepto de inteligencia artificial.....	5
1.2. Tipos y clasificación de la inteligencia artificial.....	6
1.3. Innovaciones jurídicas impulsadas por la inteligencia artificial.....	8
Capítulo 2. Desafíos éticos y jurídicos en la implementación de sistemas inteligentes dentro del ámbito jurídico.....	11
2.1. Discriminación algorítmica y sesgos.....	11
2.2. Riesgos para los derechos fundamentales.....	12
2.3. Transparencia y explicabilidad.....	18
2.4. Responsabilidad humana y control institucional.....	19
Capítulo 3. Marco normativo y político en Colombia.....	21
3.1. Diagnóstico legal actual y vacío regulatorio.....	21
3.2. Análisis del Proyecto de Ley Estatutaria sobre IA (2024).....	25
3.3 Aplicaciones tecnológicas basadas en inteligencia artificial en la justicia colombiana.....	32
Capítulo 4. Propuestas para una regulación ética y funcional.....	35
4.1. Estándares internacionales como fundamento orientador.....	35
4.2. El Estado de Derecho como Marco Regulatorio Democrático.....	40
4.3 Recomendaciones para una política pública en Colombia.....	43
5. conclusion.....	48
Referencias.....	50

## Introducción

Este estudio analiza cómo la inteligencia artificial viene desafiando las reglas del derecho colombiano. Se indaga cómo herramientas automáticas están reescribiendo, a pasos acelerados, la forma en que se investigan casos, se expide fallos que protegen derechos fundamentales. Un diagnóstico recurrente en los capítulos es el de un vacío normativo que impide cosechar los beneficios de la tecnología y el respeto de los principios cardinales como la igualdad, la privacidad y la dignidad humana.

La pregunta principal de este estudio es ¿cómo puede Colombia crear reglas claras y justas que, sin frenar el avance de la tecnología inteligente, que protejan los derechos fundamentales de la igualdad, la privacidad y la dignidad humana?. La inquietud no surge en el vacío: los algoritmos han dejado de ser una curiosidad técnica para operar ya en salas de justicia, en los organismos y entidades del Estado. Esa misma omnipresencia torna urgente evitar la discriminación automatizada, exigir transparencia en los razonamientos que las máquinas publican y, no menos importante, custodiar los datos personales que cada proceso genera.

La monografía tiene como objetivo diseñar, de forma gradual, un marco jurídico para el uso de la inteligencia artificial en Colombia. Estas reglas buscan apoyarse en los valores que ya reconoce la Constitución y en las guías internacionales que el país ha venido aceptando.

Para dar cuerpo a ese propósito general se han organizado tareas más concretas: primero se expone los fundamentos conceptuales de la inteligencia artificial y su relación con el derecho, abordando sus definiciones, clasificaciones y el impacto de las innovaciones jurídicas que impulsa. Segundo, examinar los principales desafíos éticos y jurídicos que surgen al implementar sistemas inteligentes en el ámbito legal, poniendo especial atención a los riesgos para los derechos fundamentales, la discriminación algorítmica, la transparencia y la responsabilidad institucional. Tercero, se analiza el marco normativo y político colombiano, desde el diagnóstico del vacío legal en que se encuentra, las propuestas normativas en curso, así como las implicaciones de las nuevas tecnológicas en la justicia. Finalmente, el cuarto capítulo presenta propuestas para una regulación ética y funcional, considerando estándares internacionales, el papel del Estado de Derecho y recomendaciones para la formulación de una política pública adaptada a la realidad nacional.

Estas tareas parciales buscan alimentar el debate académico y la construcción de políticas que permitan aprovechar las ventajas de la inteligencia artificial, sin que eso implique ceder en la protección de los derechos fundamentales. El proyecto, en definitiva, espera servir de brújula para una innovación responsable.

El soporte del marco teórico son los derechos fundamentales, poniendo bajo la lupa la dignidad humana que enuncia la Constitución Política de Colombia. Estos valores se proyectan hacia el entorno digital en el que operan hoy las plataformas automatizadas.

El presente estudio adopta un enfoque de investigación cualitativo documental, concepto que, en su forma más estricta, se refiere a la evaluación sistemática de todo aquel material impreso, digital o audiovisual que aborde el efecto y la normativa que la inteligencia artificial ha traído consigo al derecho colombiano. Esa información primaria exige, en primer lugar, la identificación y posterior selección de fuentes pertinentes, entre las que se

cuentan monografías, artículos de revistas arbitradas, informes de instituciones oficiales, textos legislativos y la jurisprudencia más importante de las altas cortes del país.

## **Capítulo 1. Fundamentos de la inteligencia artificial y su vínculo con el derecho**

### **1.1 Concepto de inteligencia artificial.**

La inteligencia artificial (IA) ha transformado de manera notable diversos aspectos de la vida contemporánea en el siglo XXI. Este concepto, en constante evolución, puede entenderse como el conjunto de algoritmos y sistemas capaces de reproducir, de formas a veces inesperadas, habilidades que antes eran exclusivas del ser humano, tales como la toma de decisiones o el aprendizaje a partir del análisis de datos y ejemplos. Según el *Proyecto de Ley Estatutaria sobre Inteligencia Artificial en Colombia* (Cámara de Representantes de Colombia, 2024), la IA se define como:

Conjunto de técnicas informáticas, sistema de programación, sistema computacional, máquinas físicas o procesos tecnológicos que permiten desarrollar algoritmos, toma decisiones y crear programas informáticos para ejecutar objetivos definidos por humanos, hacer predicciones, recomendaciones, tomar decisiones crear nuevo conocimiento y/o completar tareas cognitivas y científico-técnicas a partir de la extracción, selección, recorte y organización de la información disponible o cualquier tarea que requiera inteligencia como el razonamiento o el aprendizaje. (art. 3, inc. 5)

El planteamiento normativo que ahora se discute opera sobre la base de la funcionalidad pura: la inteligencia artificial deja de ser un mero reflejo de la actividad humana y se convierte en un socio capaz de procesar vastos torrentes de datos a ritmos y extensiones que el juicio de cualquier persona no podría igualar. A partir de esa ventaja, el sistema elabora conclusiones sin necesitar supervisión directa, actuando, de cierto modo, como un operador autónomo.

Cuando se habla de inteligencia artificial, suelen aparecer términos como *machine learning*, *deep learning* o los más tradicionales —aunque aún vigentes— sistemas expertos. Todos ellos comparten la idea de que un programa puede realizar tareas que recuerdan al pensamiento humano, ya sea aplicadas a un juicio o a la interpretación de un contrato. Algunos motores operan siguiendo reglas predefinidas, semejantes a un manual jurídico; otros se apoyan en la experiencia de casos previos para decidir con base en patrones reconocidos; y los más avanzados intentan reproducir, paso a paso, el proceso mediante el cual una persona forma una opinión. Cada enfoque posee sus ventajas y aplicaciones, aunque persiste el debate sobre cuál es el más adecuado en cada contexto.

A nivel internacional, la UNESCO (2021) decidió meter mano al tema y por eso hizo su famosa Recomendación sobre la Ética de la Inteligencia Artificial. En el texto aclara que, cuando hablamos de IA, de verdad no sólo miramos líneas de código, sino en gadgets y plataformas que aprenden por su cuenta, se adaptan y, a veces, eligen sin que nadie les diga qué hacer. Esa flexibilidad deja el camino abierto a montones de problemas éticos, legales y sociales que, advierte la organización, hay que poner bajo control. También, al darle ese giro tan amplio al concepto, la UNESCO empuja a pensar en cómo todo eso puede tocar los derechos humanos de cada persona.

Por ahora, nadie ha conseguido una sola definición de inteligencia artificial que todos aceptan. Lo que llama IA puede verse como un conjunto de programas, una máquina que toma decisiones o algo que todavía suena de ciencia-ficción, dependiendo de a quién se le pregunte. En el contexto del derecho colombiano, los juristas se están moviendo hacia una idea que mezcla el músculo técnico de esas herramientas con frenos normativos anclados en los derechos fundamentales.

## 1.2. Tipos y clasificación de la inteligencia artificial

La charla sobre inteligencia artificial se puede dividir de varias maneras: cuánto terreno actúa por sí sola, qué papel le asignamos y cuánto se parece de verdad a la mente humana. Entre esos modos de mirar son famosos el apodo de IA débil y su hermano mayor, la IA fuerte.

Cuando hablamos de IA débil nos referimos a esas aplicaciones que hacen lo que se les indica, pero no saben que lo están haciendo. Piensa en tu asistente de voz, el software que ordena música o el servicio que sugiere películas. La IA fuerte, en cambio, es el gran sueño: una máquina que razona en cualquier contexto y elige como lo haríamos nosotros en situaciones complejas.

La OCDE, la Unión Europea y varios grupos de expertos a menudo dividen la IA según cuánto peligro puede traer: riesgo mínimo, riesgo limitado, alto riesgo y riesgo inaceptable. Bajo esa misma guía, cosas como el sistema de puntuación social o el programa que pronostica delitos sin ninguna revisión humana caen en la categoría más alarmante.

La forma de clasificación más relevante es la que distingue la IA según su **estructura técnica**. En este sentido, se puede dividir en:

**La IA basada en reglas (también llamado sistemas expertos):** operan siguiendo una serie de instrucciones y condiciones que han sido definidas con anterioridad. Este enfoque les permite analizar grandes volúmenes de información, reconocer patrones y tomar decisiones sin añadirle un conjunto de datos para entrenar. En contextos donde la consistencia y la exactitud son requisitos innegociables, estos modelos suelen resultar más confiables que las variantes basadas en el aprendizaje automático.

El funcionamiento de un sistema basado en reglas puede describirse como un ciclo continuo de aplicación, verificación y ajuste. Primero, se emplean las reglas existentes para procesar la entrada. Luego, los resultados son comparados con los criterios establecidos y, si es necesario, se introducen nuevas reglas que se adaptan a la situación actual. En ocasiones, el razonamiento se dirige mediante inferencia lógica. Cuando parte de un conjunto de evidencia inicial, avanza hacia una conclusión posible; si, en cambio, parte de una hipótesis, retrocede buscando datos que la respalden.

Pondremos un ejemplo bastante simple para entender la lógica detrás de los Sistemas Expertos: Messi es el mejor. Si Messi juega, entonces Barcelona FC gana. Esta sería una regla muy simple. La cuestión clave es que debe existir alguien que defina estas reglas. Si existe un cambio, el Sistema Experto debe ser cambiado, lo cual era costoso en esa época. El problema con los Sistemas Expertos era que no tenían autonomía al momento de resolver el problema, pues dependían de las reglas impuestas. En términos sencillos, el Sistema Experto era como aquel estudiante que

memorizaba todo un curso y aprobaba el examen, pero realmente no entendía nada acerca de la materia. (Morales Cáceres, 2021, p. 47)

**La IA basada en aprendizaje automático (*machine learning*)**, representa una transformación clave dentro del panorama tecnológico contemporáneo. A diferencia de los sistemas tradicionales que funcionan mediante un conjunto rígido de reglas, los modelos de aprendizaje automático tienen la capacidad de adquirir nuevos conocimientos, adaptarse a circunstancias cambiantes y autoperfeccionarse simplemente a partir de la información que procesan. Esto elimina, en buena medida, la necesidad de que un programador defina cada escenario posible por adelantado. Desde esta perspectiva, el aprendizaje automático se erige como un subcampo autónomo de la informática y, a su vez, como una de las ramas más dinámicas de la inteligencia artificial en sentido amplio.

La idea central que guía estos desarrollos es que una máquina puede “aprender” cuando su comportamiento, o su rendimiento en tareas concretas, muestra una tendencia al crecimiento o la corrección tras haber sido expuesta a nuevos datos. Este principio establece un claro contraste con los modelos basados en reglas, que operan sin variación a menos que un humano modifique el código fuente. En esencia, el aprendizaje automático permite que las computadoras evolucionen, y esa promesa es lo que ha capturado el interés de investigadores, empresas y sociedad en general.

El *Machine Learning* funciona de la siguiente manera: imaginemos que una compañía de telecomunicaciones quiere saber qué clientes están a punto de darse de baja de sus servicios a fin de crear campañas de marketing para retenerlos. ¿Cómo podría hacerlo? La empresa tiene muchos datos de los clientes, tales como la antigüedad, planes contratados, consumo diario, llamadas mensuales, reclamos, últimos cambios de planes contratados. En el pasado esta información, al no poder ser procesada adecuadamente, solo sería utilizada para facturar y para hacer estadísticas. Sin embargo, con el *Machine Learning*, estos datos se pueden usar para predecir cuándo un cliente se va a dar de baja y gestionar la mejor acción que lo evite. Esta rama de la IA detecta patrones de comportamiento contando y en base a ello predice futuros comportamientos. Por tanto, la empresa podrá descubrir cuáles son las causas que han llevado, en este caso, a darse de baja como cliente. (Morales Cáceres, 2021, p. 48)

**La IA basada en redes neuronales profundas (*deep learning*)**, ha representado, en la última década, el paso más significativo en la evolución de los sistemas inteligentes contemporáneos. Esta innovación ha transformado áreas tan diversas como el reconocimiento de imágenes, el lenguaje que usan los programas para comunicarse con personas y la toma de decisiones en tiempo real, reconfigurando así los paradigmas sobre cómo las máquinas aprenden y gestionan la información.

El aprendizaje profundo es una rama del aprendizaje automático que utiliza estructuras de redes con múltiples capas —conocidas comúnmente como “profundas”— para imitar las complejas conexiones de las neuronas en el cerebro humano. En el fondo, es un esfuerzo por replicar con métodos computacionales los procesos mentales humanos, muchos de los cuales aún no han sido completamente comprendidos.

Una de las ventajas más notables de este enfoque es su facultad para aprender directamente de los datos, sin depender de la mano de obra previa de un experto que debe definir manualmente qué características son relevantes. En contraste con los métodos de aprendizaje automático más tradicionales, donde la ingeniería de características ocupa un papel central, el aprendizaje profundo extrae automáticamente lo útil de cada conjunto de datos, dejando que la computadora descubra patrones por sí misma a medida que avanza en iteraciones sucesivas.

Por ejemplo, imaginemos a un niño cuya primera palabra es “perro”. ¿Cómo hace el niño para diferenciar un perro de otros animales? Comúnmente, el niño señalará a aquellos animales que considere de esas características y dirá la palabra “perro”.

Entonces, los padres le dirán ya sea: “sí, eso es un perro” o “no, eso no es un perro”. Mientras el niño continúa apuntando a los objetos, se vuelve más consciente de las características que poseen todos los perros. Lo que el niño hace, sin saberlo, es aclarar una abstracción compleja —el concepto de perro— construyendo una jerarquía en la que cada nivel de abstracción se crea con el conocimiento que se obtuvo de la capa precedente de la jerarquía. (Morales Cáceres, 2021, p. 48)

Esta clasificación técnico-funcional también tiene implicaciones jurídicas, especialmente en cuanto a **aplicabilidad y transparencia**. Los sistemas que usan reglas claras son más fáciles de revisar o auditar que los de aprendizaje profundo, porque estos últimos funcionan como una especie de “caja negra”, donde es difícil entender cómo llegan a sus decisiones.

### **1.3. Innovaciones jurídicas impulsadas por la inteligencia artificial**

La inteligencia artificial ya colabora de forma directa con jueces y abogados, dejando atrás su antigua imagen de ciencia ficción para convertirse en una herramienta integrada en las dinámicas cotidianas del ámbito jurídico. Su función principal consiste en procesar grandes volúmenes de información, depurarla y presentarla de forma ordenada, facilitando así una toma de decisiones más ágil. Las aplicaciones más comunes en este contexto suelen agruparse en tres categorías: sistemas que localizan y organizan normas jurídicas, herramientas que automatizan la emisión de decisiones repetitivas y plataformas que gestionan el flujo de documentos y expedientes dentro de los tribunales.

John Zeleznikow (Cáceres, 2006) identifica múltiples líneas de trabajo en las que la inteligencia artificial comienza a integrarse progresivamente en el ámbito jurídico. Entre estas se encuentran la codificación automatizada de normas para facilitar su actualización y la inferencia de consecuencias legales; el tratamiento flexible de conceptos jurídicos de textura abierta; la construcción de ontologías que ordenan el vocabulario legal disponible en la Web y optimizan la recuperación de información; y el desarrollo de sistemas de asesoría jurídica capaces de prever, por ejemplo, la distribución de bienes tras un divorcio.

En el contexto colombiano, la Corte Constitucional presentó recientemente **PROMETEA**, una herramienta de inteligencia artificial desarrollada en conjunto con la Universidad del Rosario, el Laboratorio de IA de la Universidad de Buenos Aires y otras entidades, con el propósito de apoyar la selección de tutelas en la Sala de Revisión. Este sistema surge en un contexto de creciente congestión judicial, agravada por el éxito de la acción de tutela como mecanismo de acceso a la justicia y la ineficiencia estructural de la justicia ordinaria.

PROMETEA representa una apuesta por el uso de tecnologías de la información para mejorar la eficiencia, rapidez y cobertura del sistema judicial sin necesidad de reformas legales profundas. Mediante el análisis automático de sentencias, la herramienta identifica casos prioritarios para su estudio, como lo demostró en una prueba donde seleccionó 32 tutelas relevantes entre 2.700 solicitudes en menos de dos minutos, tarea que a una persona le tomaría meses. Este avance, que podría replicarse en otras áreas judiciales como ocurre ya en Argentina, se enmarca en una visión más amplia de transformación digital, que incluye iniciativas como el expediente digital interoperable y confiable, y que busca cambios estructurales en la administración de justicia y en la percepción ciudadana sobre su acceso y legitimidad. (Blog de Derecho de los Negocios – Universidad Externado de Colombia, 2019)

Asimismo, se ha venido desarrollando el aplicativo **TYBA**, que permite el manejo digital de expedientes. Su idea es que la gente normal, no solo los abogados, revise su propio proceso con un par de clics. Cualquiera que esté metido en el caso (pensando en exigentes letrados, apoderados y hasta el mismo demandante) puede chequear el estado sin fila ni papeles.

Varios países europeos han volcado buena parte de su justicia civil a Internet y la rutina del papel se siente cada vez más ajena. Inglaterra, Finlandia o Portugal en su propio andar han puesto en marcha portales donde el usuario ni siquiera pisa un juzgado para presentar una reclamación, mucho menos para monitorizarla. El County Court Bulk Center británico procesa por arte de magia y casi sin intervención humana los pleitos por deudas monetarias que le llegan en bloques, y la aplicación Money Claim On-Line deja al demandante manejar todo su expediente desde la comodidad de una pantalla, ceder derechos, enmendar errores o aceptar la sentencia sin moverse de casa.

En el ecosistema jurídico español emergen herramientas bastantes llamativas. Esta **Jurimetría**, que combina aprendizaje automático con procesamiento de lenguaje natural para estudiar millones de sentencias. El resultado inmediato son estadísticas predictivas que ayudan a abogados y magistrados a calibrar posibilidades antes de que el fallo ocurra.

Básicamente, el programa ayuda a diseñar la mejor línea argumental y las posibilidades de éxito en relación a un caso, expuesto en sus notas esenciales, teniendo en cuenta el contexto que sobre resoluciones de casos similares anteriores ofrece un tribunal, un juez o magistrado y una hipotética parte contraria (abogado). También ofrece una visión panorámica clara de la estadística judicial de todos los tribunales y de todas las jurisdicciones a nivel nacional. De este modo, con la combinación de los elementos citados, cabe construir una estrategia procesal orientada por el logro del éxito del caso, con ayuda de indicadores gráficos interactivos que proporcionan un uso amigable e intuitivo de la herramienta. (Galindo Ayuda, 2019, p. 40)

**Tyrant Analytics.** Esta plataforma reduce la complejidad de localizar jurisprudencia relevante al presentar visualizaciones que ponen de relieve los patrones y frecuencias de los fallos, de modo que los abogados pueden formular estrategias jurídicas informadas por datos palpables.

El programa es un buscador de jurisprudencia y legislación dotado de mejoras tanto visuales como funcionales que mejoran el acceso que proporcionan las clásicas bases de datos de jurisprudencia y legislación a las que se accede mediante el uso de operadores lógicos por palabras o conceptos, y que normalmente proporcionan en forma desordenada como resultado un conjunto de sentencias o normas que los contienen. De este modo, la funcionalidad principal de la base de datos no se ve difuminada por otras informaciones que pudieran restarle protagonismo al buscador. Mediante ayudas visuales y otras opciones esa búsqueda básica se ve potenciada para ofrecer una visión más amplia al usuario de la herramienta. Cuestiones como el sentido del fallo ante una determinada búsqueda, ponentes, órganos emisores..., se muestran desde el primer momento de una manera gráfica, de modo que la navegación por parte del profesional que utiliza la herramienta, hasta los documentos que necesita, se hace de forma dinámica y visual. (Galindo, 2019, p. 41)

Se destaca así mismo **vLex Analytics**, una herramienta impulsada por inteligencia artificial que examina fallos, escritos y otros documentos procesales para producir métricas orientadas a la práctica profesional en el contexto español. La aplicación permite sondear los expedientes por órgano jurisdiccional, categoría de recurso o asunto concreto, y despliega datos sobre plazos de resolución, porcentajes de estimación y artículos más citados, incluyendo comparativas que enriquecen el diagnóstico estratégico del abogado.

La inteligencia artificial orientada al ejercicio legal también transforma, de forma cada vez más palpable, el modo en que los abogados gestionan sus tareas cotidianas. **Sof-IA**, desarrollada por la editorial Tirant lo Blanch, se presenta como una solución innovadora que automatiza múltiples alcances del trabajo jurídico. La plataforma, en virtud de sus algoritmos ajustados, ofrece respuestas oportunas y con un margen elevado de confiabilidad. Gracias a esta agilidad, los profesionales del Derecho logran acortar sus plazos y mantienen la consistencia en los resultados que entregan a sus clientes.

(..) Su objetivo principal es ayudar a los abogados con búsquedas normativas y jurídicas. Además, ofrece información clara y siempre actualizada basada en su extenso conocimiento del derecho.

A diferencia de otras herramientas de IA, **Sof-IA** está pensada específicamente para abogados. Por lo tanto, proporciona respuestas precisas a preguntas legales con una base de datos especializada en legislación y jurisprudencia. Esto la convierte en una excelente solución para despachos de cualquier tamaño. (Tirant lo Blanch, 2024, párrs. 3-4)

A día de hoy existen herramientas bastante impresionantes que usan inteligencia artificial dentro del sistema penal. Por ejemplo, en Estados Unidos, un sistema llamado **COMPAS** puede calcular en segundos qué tan probable es que una persona vuelva a cometer un delito cuando salga de prisión. En una línea similar, el software **ALIBI** (países bajos) sugiere narrativas que podrían justificar las acciones del interno. Otra herramienta, conocida como **STEVIE** (Australia), arma relaciones coherentes utilizando fragmentos dispares de datos biográficos y de conducta. Por si esto fuera poco, han ido surgiendo métodos cuantitativos

destinados a medir la confiabilidad de los testimonios que presentan las víctimas y los testigos.

En definitiva, la inteligencia artificial ha conquistado los despachos de jueces al ayudarles a preestablecer sus decisiones. El sistema no reemplaza la autoridad del juez, pero sí le ofrece bosques de providencias basados en fallos anteriores o en la jurisprudencia pertinente. Tal apertura a los algoritmos responde sobre todo a la apretada mora judicial, al torrente de litigios y al déficit crónico de personal en los tribunales.

## **Capítulo 2. Desafíos éticos y jurídicos en la implementación de sistemas inteligentes dentro del ámbito jurídico**

La inteligencia artificial ya se ha transformado en una de las tecnologías más importantes del siglo XXI, pero su avance rápido nos pone a pensar sobre dilemas éticos profundos. Estos problemas no son solo teóricos: afectan de manera real y muy tangible a nuestra sociedad, la economía y los derechos fundamentales de las personas.

### **2.1. Discriminación algorítmica y sesgos.**

El problema ético habitual de la IA es la discriminación algorítmica: decisiones hechas sobre patrones históricamente parciales o grupos de datos poco representativos. Un caso representativo fue el sistema COMPAS en los Estados Unidos, donde las tasas de reincidencia se basaron en factores indirectamente raciales. La falta de transparencia hace que el problema sea aún más significativo. Algunos algoritmos no permiten a las personas entender cómo se obtiene una conclusión debido a la complicación técnica: algunos están basados en el aprendizaje profundo.

El sesgo algorítmico ocurre cuando un sistema informático refleja los valores de los humanos que estuvieron implicados en su codificación y en la recolección de los datos usados para entrenar el algoritmo. La IA es buena para establecer patrones, así como para agilizar procesos y operaciones con volúmenes masivos de información —Big Data—. Sin embargo, el problema es que la IA al nutrirse de la información de hecha o recopilada por seres humanos, puede que reflejen sus sesgos.

(...) A modo de ejemplo, en el año 2015, Jacky Alcine —una mujer afroamericana— cuando miró su fotografía en la aplicación de Google Photos no podía creer que el software de reconocimiento facial la había etiquetado con la palabra “gorila”. Esto sucedió porque el algoritmo no había sido entrenado con suficientes imágenes de personas de piel oscura. En otro caso, a inicios del año 2016, Microsoft lanzó a “Tay”, un chatbot cuyo fin era imitar el comportamiento de una adolescente curiosa y buscaba entablar en las redes sociales una conversación informal y divertida con una audiencia de entre 18 y 24 años, según explicó la compañía en su página web. El proyecto mostraría las promesas y el potencial de las interfaces conversacionales alimentadas por IA. Sin embargo, en menos de 24 horas, el “inocente” Tay a través de tweets, mostraba su empatía hacia Hitler o su apoyo al genocidio al responder a preguntas de los usuarios de las redes sociales son algunos ejemplos, además de insultos raciales y comentarios sexistas y homófobos. También defendió el Holocausto, los campos de concentración o la supremacía blanca, y se mostró contraria al feminismo. (Morales Cáceres, 2021, p. 60)

El sesgo de la tecnología de reconocimiento facial es uno de los peores. Afecta principalmente a las mujeres, los individuos de color, y otras poblaciones raciales. El Instituto Nacional de Estándares y Tecnología de Estados Unidos analizó 189 algoritmos de reconocimiento facial y verificó el hallazgo de que la mayoría – si no todos – tenían sesgo. Es posible que haya múltiples razones para ello; sin embargo, resulta que eliminar el sesgo es casi imposible. Por ejemplo, los algoritmos de reconocimiento facial ciento diez veces más inexactos señalarían que una persona es de descendencia asiática en lugar de caucásica.

Las mujeres negras de dieciocho a treinta años son, según diversas investigaciones, el segmento poblacional que menos acierta en pruebas de reconocimiento facial. Esta disfunción del software no es solo una cifra; arrastra efectos perjudiciales muy concretos. Con frecuencia conduce a perder un avión, sufrir un interrogatorio que se prolonga, quedar inscrito en listas de vigilancia, experimentar un roce incómodo con la policía, ser arrestado sin justificación o, en el peor de los casos, a cosas todavía más graves.

El caso del mentiroso en Detroit, Robert Williams, ilustra dramáticamente estos peligros. Williams fue arrestado después de que la tecnología de reconocimiento facial lo vinculó a la imagen de un sospechoso; una imagen, según Williams, no se parecía a él. Al mostrarle la imagen de vigilancia a los oficiales, Williams les preguntó si ellos creían que todos los hombres negros son iguales.

¿Qué es lo que vuelve sesgado un sistema de inteligencia artificial? Las técnicas de aprendizaje automático y aprendizaje profundo son, paradójicamente, las fuentes de su sesgo parcial. Cuando un modelo de clasificación recibe cientos de millas de imágenes de perros correctamente etiquetadas, aprende a reconocer lo que la mayoría de las personas llamarían un perro, inclusive si nunca ha procesado una imagen de un can de raza poco común. De igual manera, un sistema de reconocimiento de voz se vuelve ágil para transcribir la habla en función de las grabaciones que ha “ingerido” previamente; en ocasiones, escribe lo que escucha con mayor rapidez que un taquígrafo profesional. La regla general es sencilla: más datos de etiquetado generalmente equivalen a un mejor rendimiento. El problema, sin embargo, es que al mismo tiempo el modelo desarrolla huecos y ceguera en áreas donde los datos son escasos, sesgados o simplemente inexistentes. (Morales Cáceres, 2021)

## **2.2. Riesgos para los derechos fundamentales**

Los sistemas de inteligencia artificial son cada vez más capaces de agrupar y clasificar a las personas a través de algoritmos cuya lógica permanece oculta a quienes los usan. Esta falta de transparencia plantea serios interrogantes sobre la dignidad humana, ya que las decisiones automatizadas pueden sentirse distantes y casi mecánicas, mientras que el impacto sobre la vida de las personas es muy directo. A esto se añade la inquietud de que tales tecnologías no actúan en un vacío; en su entrenamiento suelen extraer datos de realidades plagadas de sesgos históricos. De ahí que haya un riesgo evidente de que un sistema aparentemente neutro termine reproduciendo y amplificando prejuicios raciales, de género o socioeconómicos. El resultado es una degradación social que fragiliza los entornos comunitarios y amenaza, en última instancia, derechos fundamentales que se consideran universales. Se analizarán los derechos que, prima facie, resultan más vulnerados o en mayor riesgo ante la implementación de sistemas de inteligencia artificial.

## El derecho a la protección de datos personales

El artículo 15 de la Constitución consagra derechos a la intimidad, buen nombre y la facultad de conocer, actualizar y rectificar informaciones en bancos de datos y archivos de entidades públicas y privadas. Sin embargo, de una lectura literal del texto no se desprende de manera explícita la existencia de un derecho fundamental autónomo al habeas data o a la protección de datos personales.

No obstante, la Corte Constitucional ha sostenido, desde temprana jurisprudencia (por ejemplo, Sentencia T-414 de 1992 y SU-082 de 1995), que del artículo 15 se deriva un derecho autónomo denominado habeas data. Este se entiende como el derecho de toda persona a conocer, actualizar, rectificar y, posteriormente, a controlar la información personal que le concierne, otorgando facultades de autodeterminación informativa.

Más adelante, con el desarrollo jurisprudencial y normativo (Ley Estatutaria 1266 de 2008 sobre habeas data financiero y Ley 1581 de 2012 sobre protección de datos personales), se consolidó la noción de que la protección de datos personales es una expresión del habeas data y, por tanto, un derecho fundamental autónomo en el orden constitucional colombiano.

El título del proyecto de ley –“Por el cual se dictan disposiciones generales para la protección de datos personales”- indica que su objetivo principal es regular de manera general las garantías del **derecho fundamental a la protección de los datos personales**, derecho que en la jurisprudencia constitucional ha sido con frecuencia denominado **derecho al habeas data** y en algunas oportunidades derecho a la autodeterminación informática o informativa. (Sentencia C-748/11, 2011, énfasis añadido)

Este derecho implica la obligación de manejarlo con suma cautela, dado que forma parte esencial del patrimonio moral e informático de cada individuo. En la actualidad, resulta inevitable compartir información personal en múltiples escenarios, tales como al abrir una cuenta bancaria, ingresar a una empresa, afiliarse a una institución médica, matricularse en un centro educativo o al acceder a diversos servicios. En síntesis, la vida contemporánea exige a las personas divulgar datos en una variedad considerable de contextos y relaciones sociales. (Carvajal Martínez, 2010)

Hoy en día los datos personales no son solo información privada, sino que se han convertido en una **mercancía valiosa**. Empresas y gobiernos los usan para tomar decisiones, diseñar productos, dirigir campañas políticas o incluso predecir comportamientos sociales. El problema es que la mayoría de las personas no son conscientes de cuánto se recolecta de ellas ni de cómo circula esa información en los mercados digitales

A eso se suma que muchas veces **el consentimiento es más aparente que real**, porque, aunque aceptemos “políticas de privacidad” al usar redes sociales o aplicaciones, en la práctica no tenemos capacidad de negociación. Todo esto genera un escenario donde el individuo pierde poder frente a grandes corporaciones y Estados, que acumulan cantidades enormes de datos con los que obtienen ventajas económicas y políticas.

En este contexto, el consentimiento informado —tradicional pilar del *habeas data*— resulta insuficiente, pues en la práctica se diluye en cláusulas generales de adhesión que los usuarios aceptan sin posibilidad de negociación real. El reto jurídico y ético radica entonces en establecer **mecanismos efectivos de protección de la autodeterminación informativa**, capaces de equilibrar el interés económico en la explotación de datos con la salvaguarda de la dignidad humana y los derechos fundamentales de privacidad e intimidad.

### **El derecho a la no discriminación**

El gran problema con los algoritmos es que solemos pensar que son “justos” porque funcionan con números y datos. Pero en realidad no son imparciales: detrás de cada algoritmo hay personas que deciden qué información usar y cómo usarla. Eso significa que los prejuicios o errores humanos pueden colarse sin que nadie se dé cuenta.

El riesgo de discriminación derivado del uso de algoritmos y big data se explica por la **falsa percepción de objetividad tecnológica**. Aunque los sistemas se presentan como neutrales, las decisiones sobre la selección de variables, categorías y modelos de análisis provienen de seres humanos, quienes trasladan consciente o inconscientemente sus propios sesgos, valores y limitaciones cognitivas. Esta “carga normativa oculta” puede materializarse en discriminación indirecta contra colectivos vulnerables, incluso cuando no hay intención de excluir.

La sorprendente capacidad de la inteligencia artificial para analizar enormes cantidades de datos al instante ha convertido a la vigilancia moderna en un mecanismo de control social casi sin parangón. Gracias a esta tecnología, los sistemas de reconocimiento facial son capaces de seguir a personas en calles, aeropuertos y plazas, borrando de facto el viejo refugio del anonimato y socavando esa privacidad que muchos consideran un pilar de la vida democrática.

El ejemplo más notable y ampliamente reportado de control masivo se sitúa en China, donde el gobierno ha implementado herramientas de inteligencia artificial dirigidas a la minoría uigur. Informes provenientes de fuentes oficiales indican que el régimen utiliza redes de reconocimiento facial para confeccionar "generar perfiles detallados de movimiento individual" de esta comunidad a través de diversas urbes. Este esquema de vigilancia, centrado en un grupo étnico concreto, constituye la primera instancia documentada en que un estado recurre deliberadamente a la IA para realizar perfilamientos basados en raza.

La Integrated Joint Operations Platform (IJOP) que se ha implementado en Xinjiang es un caso ilustrativo de cómo la inteligencia artificial puede convertirse en un mecanismo de control estatal. El sistema aborda la vigilancia desde un nivel insólitamente granular, almacenando datos que van desde el color del automóvil de un residente hasta su estatura en centímetros. A partir de esta información, la plataforma genera perfiles que asocian comportamientos ordinarios—como no intercambiar visitas con vecinos o recurrir a aplicaciones de mensajería cifrada como WhatsApp—con etiquetas de “sospechoso”.

Los algoritmos de inteligencia artificial que hoy día funcionan dentro de los sistemas de justicia penal no son neutros; más bien, a menudo refuerzan y amplifican desigualdades raciales que ya están muy arraigadas. Por ejemplo, las herramientas de policía predictiva, que analizan datos acumulados a lo largo de años para determinar dónde enviar patrullas,

terminan reproduciendo los mismos sesgos que nacen de décadas de vigilancia excesiva en barrios de color.

Una investigación realizada en 2018 ilustra este riesgo de forma muy concreta: si el sistema se hubiera puesto en práctica en Indianápolis, la comunidad latina habría enfrentado entre un doscientos y un cuatrocientos por ciento más de patrullajes en comparación con la población blanca. Ese exceso de supervisión no sólo afecta la experiencia inmediata de quienes viven en esos barrios, sino que además alimenta un ciclo vicioso de datificación, donde el rastreo constante genera más datos, que a su vez justifican más patrullas. Más preocupante aún es que estos algoritmos pueden erosionar las garantías que la Cuarta Enmienda pretende ofrecer. Al basar las detenciones en lo que se califica como “sospecha razonable” algoritmizada, el sistema puede crear una apariencia de evidencia sólida aun cuando no hay indicios concretos de delito, legitimando intervenciones severas sobre comunidades que ya han sido históricamente vulnerables.

Los sistemas de inteligencia artificial también pueden alterar la práctica del Derecho al introducir sesgos indeseados en procesos que se dicen automatizados. Los algoritmos empleados en decisiones judiciales, aunque prometen rapidez, tienden a amplificar la marginación de determinados colectivos y, con ello, amenazan la igualdad ante la ley que debe guiar todo sistema democrático. Así, su despliegue, si bien puede acortar plazos, obliga a plantearnos si, al ganar eficacia, no estamos renunciando a uno de los pilares más antiguos y valorados de la Justicia: la imparcialidad

En el ámbito sanitario la inteligencia artificial no es menos problemática. Diversos programas, en su intento por clasificar pacientes o predecir complicaciones, han terminado por reproducir y hasta agravar las disparidades raciales que ya sufren las comunidades menos favorecidas. Un informe crucial, publicado en Science en 2019, reveló que un sistema ampliamente usado en hospitales de renombre privilegiaba a pacientes blancos sobre sus homólogos negros, aun cuando estos últimos llegaban con cuadros clínicos más graves. La razón detrás de esta falacia descansaba en el uso del gasto total como medida de “necesidad”; dado que, por cuestión de historias de injusticia, se invierte menos en pacientes negros, el programa erraba al concluir que requerían un menor esfuerzo terapéutico.

Los modelos de inteligencia artificial en el ámbito sanitario a menudo se ajustan con registros históricos que están plenos de desigualdades raciales. Debido a que, a lo largo de los años, pacientes negros han recibido tratamientos inferiores, la IA que se nutre de estos datos termina por replicar esos mismos sesgos. Como resultado, las decisiones automatizadas pueden desembocar en diagnósticos erróneos, terapias inadecuadas o la patente desestimación de las inquietudes de salud de quienes pertenecen a esa comunidad.

De forma similar, las plataformas algorítmicas destinadas a la contratación laboral han demostrado perpetuar la discriminación contra colectivos ya vulnerables. El caso Mobley contra Workday, Inc. pone de relieve cómo estos sistemas pueden, sin que sus creadores lo busquen a propósito, dejar fuera a solicitantes negros, personas mayores y trabajadores con discapacidad. Aunque Workday no actuó con mala intención, el resultado de su IA tuvo un coste desproporcionado para esos grupos.

Por otra parte, los sistemas de gestión que funcionan a partir de algoritmos convierten a los empleados en simples cifras, despojándolos de la autonomía que históricamente existió. Esta tendencia resulta particularmente intensa en la economía colaborativa, donde los operarios son controlados cada minuto por plataformas que alegan objetividad mientras perpetúan patrones discriminatorios.

La irrupción de la automatización basada en inteligencia artificial está transformando el mercado laboral de una manera que hace poco hubiésemos imaginado. Se estima que, en el horizonte de 2030, alrededor de 800 millones de puestos de trabajo podrían verse amenazados. El propio Fondo Monetario Internacional prevé que la IA, de forma directa o indirecta, afecte ya a casi el 40% de las ocupaciones actuales. La pérdida repentina de un empleo puede dejar a un trabajador en un laberinto financiero, erosionar su autoestima y desdibujar el sentido de propósito que el trabajo suele aportar.

En el ámbito de los servicios públicos, la adopción de sistemas de decisión automatizados plantea riesgos sistémicos para los derechos humanos, y en particular para los derechos sociales. El Comité de Ministros del Consejo de Europa ha señalado que la escasa supervisión humana sobre estas herramientas puede traducirse en errores que impactan de lleno en el acceso de las personas a prestaciones básicas como la salud, la vivienda o el subsidio de desempleo.

Además, los algoritmos que se utilizan para distribuir esos servicios en ocasiones reproducen sesgos históricos, marginando aún más a colectivos vulnerables. Cuando los datos de entrada son asimétricos o los criterios de diseño son discriminatorios, el resultado es una exclusión sistemática y silenciosa. La opacidad con la que funcionan muchos de estos sistemas agrava el problema, ya que dificulta que las personas entiendan por qué se les ha denegado ayuda o, peor aún, cómo pueden cuestionar una decisión que, ostensiblemente, “tomó la máquina”.

### **En derecho a la libertad de expresión y de información**

Hoy, el derecho a la información es también un derecho político, en cuanto es un instrumento para participar, activamente, en las actividades que cumple el Estado, y conocer qué hacen los gobernantes, pues, comprende las facultades de investigar, de difundir y de recibir información, y todo tipo de mensajes, sin que, para ello, sean admisibles impedimentos, obstáculos, o fronteras. Ni mucho menos, que haya límites impuestos por las autoridades, caprichosamente, porque la Constitución Política no lo permite, al disponer que en Colombia “no habrá censura”. (Carvajal Martínez, 2010, p. 162)

Los riesgos de la inteligencia artificial impactan directamente en los derechos comunicativos, en especial en la libertad de expresión y de información. Una de las amenazas más relevantes es la llamada *censura algorítmica*, que ocurre cuando los sistemas de IA restringen o manipulan el acceso y difusión de contenidos en el entorno digital, generando consecuencias derivadas tanto de su diseño como de su aplicación práctica. (Pérez-Ugena, 2024)

Por otra parte, la capacidad de la IA para generar y difundir desinformación añade una nueva dimensión de riesgo. Los sistemas de generación automática de contenidos —

incluidos los *deepfakes*— ponen en entredicho la confianza en la veracidad de la información, afectando bienes jurídicos tan relevantes como la honra, la integridad moral, la seguridad pública e incluso la estabilidad institucional. El escenario descrito obliga a replantear los marcos regulatorios tradicionales, que resultan insuficientes frente a los efectos disruptivos de estas tecnologías en la esfera comunicativa. (Pérez-Ugena, 2024)

En consecuencia, se hace necesaria una regulación orientada a la transparencia algorítmica, la rendición de cuentas y la protección efectiva de la libertad de expresión e información, con el fin de equilibrar el desarrollo tecnológico con el respeto a los derechos fundamentales.

### **Derecho a la tutela judicial efectiva**

La introducción de la inteligencia artificial en la administración de justicia plantea una tensión evidente entre las ventajas de la tecnología y la necesidad de preservar las garantías procesales inherentes al derecho a la tutela judicial efectiva. Si bien la IA puede aportar eficiencia, celeridad y uniformidad en las decisiones, también conlleva riesgos de opacidad, sesgo y error, que resultan inaceptables cuando se trata de la resolución de conflictos jurídicos que comprometen derechos fundamentales.

En este sentido, la doctrina ha insistido en la importancia de asegurar un control humano significativo sobre los sistemas de IA aplicados al ámbito judicial. Este control no se limita a una supervisión meramente formal, sino que exige la posibilidad real de revisión, explicación y corrección de los resultados emitidos por la máquina. Solo así se garantiza que la decisión final respete los principios de contradicción, imparcialidad y motivación, elementos esenciales del proceso justo.

Además, resulta imprescindible establecer criterios normativos de calidad, seguridad y ética que rijan el diseño y uso de estas herramientas, de manera que se asegure su fiabilidad técnica y su compatibilidad con el marco constitucional y convencional de derechos. De lo contrario, existe el riesgo de sustituir la función judicial por un mecanismo tecnocrático carente de legitimidad democrática y sin capacidad de responder adecuadamente a la complejidad de los conflictos humanos.

En definitiva, el uso de la tecnología de IA en los sistemas legales puede socavar derechos y garantías fundamentales como el debido proceso, la igualdad ante la ley y la tutela judicial efectiva. Los sistemas de IA mal entrenados o aquellos utilizados sin supervisión humana pueden generar decisiones sesgadas, discriminatorias o de otro modo injustas. La Corte ha establecido que la IA puede ser una herramienta útil para asistir a los jueces, pero no debe reemplazar la toma de decisiones humana ni vulnerar el debido proceso.

El juez, cuando encuentre necesario y pertinente recurrir a sistemas de IA, puede emplearlos razonada y ponderadamente. Esto supone un enfoque de protección de derechos fundamentales, que valore y considere las mejores prácticas, así como la aplicación de criterios éticos y de respeto a los mandatos superiores. De acuerdo con ello, so pena de comprometer eventualmente su responsabilidad a raíz del uso indiscriminado e imprudente de estas tecnologías, el funcionario judicial que las

emplee deberá respetar, como presupuesto mínimo en la materia, el criterio de no sustitución de la racionalidad humana, así como atender las cargas de transparencia, responsabilidad y privacidad. En virtud de aquel criterio, es factible emplear inteligencia artificial en labores propias de la justicia siempre que el uso de dichas herramientas no remplace labores jurisdiccionales indelegables e irremplazables, como lo son aquellas que requieren del razonamiento lógico y humano a efectos de interpretar los hechos, las pruebas, motivar la decisión o adoptarla. (Sentencia T-323/24, 2024)

### **El algoritmo como núcleo de los sistemas de inteligencia artificial y su vínculo con la protección de los derechos fundamentales**

La idea central es que el algoritmo es el motor esencial de todo sistema de inteligencia artificial, ya que traduce instrucciones en un lenguaje que las máquinas pueden ejecutar. Esto permite procesar grandes volúmenes de información y obtener resultados que simulan decisiones humanas. (Pérez-Ugena, 2024)

En la práctica, los algoritmos son la pieza que convierte la abstracción matemática en una herramienta aplicada, y su relevancia es tal que sin ellos no existiría la IA. De hecho, su capacidad para identificar patrones invisibles al ojo humano es lo que ha permitido que la IA se use en campos tan diversos como la medicina, el comercio electrónico o la seguridad digital.

Aunque los algoritmos no “piensan” ni buscan entender la realidad, sí son extremadamente útiles para resolver problemas prácticos. Sin embargo, esta característica también marca sus límites: dependen totalmente de cómo estén programados, lo que los hace vulnerables a errores de diseño o a la calidad de los datos que procesan.

Hay muchos tipos de algoritmos, pero para el Derecho Público interesan especialmente aquellos que inciden en el acceso a derechos fundamentales como la educación, el empleo o la justicia. Si un algoritmo decide quién accede a una beca, un subsidio o incluso quién es señalado como sospechoso en un proceso penal, entonces su diseño y uso tienen consecuencias constitucionales directas. (Pérez-Ugena, 2024)

En este punto, se evidencia que los algoritmos no son neutrales; influyen en la vida social y en la garantía de derechos. Por eso, su análisis desde el derecho no es opcional, sino necesario, pues pueden ser instrumentos de inclusión o, por el contrario, de exclusión y discriminación. La falta de explicabilidad no solo es un desafío técnico, sino también jurídico y ético. Si una persona recibe una negativa en un crédito, es discriminada en una selección laboral o enfrenta una decisión judicial influenciada por un algoritmo, debe poder conocer por qué sucedió. De ahí que la transparencia, supervisión y corrección de sesgos sean principios esenciales para que la IA no solo sea eficiente, sino también legítima desde una perspectiva democrática. En esa línea “Se puede precisar que el objetivo de análisis a la hora de estudiar cómo se afecta a los derechos fundamentales, no son los algoritmos en sí mismos, sino los procesos de toma de decisiones en torno a los algoritmos”. (Pérez-Ugena, 2024, p. 317)

### 2.3. Transparencia y explicabilidad

La transparencia algorítmica es uno de esos principios que deberían ser básicos cuando hablamos de inteligencia artificial. Básicamente, quiere decir que las personas tienen derecho a saber cómo y por qué una máquina llega a una conclusión. Esto resulta especialmente importante cuando la decisión de la máquina puede meternos en problemas, como cuando nos imponen una multa, se decide quién cuida de un niño o se dicta una sentencia judicial. Si no entendemos el razonamiento de la IA, es muy difícil defendernos o, al menos, saber a qué nos estamos enfrentando. El Proyecto de Ley Estatutaria sobre Inteligencia Artificial recoge este postulado en el denominado **principio de transparencia y explicabilidad**, al señalar que:

Los desarrollos, resultados y subprocesos de las Inteligencias artificiales serán inteligibles, trazables, explicables y comprensibles. La ciudadanía deberá conocer el propósito y alcance de cada componente algorítmico y la forma en que contribuye a los resultados de los sistemas, para saber los motivos por los que la Inteligencia Artificial llega a una u otra conclusión o decisión. (Cámara de Representantes de Colombia, 2024, art 4, inc 10)

En los últimos años, la noción de buen gobierno se ha relacionado estrechamente con la transparencia, entendida tanto en el acceso a la información como en la gestión pública. Este principio también se ha trasladado al ámbito de la inteligencia artificial, al permitir que los ciudadanos comprendan cómo las entidades estatales toman decisiones y aplican la tecnología. De ahí que la transparencia se haya consolidado como un eje central en las estrategias y documentos oficiales de los gobiernos, como lo señala la Universidad de Harvard en uno de sus informes, lo cual implica que el uso de la inteligencia artificial en políticas públicas y servicios estatales estará necesariamente mediado por este principio (Guío Español, 2020).

Aunque el principio de explicación se presenta a menudo como un complemento a la transparencia, tiene un matiz que lo distingue. La información que se pone a disposición del público de manera transparente no solo necesita estar visible, sino que también debe ser comprensible, tanto para quienes programan y operan el sistema como para quienes pueden sufrir las consecuencias de sus elecciones y resultados. En consecuencia, el acceso disponible trasciende lo meramente visible; persigue que el contenido y la forma de esa información se presenten de modo que cualquiera pueda asimilarla sin dificultad, tomando en cuenta las diversas capas de complejidad que rodean hoy a estas tecnologías. (Guío Español, 2020)

Las herramientas de inteligencia artificial suelen funcionar como verdaderas “cajas negras”. Metemos información dentro, y cuando sale el resultado, a veces no queda claro cómo o por qué se tomó esa decisión. Esa falta de transparencia es un gran problema si la IA se usa en áreas tan sensibles como la justicia, la salud o la seguridad, donde un simple error puede cambiar por completo la vida de alguien.

Sin un poco de claridad sobre el funcionamiento de los sistemas automáticos, resulta difícil para la gente saber por qué se tomaron ciertas decisiones. Cuando no entendemos el “por qué” detrás de un resultado, es casi imposible cuestionarlo o pedirlo de vuelta. Esto deja a

las personas en una posición rara, ya que se sienten afectadas, pero no tienen las herramientas para defenderse. Por eso el tema de la rendición de cuentas y la justicia se vuelve tan importante en un mundo donde cada vez más tareas las hacen máquinas. Si no se sabe quién es responsable de una decisión automática, la confianza en ese sistema se tambalea.

#### **2.4. Responsabilidad humana y control institucional**

Uno de los líos más grandes que tenemos hoy en día es decidir quién se hace cargo cuando un programa que toma decisiones por su cuenta hace daño. ¿Debería pagar el que escribió el código? ¿La empresa que lo puso en marcha? ¿O tal vez el juez que acepta los resultados sin pararse a verificar? Como la responsabilidad rebota entre tantos, muchas veces la cosa se desdibuja y termina pareciendo que la tecnología tiene su propia zona de impunidad.

Por eso, se sigue diciendo que siempre tiene que haber un ser humano que decida en los momentos en que se juegan derechos. La nueva Ley sobre inteligencia artificial que se está debatiendo en Colombia lo deja muy claro:

Principio de supervisión y prevalencia de la inteligencia humana: Se preferirá la decisión humana respecto de los resultados derivados de la Inteligencia Artificial. Toda decisión tomada por las inteligencias artificiales será susceptible de revisión humana y será responsabilidad ética y legal de personas naturales y jurídicas. (Cámara de Representantes de Colombia, 2024, art 4, inc 5)

El principio aquí descrito es relevante para cualquier sistema de inteligencia artificial que opere con cierto grado de autonomía en sus decisiones. En su etapa inicial, el enfoque debe ser que los usuarios mantengan un control total sobre el proceso, una filosofía que se traduce en el modelo conocido como “Human-in-the-loop”. A medida que la tecnología madura y se consolida en el país, es de esperar que estos sistemas empiecen a operar con mayor independencia; sin embargo, esa mayor autonomía siempre deberá ir acompañada de salvaguardias que permitan a los humanos intervenir, sobre todo cuando surjan efectos no deseados o inesperados. Esta evolución no se planificará en términos estrictamente técnicos, sino que tendrá en cuenta sus posibles repercusiones sociales, con particular atención a su impacto en el futuro del trabajo y al desplazamiento de personas de determinados puestos y tareas. (Guío Español, 2020)

La Corte Constitucional ha marcado límites en la forma en que la justicia colombiana se relaciona con la inteligencia artificial. Por primera vez, se establecen reglas claras sobre cuándo y cómo se puede usar esta tecnología en los juzgados. Pero lo más importante es que esas reglas no solo tratan de lo técnico, sino también de lo humano: proteger los derechos de las personas y garantizar que todo lo que ocurra en los tribunales sea claro y transparente:

(...) criterios orientadores en cuanto al uso de herramientas de IA como ChatGPT por parte de los despachos judiciales en el país: a. Transparencia, entendida como la obligación de evidenciar con claridad y precisión el uso, alcances y ubicación en las actuaciones o decisiones de los resultados obtenidos por la utilización de tales herramientas, que permita a los usuarios e interesados su pleno conocimiento y la posibilidad efectiva de contradicción. B. Responsabilidad, comprendida como

aquella obligación que existe de que el usuario de la herramienta de IA se encuentre capacitado y comprenda los impactos del uso de estas tecnologías, para a su vez dar cuenta del origen, idoneidad y necesidad del uso de la IA y la información suministrada por la misma, la cual debe ser verificada. c. Privacidad, es aquel deber de custodiar y proteger la reserva de los datos personales y sensibles que se ponen en conocimiento de la administración de justicia para cumplir con los fines propios de la Rama Judicial. d. No sustitución de la racionalidad humana, como expresión de la imposibilidad ética y jurídica de sustituir la acción y la responsabilidad del individuo de la especie humana en la gestión de las actuaciones y decisiones judiciales. e. Seriedad y verificación, que implica la obligación de realizar un estricto escrutinio sobre las fuentes, alcances, restricciones, posibilidades, falencias y riesgos que presente la herramienta de cara a la actuación en curso o a la solución del problema jurídico correspondiente. f. Prevención de riesgos, como mandato en cuanto aplicar los estándares adecuados de control sobre situaciones que generen riesgo por la aplicación de tecnologías tales, en aspectos como imprecisiones, desactualizaciones, alucinaciones, sesgos, inconsistencias y demás. g. Igualdad y equidad, en cuanto se erradiquen todas las formas de discriminación relacionadas con la aplicación de sesgos derivada del uso de tales tecnologías y su impacto negativo en la eficacia de los derechos humanos. h. Control humano, en tanto considerando los anteriores criterios, siempre se permita la realización efectiva de escrutinios sobre las actuaciones y decisiones en que se usen herramientas de IA, mediante el acceso a la debida información y el uso de recursos que deban ser resueltos por autoridades humanas. i. Regulación ética, que implica el desarrollo de estándares de comportamiento individual que se adecúen a los mandatos superiores y legales y a las pautas razonables para el uso de tales tecnologías por parte de los funcionarios y servidores de la Rama Judicial. j. Adecuación a buenas prácticas y estándares colectivos, en tanto se apliquen los esquemas razonables que se definan para el funcionamiento de la Rama Judicial, desde su autonomía e independencia, a partir de las definiciones que adopten sus autoridades, tanto en sede de administración como de orientación jurisprudencial. k. Seguimiento continuo y adaptación, a efecto que el uso de tales tecnologías consulte los avances jurídicos, sociológicos y tecnológicos que se vayan implementando, así como los esquemas de mejora y control que se construyan en forma progresiva. l. Idoneidad. El uso de las tecnologías de la información y de las comunicaciones en la gestión y trámite de los procesos judiciales y asuntos en curso debe ser adecuado para facilitar y agilizar el acceso a la justicia. (Sentencia T-323/24, 2024)

### **Capítulo 3. Marco normativo y político en Colombia**

En Colombia, el crecimiento de la inteligencia artificial ha dejado claro que se necesita unas reglas del juego que hagan que su uso sea ético, transparente y respeten los derechos de todas las personas. Aunque todavía no hay una ley definitiva y completa que regule esta tecnología de manera formal, ya hay varias iniciativas, propuestas y documentos que están sirviendo como primeras piezas para armar ese nuevo marco legal.

### 3.1. Diagnóstico legal actual y vacío regulatorio

Colombia tiene que avanzar en inteligencia artificial si quiere usar la tecnología de manera segura, sobre todo en el sector público, donde el impacto es mayor. Hoy por hoy no contamos con una norma vinculante específica que regule la IA; tenemos normas que tocan las tecnologías de la información, pero son poco evidentes cuando hablamos de algoritmos autónomos. Por eso resulta urgente sentar a todas las partes –gobierno, empresas, academia y comunidad– e implementar reglas que, además de ser claras, se puedan cumplir. Mientras se dan esos pasos, vale la pena poner en marcha mecanismos de auto restricción y autocontrol que garanticen que los sistemas se usen con la ética, la transparencia y la responsabilidad que la ciudadanía merece.

En Colombia (Nieto Martínez, 2025), diez proyectos de ley sobre inteligencia artificial actualmente se analizan en el Congreso. Nueve de estas iniciativas han sido propuestas por parlamentarios y una más proviene del Gobierno Nacional. Algunos textos se concentran en regular aplicaciones concretas de la IA, mientras que otros persiguen un marco normativo más amplio. Un repaso general revela que la mayoría de las iniciativas todavía discuten en sus primeras comisiones; solo una, que introduce una pena agravada para la falsedad personal facilitada por estas herramientas y establece otras medidas de política pública, ha llegado al cuarto debate. A continuación, un breve resumen de las normativas señaladas:

**Tabla 1**

#### Resumen de Proyectos de Ley sobre Inteligencia Artificial en Colombia

#	Proyecto(s)	Título resumido	Cámara	Tipo	Estado actual	Plazo límite
1	59/2023	Lineamientos de política pública sobre IA	Senado	Ordinaria	Pendiente discutir ponencia (2.º debate)	20 jun. 2025
2	91/2023	Deber de información para uso responsable de IA	Senado	Ordinaria	Pendiente rendir ponencia (2.º debate)	20 jun. 2025
3	130/2023	IA y derecho al trabajo	Senado	Ordinaria	Pendiente discutir ponencia (2.º debate)	20 jun. 2025
4	225/2024 – 360/2024	Modificación art. 296 Código Penal (Ley 599/2000)	Senado	Ordinaria	Pendiente discutir ponencia (4.º debate)	20 jun. 2025
5	255/2024	Uso de IA para reducir	Senado	Ordinaria	Pendiente discutir	20 jun. 2025

		siniestros viales			ponencia (2.º debate)	
6	005/2024	Ley de IA ética y sostenible para el bienestar social	Cámara	Ordinaria	Pendiente discutir ponencia (1.º debate)	20 jun. 2025 / 20 jun. 2026
7	152/2024	Régimen general de protección de datos personales	Cámara	Estatutaria	Pendiente discutir ponencia (1.º debate)	20 jun. 2025
8	154/2024	Regulación integral de la IA y modificación parcial a la Ley 1581 de 2012	Cámara	Estatutaria	Pendiente discutir ponencia (1.º debate)	20 jun. 2025 / 20 jun. 2026
9	293/2024	Entrenamiento de modelos IA y gestión colectiva de derechos de autor	Senado	Ordinaria	Pendiente discutir ponencia (1.º debate)	20 jun. 2025 / 20 jun. 2026
10	442/2025	Desarrollo ético y responsable de la inteligencia artificial	Senado	Ordinaria	Pendiente rendir ponencia (1.º debate)	20 jun. 2025 / 20 jun. 2026

Fuente: (Nieto Martínez, 2025)

La ausencia de una regulación específica sobre inteligencia artificial en Colombia no significa que estas tecnologías operen en un vacío legal. Al igual que cualquier otro bien o servicio, los sistemas inteligentes están sujetos al marco normativo vigente, el cual abarca la Constitución, las leyes generales y la normativa reglamentaria. En consecuencia, sus desarrolladores, distribuidores y usuarios deben garantizar un manejo adecuado de los datos personales, observar los derechos de los consumidores, respetar la propiedad intelectual y asumir las responsabilidades que correspondan en caso de daños. Asimismo, cuando estas tecnologías se aplican en sectores regulados —como el financiero— deben ajustarse a las disposiciones sectoriales correspondientes. (Nieto Martínez, 2025)

En este sentido, es importante señalar que las tecnologías emergentes —en particular los sistemas de inteligencia artificial— se inscriben dentro del conjunto más amplio de las

tecnologías de la información y las comunicaciones (TIC); en consecuencia, desde esta perspectiva, su funcionamiento ya estaría sujeto a alguna forma de regulación prevista en el marco jurídico colombiano. Traemos a colación una definición legal que las describe como “el conjunto de recursos, herramientas, equipos, programas informáticos, aplicaciones, redes y medios que permiten la compilación, procesamiento, almacenamiento, transmisión de información como voz, datos, texto, video e imágenes” (Ley 1431, 2009, art. 6).

Ahora bien, el CONPES 3975 (Departamento Nacional de Planeación [DNP], 2019) caracteriza la inteligencia artificial como un área de la informática que busca afrontar tareas cognitivas usualmente relacionadas con la inteligencia humana o con entidades que demuestran esa capacidad, es decir, que pueden adaptarse a entornos variables. Esta labor depende, a su vez, del diseño de sistemas computacionales, de la existencia de datos pertinentes y de algoritmos que procesen dicha información.

En línea con lo anterior, en octubre de 2021 se publicó el Marco Ético para la Inteligencia Artificial en Colombia (Guío Español, 2020), un documento orientador de soft law que ofrece recomendaciones prácticas a las entidades públicas al formular y gestionar proyectos que incorporan IA. La guía aborda, además, principios éticos que deben tenerse en cuenta durante el diseño, desarrollo e implementación de estas tecnologías en el territorio nacional.

De manera paralela, durante ese mismo lapso y en medio de la crisis provocada por la pandemia de Covid-19, el Estado colombiano avanzó en un plan de digitalización de la justicia que quedó oficialmente plasmado en el CONPES 4024 de 2021 y en el Decreto 806 de 2020; dicho enfoque se institucionalizó luego con la Ley 2213 de 2022. Estas disposiciones, surgidas de la imposibilidad temporaria de seguir brindando atención presencial en los despachos judiciales, apuntaron igualmente a establecer vías más ágiles y eficaces para llevar a cabo los trámites ante los juzgados.

Finalmente, el documento CONPES 4144 (Departamento Nacional de Planeación [DNP], 2025), aprobado el 14 de febrero de 2025, marca el camino que Colombia seguirá con la inteligencia artificial hasta el año 2030, viéndola como una herramienta clave para impulsar el cambio social y económico. Según este plan, el Gobierno se compromete a que toda aplicación de IA respete principios como la ética, la transparencia, la inclusión y la sostenibilidad. Para lograrlo, se crea una estructura de coordinación encabezada por los ministerios de TIC, Ciencia y el DNP. En total, el plan incluye 106 acciones divididas en seis grandes temas: ética y gobernanza, datos e infraestructura, innovación, talento digital, reducción de riesgos y adopción tecnológica. Para esto se destinan cerca de 479 mil millones de pesos. Entre las metas concretas están duplicar la inversión en investigación, capacitar a 150 mil personas y 15 mil expertos en IA, reducir a la mitad la brecha digital en zonas rurales y lograr que el 80 % de las entidades públicas usen IA de forma responsable.

Por lo tanto, aunque en este momento no haya una normativa específica que regule la inteligencia artificial, cualquier proyecto que implique el manejo de información de carácter personal debe adecuarse al marco legal vigente. En particular, es fundamental que se respeten los principios establecidos en la Ley Estatutaria 1581 de 2012 y 1266 de 2008. A continuación, se presentan, respectivamente, los criterios que cada norma contempla:

Principios para el Tratamiento de datos personales. En el desarrollo, interpretación y aplicación de la presente ley, se aplicarán, de manera armónica e integral, los

siguientes principios: (...) d) Principio de veracidad o calidad: La información sujeta a Tratamiento debe ser veraz, completa, exacta, actualizada, comprobable y comprensible. Se prohíbe el tratamiento de datos parciales, incompletos, fraccionados o que induzcan a error. (ley 1581, 2012, art. 4)

Principios de la administración de datos. En el desarrollo, interpretación y aplicación de la presente ley, se tendrán en cuenta, de manera armónica e integral, los principios que a continuación se establecen: a) principio de veracidad o calidad de los registros o datos. La información contenida en los bancos de datos debe ser veraz, completa, exacta, actualizada, comprobable y comprensible. Se prohíbe el registro y divulgación de datos parciales, incompletos, fraccionados o que induzcan a error. (ley 1266, 2008, art. 4)

Como ha señalado la Corte Constitucional, crear una inteligencia artificial generativa que se adapte a las leyes y sentencias colombianas e internacionales podría ser de gran ayuda para los jueces y, además, haría más fácil que las personas accedan a la justicia. Una herramienta así permitiría a abogados, usuarios y entidades encontrar información útil de forma rápida y precisa: consultar jurisprudencia actualizada, comparar casos parecidos, leer resúmenes claros de las normas y detectar con rapidez problemas legales o posibles violaciones de derechos. Eso sí, se entiende que estas tecnologías deben servir como apoyo y no reemplazar el trabajo humano. Además, si la inteligencia artificial funciona bajo el control de la Rama Judicial, sería más fácil asegurar la transparencia, cuidar los datos personales y controlar posibles sesgos que vienen del tipo de información con la que fue entrenada. (Sentencia T-323/24, 2024)

### **3.2. Análisis del Proyecto de Ley Estatutaria sobre IA (2024)**

Dada su reciente introducción, hace conveniente referirse al Proyecto de Ley Estatutaria No 154 de 8 de octubre de 2024, que tramita la Cámara de Representantes en Colombia. Con esta iniciativa se persigue delimitar y regular la inteligencia artificial acometiendo un todo que incorpore los estándares internacionales sobre derechos humanos. Adicionalmente, se establecen barreras para el desarrollo, el uso y la implementación de esas tecnologías, se reforma la Ley 1581 de 2012 sobre protección de datos personales y se dictan otras disposiciones ulteriores. En suma, el proyecto intenta ofrecer un marco normativo coherente y actualizado frente a la creciente penetración de la inteligencia artificial en la vida pública y privada del país.

#### **Aspectos relevantes del proyecto de ley:**

##### **No aplicación del régimen regulatorio en contextos de docencia e investigación**

Según lo preceptúa la norma, en el contexto de la inteligencia artificial no habrá régimen regulatorio cuando:

Se desarrolla en el marco de actividades de carácter educativo o académico o científico en instituciones educativas de educación básica primaria y básica secundaria e instituciones de educación superior ubicadas en el territorio colombiano, siempre que la inteligencia artificial no sea instalada y entre en funcionamiento de manera permanente en la misma institución, actividades de

terceros o por fuera del ámbito educativo o científico. (Cámara de Representantes de Colombia, 2024, art. 2, párr. final, lit. a)

Ahora bien, esta excepción no implica carta blanca. Las instituciones siguen estando obligadas a respetar principios éticos, proteger los datos personales de los participantes y garantizar condiciones seguras en todos los experimentos o desarrollos. La idea es promover la investigación sin imponer cargas regulatorias innecesarias, pero sin relajar la responsabilidad.

### ¿Qué condiciones deben cumplirse para que aplique esta exención?

- El uso debe limitarse al ámbito educativo o científico.
- No puede tratarse de una implementación permanente fuera de esos espacios.
- Los desarrollos no deben tener un propósito comercial inmediato.

### ¿Qué beneficios ofrece esta flexibilidad?

- **Fomenta la innovación y el conocimiento:** universidades como el Área Andina ya han creado sus propios “pentálogos éticos”, mostrando que la autorregulación académica puede ir incluso un paso adelante de la legislación.
- **Impulsa la formación de talento en IA:** según el Ministerio de Educación, esta flexibilidad es clave para ejecutar programas de formación que cuentan con una inversión de 3.200 millones de pesos, permitiendo a los estudiantes trabajar con tecnología real, sin trabas excesivas (Ministerio de Educación Nacional, 2025).
- **Evita la fuga de talento e innovación:** al no exigir requisitos regulatorios prematuros a proyectos emergentes dentro de instituciones educativas, se reduce el riesgo de que estos busquen desarrollarse en países con normas más laxas.

### ¿Qué límites existen?

Aunque se trata de una exención, el proyecto introduce salvaguardas importantes:

- **Cláusula de transición:** si un desarrollo académico muestra potencial de impacto masivo o adopción a gran escala, deberá someterse a una evaluación de riesgos, como lo establece el Título II del proyecto.
- **Responsabilidad ética institucional:** las universidades deben dejar constancia de los principios que guían sus desarrollos en IA, como la transparencia, la no discriminación y la protección de derechos, lo cual ya hacen muchas de forma voluntaria.

### Valoración del grado de peligrosidad

De acuerdo con el texto del Proyecto de Ley Estatutaria 154 de 2024, el grado de riesgo que presentan los sistemas de inteligencia artificial se determina, ante todo, por los impactos que esas tecnologías pueden tener sobre los derechos fundamentales de las personas. El proyecto introduce una tipología escalonada de riesgo dividida en cuatro niveles, la cual se fundamenta en criterios éticos, técnicos y jurídicos (Cámara de Representantes de Colombia, 2024, art. 6).

Clasificar los modelos en función del riesgo que ocasionan resulta fundamental, pues facilita un control más riguroso sobre aquellos sistemas que pueden transgredir principios básicos. Esa clasificación también crea una norma que obliga a estudiar el impacto antes de poner en marcha cualquier nueva tecnología.

**Tabla 2**

**Clasificación de riesgos**

<b>Nivel de riesgo</b>	<b>Descripción</b>	<b>Consecuencia legal</b>
Riesgo inaceptable	Corresponde a sistemas que generan afectaciones a la seguridad, la subsistencia o los derechos humanos y fundamentales.	Prohibido por norma legal.
Alto riesgo	Aplicaciones automatizadas que pueden limitar derechos, aunque son admisibles si se garantiza calidad de datos y supervisión humana.	Permitido con condiciones estrictas.
Riesgo limitado	Comprende sistemas como chatbots o robots conversacionales que tienen baja capacidad de daño, pero interactúan directamente con personas.	Permitido con medidas de precaución.
Riesgo nulo	Sistemas que no generan ningún impacto en la seguridad ni en los derechos fundamentales del usuario.	Permitido sin restricciones adicionales.

Adaptado del Proyecto de Ley Estatutaria 154 de 2024 (Cámara de Representantes de Colombia, 2024, art. 6).

**Estudio de Impacto como Herramienta Clave**

El proyecto de Ley Estatutaria dispone que los sistemas de inteligencia artificial clasificados como de alto riesgo o riesgo limitado solo podrán ponerse en funcionamiento una vez se haya elaborado un estudio de impacto riguroso (Cámara de Representantes de Colombia, 2024, art. 9). Este documento deberá contener, al menos:

1. Una evaluación de derechos humanos que identifique afectaciones potenciales a la privacidad, la no discriminación y la autonomía de las personas.
2. Documentación técnica accesible que detalle los algoritmos, las fuentes de datos y las medidas para mitigar sesgos.
3. Un plan de gestión de riesgos con protocolos frente a fallos y mecanismos de reparación efectivos.

En este contexto, el estudio de impacto no se presenta como un trámite formal, sino como un instrumento sustantivo que materializa el principio de responsabilidad proactiva frente a los riesgos que plantea la inteligencia artificial.

## **Mecanismos de Control Institucional**

El Título III del Proyecto de Ley Estatutaria 154 de 2024 configura una estructura de control que busca garantizar la implementación ética, segura y conforme a los derechos fundamentales de los sistemas de inteligencia artificial (IA) en Colombia. Este diseño normativo apuesta por una institucionalidad reforzada, centrada en dos actores principales: la Superintendencia de Industria y Comercio (SIC) y el Ministerio de Tecnologías de la Información y las Comunicaciones (MinTIC), con una lógica de vigilancia articulada al enfoque de riesgo previsto en el articulado anterior.

### **1. La Superintendencia de Industria y Comercio como autoridad de protección y fiscalización**

El artículo 16 modifica el artículo 19 de la Ley 1581 de 2012 para confirmar a la SIC —a través de su Delegatura para la Protección de Datos Personales— como la entidad encargada de vigilar que los sistemas de IA respeten los principios, derechos y garantías constitucionales en el tratamiento de datos. Además, el proyecto amplía sus competencias para incluir expresamente la inspección, control y vigilancia sobre el desarrollo, uso e implementación de estos sistemas, así como la administración de la futura Plataforma de Certificación de Sistemas de Inteligencia Artificial (Cámara de Representantes de Colombia, 2024, art. 16).

Esta función se complementa con lo previsto en el artículo 18, que impone a la SIC la obligación de llevar a cabo auditorías sobre sistemas en proceso de certificación o que representen riesgos altos, de forma que se garantice la protección efectiva de los derechos de los usuarios (art. 18). Esta auditoría se constituye en un mecanismo preventivo clave, orientado a reforzar la transparencia, supervisión humana y trazabilidad de decisiones automatizadas.

### **2. MinTIC: regulación, pedagogía y administración de registros**

El artículo 17 amplía las funciones del MinTIC al incluir expresamente su rol en la supervisión del cumplimiento normativo en materia de IA (función 31), así como en la promoción de campañas pedagógicas para educar a la ciudadanía sobre sus derechos y los riesgos asociados a estas tecnologías (función 32). Además, este ministerio será responsable de gestionar el Registro Nacional de Certificación de Sistemas de IA (función 33), lo que lo posiciona como eje articulador del componente técnico y formativo del régimen regulatorio.

## **Beneficios del Enfoque Basado en Riesgo**

**Precisión regulatoria.** El modelo evita someter a un elevado nivel de control a aplicaciones de bajo riesgo, como los filtros de fotografías, y concentra los recursos de inspección en Sistemas críticos, tales como la justicia algorítmica.

**Seguridad jurídica.** El mecanismo ofrece a inversores y emprendedores un marco predecible, al clasificar las reglas de juego según el nivel de impacto esperado.

Adaptabilidad. La clasificación de categorías puede revisarse por acto administrativo, sin exigir reformas legales, lo que resulta esencial en un ámbito que evoluciona a gran velocidad.

La estructura del Título III complementa el enfoque de riesgo adoptado por el proyecto. Este diseño evita sobre regular aplicaciones de baja complejidad técnica (como filtros de imágenes) y concentra los mecanismos de control en sistemas de alto impacto, como los que se emplean en justicia, salud o seguridad. Al priorizar la proporcionalidad, se fortalece la eficiencia regulatoria y se brinda mayor seguridad jurídica a los desarrolladores e inversionistas.

### **Consentimiento informado en el tratamiento de datos mediante sistemas de inteligencia artificial**

Consentimiento informado. En el uso de los sistemas de inteligencia artificial los usuarios deberán manifestar de manera libre y voluntaria su consentimiento para asumir los riesgos expresamente identificados y comunicados por los productores y proveedores que llegare a representar para sus derechos humanos y fundamentales y para el tratamiento de sus datos personales. (Cámara de Representantes de Colombia, 2024, art. 10, inc. 1).

El consentimiento informado constituye un pilar normativo en la gobernanza del tratamiento de datos personales cuando interviene la inteligencia artificial. En este sentido, el Proyecto de Ley 153 de 2024 subraya que el consentimiento, lejos de ser un gesto rutinario, debe ser un proceso en el que la persona interesada capta, sin ambigüedad, por qué, cómo y para qué sus datos serán recogidos, analizados o empleados mediante tecnologías automatizadas.

Este enfoque genera un deber más sólido para desarrolladores, responsables y operadores, quienes han de ofrecer a los usuarios información accesible, redactada en lenguaje llano, y estructurada para que se comprenda a primera vista. La explicación no debe limitarse a los datos en sí, sino que ha de extenderse a los algoritmos empleados, a los riesgos de discriminación o abuso, y a las posibles repercusiones de la automatización sobre los derechos del interesado.

El consentimiento será válido únicamente si resulta libre, previo, expreso, específico e informado; esto implica que cada persona debe poder aceptar, negar y revocar en cualquier instante el tratamiento de sus datos en cada fase del funcionamiento del sistema. Dicho principio protege la autonomía individual y el dominio sobre la propia información, tal como exigen la Constitución colombiana, el régimen de habeas data y los estándares internacionales sobre derechos digitales.

### **Protección del empleo**

El Proyecto de Ley 153 de 2024 presenta una cláusula innovadora que protege el empleo frente a la automatización, en particular cuando un sistema de inteligencia artificial desplaza a un trabajador. Según esta disposición, el empleador debe ofrecer al afectado un puesto de igual o de superior jerarquía durante el resto de la vigencia del contrato. La norma busca salvaguardar la estabilidad laboral, evitando despidos motivados únicamente

por la introducción de nuevas tecnologías (Cámara de Representantes de Colombia, 2024, art. 12).

A pesar de su intención positiva, la disposición requiere un análisis minucioso a la luz del Código del Trabajo y la jurisprudencia constitucional sobre el derecho al trabajo. Su implementación plantea interrogantes prácticos: ¿qué ocurre si no existe un puesto equivalente? ¿Es obligatoria la reubicación o puede pactarse? ¿Puede interpretarse como estabilidad absoluta? Además, la norma podría entrar en conflicto con la autonomía privada, la libertad de empresa y la autoridad interna del empleador para organizar su fuerza laboral.

Por tanto, si bien la norma añade una dimensión social relevante al diálogo sobre inteligencia artificial y trabajo, su efectiva aplicación requerirá una regulación precisa y una adecuada sincronización con la legislación laboral vigente, a fin de evitar lagunas que puedan derivar en conflictos judiciales o desincentivar a las empresas a adoptar la tecnología.

### **Sistema nacional de evaluación y validación de tecnologías basadas en IA**

Una de las propuestas más llamativas del nuevo Proyecto de Ley 154 de 2024 es la creación de una plataforma nacional que certifique los sistemas de inteligencia artificial, y que estaría bajo la supervisión directa del Estado. Con esta herramienta, las autoridades podrían registrar, monitorear y evaluar de forma técnica cada sistema de IA que quiera operar en el país, asegurándose de que antes de llegar al mercado se cumplan varios controles preventivos (Cámara de Representantes de Colombia, 2024, art. 19).

La idea es que cualquier modelo pase por una auditoría que examine su comportamiento desde ángulos éticos, legales, de ciberseguridad, de desempeño y de protección de derechos fundamentales. De esa manera se fijaría un estándar básico de calidad, sobre todo en sectores tan sensibles como la justicia, la salud, la educación y la administración pública. Al hacer obligatoria esta certificación, el Estado ayudaría a reducir riesgos grandes y, lo más importante, a aumentar la confianza de la ciudadanía en que las herramientas automatizadas se usan de manera responsable.

Aun así, el sistema debe encontrar una forma de equilibrar lo que le conviene al público con lo que hace sentir seguro al creador que su invento no se roba. Certificar una IA no puede convertir el proceso en una lupa que mire los secretos comerciales, las recetas de negocio o los códigos guardados tras una patente. Por eso, la misma plataforma tiene que ofrecer puertas de atrás que sigan siendo privadas, usando técnicas de anonimato o de medición neutrales que mantengan a salvo la novedad sin cerrar la ventana de transparencia y rendición de cuentas.

Así las cosas, una red nacional de certificación podría acabar siendo la brújula de la gobernanza algorítmica, animando a la industria a experimentar, atrayendo inversión privada en inteligencia artificial y defendiendo a la vez los derechos de quienes usan y de quienes fabrican estas herramientas.

### **Análisis Crítico del Proyecto de Ley 154 de 2024: Aspectos Negativos y Propuestas de Mejora**

## **Confusión por definiciones vagas**

El texto da unos pocos ejemplos de lo que entiende por “inteligencia artificial”, “riesgo inaceptable”, “explicabilidad” o “evaluación de impacto algorítmico”, pero muchas de esas definiciones suenan poco precisas y, a veces, demasiado generales. Esa falta de claridad puede hacer que luego sea complicado a la hora aplicar la norma jurídica. Por poner un caso, nunca se dice exactamente cómo se mide el “riesgo” ni qué pruebas técnicas y legales tiene que pasar un modelo para caer en uno u otro nivel.

## **Desequilibrio en el Enfoque Regulatorio**

El texto del proyecto concede prioridad casi exclusiva a la mitigación de riesgos, descuidando el acompañamiento sistemático de medidas que estimulen la innovación. En el Artículo 1, la inteligencia artificial se describe, ante todo, como una amenaza, sin que se reconozca su capacidad de transformar sectores como la salud, la educación o la productividad. Esta visión restringida choca con el propósito del CONPES 3975 de avanzar en la transformación digital y, al hacerlo, corre el riesgo de desalentar inversiones tecnológicas que la economía nacional necesita.

La justificación actual del Proyecto de Ley Estatutaria No. 154 de 2024 presenta un enfoque desbalanceado que prioriza los riesgos de la IA sobre sus beneficios, aspectos que deben estar adecuadamente equilibrados. La descripción de la IA como una "amenaza" para los derechos fundamentales genera una percepción negativa y desproporcionada de la tecnología. Si bien es crucial reconocer los desafíos éticos que plantea la IA, es igualmente importante resaltar sus potenciales transformadores, algo que el texto no logra adecuadamente. Esta visión pesimista puede desalentar la innovación y subestima el papel clave que puede desempeñar la IA en el desarrollo social y económico y obstaculizar el desarrollo y progreso del país. (Instituto Colombiano de Derecho Procesal [ICDP], 2024, p. 3)

## **Ausencia de mecanismos técnicos claros para la evaluación de riesgo**

A pesar de que el proyecto distingue cuatro categorías de riesgo (nulo, limitado, alto e inaceptable) no ofrece un procedimiento técnico claro para realizar esa clasificación. Esta omisión puede generar decisiones arbitrarias, debilitar la seguridad jurídica y provocar una aplicación irregular en las entidades públicas y privadas que desarrollen o usen sistemas de inteligencia artificial.

El proyecto debería adoptar un modelo técnico y normativo estandarizado para la evaluación del impacto algorítmico y del riesgo, sirviéndose como guía el Artificial Intelligence Risk Framework elaborado por el NIST en Estados Unidos o el esquema propuesto en el Reglamento europeo sobre inteligencia artificial.

## **Exención de regulación en ámbitos académicos y de investigación**

Aunque resulta razonable que se excluya explícitamente el uso de sistemas de inteligencia artificial en entornos educativos permanentes, el hecho de prohibir su funcionamiento continuo podría afectar investigaciones que dependen de una implementación sostenida. El Instituto Colombiano de Derecho Procesal (2024) advierte, además, que, si bien la inclusión de un comité de ética para proyectos temporales es acertada, el texto del proyecto

de ley no define aspectos esenciales como su conformación ni los mecanismos que garanticen su autonomía. Asimismo, señala la necesidad de actualizar la normativa referenciada, ya que citar la Ley 10 de 1990 refleja un marco legal posiblemente obsoleto.

### **Desafíos Operativos y de Capacidad Institucional**

La implementación enfrenta **vacíos críticos de capacidad**:

**MinCiencias**, designado como autoridad rectora, carece de un presupuesto para contratar expertos en auditoría algorítmica necesarios e Infraestructura para gestionar un Registro Nacional de Sistemas de Alto Riesgo.

La **Superintendencia de Industria y Comercio** no tiene formación técnica para fiscalizar sistemas complejos como LLMs

**Ausencia de laboratorios de pruebas** certificados, esenciales para validaciones técnicas.

### **A modo de resumen**

El Instituto Colombiano de Derecho Procesal (2024) advierte que, si bien el Proyecto de Ley Estatutaria No. 154 de 2024 constituye un paso importante hacia la regulación de la inteligencia artificial en Colombia, aún persisten vacíos normativos que deben ser atendidos. La propuesta legal busca trazar un marco claro para el desarrollo y uso ético de estas tecnologías, pero su eficacia dependerá de la capacidad del legislador para afinar los mecanismos de protección de derechos y, al mismo tiempo, fomentar un ecosistema que no frene el ingenio ni la transformación digital. En última instancia, el reto no es solo normar la IA, sino hacerlo de forma inteligente: con equilibrio, visión de futuro y compromiso democrático.

### **3.3 Aplicaciones tecnológicas basadas en inteligencia artificial en la justicia colombiana.**

En los últimos años, la inteligencia artificial ha comenzado a abrirse camino como aliada estratégica en la justicia colombiana. Aunque aún está dando sus primeros pasos, varias entidades del sector han empezado a ensayar con tecnologías inteligentes para agilizar trámites, mejorar la transparencia y facilitar el acceso a los servicios judiciales. Hasta ahora, los experimentos se han enfocado en cosas como organizar procesos, anticipar decisiones, generar documentos automáticamente y explorar grandes volúmenes de información jurídica.

Una de las aplicaciones más visibles de la inteligencia artificial en Colombia es la automatización de tareas administrativas y judiciales. En este ámbito, la IA se ha integrado en sistemas de gestión documental y procesamiento de información, destacándose en entidades como el Consejo Superior de la Judicatura. Esas herramientas organizan expedientes electrónicos, clasifican documentos jurídicos con algoritmos de procesamiento de lenguaje natural y envían alertas automáticas sobre vencimientos, aliviando así la carga de trabajo de los funcionarios.

La magistrada María Victoria Quiñones, jueza del Tribunal Administrativo del Magdalena, dirige una experiencia pionera en la justicia colombiana al integrar Microsoft Copilot en su despacho. La solución de inteligencia artificial automatiza las tareas más rutinarias y ha acortado los plazos de trámite de manera notable. En una reciente audiencia pública de gran

repercusión, la herramienta elaboró un resumen en pocos minutos (antes requería seis horas), lo que permitió al equipo centrarse de inmediato en la verificación y pulido del texto final. Reconocida por su audacia tecnológica, Quiñones comparte sus hallazgos con otros jueces, estudiantes y funcionarios para extender los beneficios de la IA a todo el sistema. En el marco de un piloto nacional que agrupa a veinte magistrados, utiliza Copilot cada día para transcribir, sintetizar, redactar actas, revisar gramática, cotejar borradores de sentencia y manejar datos internos, y ya emite hasta veinte fallos por semana, casi el quintuplo de lo que producía antes. La experiencia forma parte de un programa más amplio del Consejo Superior de la Judicatura que busca modernizar la administración de justicia; desde 2024 su uso se rige por reglas que aseguran supervisión humana, transparencia y prohíben que la IA valore pruebas o redacte decisiones. El programa utiliza Microsoft Teams como la única puerta de entrada a las audiencias virtuales (que representarán el 80% del total en 2024) e intenta desatascar un trámite abrumador que dejó 2,5 millones de causas pendientes en 2023. El plan prevé extender la herramienta a otros 150 funcionarios, de modo que la justicia llegue al ciudadano más rápido y, sobre todo, con garantías sólidas. La salvaguarda de datos en Azure y la evidencia de que las reuniones por vídeo reducen a la mitad la duración de los juicios, figuran entre los argumentos que llevaron a la alianza y muestran el interés del juez Quiñones por abrir caminos seguros a quienes desde siempre sufrieron la lentitud del sistema. (Montes, 2025)

### **Plataformas PRISMA y PRETORIA: Inteligencia Artificial aplicada a la justicia colombiana**

Prisma y Pretoria son dos de las iniciativas más sofisticadas en inteligencia artificial que se han implementado dentro del sistema de justicia colombiano, cada una orientada a objetivos y ámbitos específicos dentro de la justicia penal y constitucional. Ambas plataformas combinan distintas tecnologías de IA para potenciar la eficiencia del proceso judicial, optimizar el procesamiento de extensos conjuntos de información jurídica y proporcionar a los jueces información que sustente sus decisiones, siempre preservando el papel central del operador humano.

**Prisma**, un sistema de algoritmos de predicción y aprendizaje automático, fue concebido por la Fiscalía General de la Nación con un objetivo claro: estimar, de manera cuantitativa, el riesgo de reincidencia en personas en etapa de investigación penal. Su principal aplicación radica en aportar evidencias al pedido de medidas de aseguramiento, sobre todo cuando se debate la procedencia de la privación provisional de la libertad en centro penitenciario. La herramienta se desarrolla dentro del contexto colombiano, marcado por el alto índice de hacinamiento, la sobrecarga en los centros penitenciarios y la persistente preocupación por la reincidencia delictiva. Gracias a Prisma, los fiscales disponen de un soporte analítico que les permite justificar, de forma más estructurada, si procede o no restringir la libertad del investigado mientras avanza el proceso.

El sistema asigna al individuo una puntuación de cero a cien que mide la probabilidad de cometer otro delito. Con base en este dato, el fiscal elabora un informe dirigido al juez en el que se expone el nivel de riesgo o reincidencia del procesado. Adicionalmente, la herramienta permite a los fiscales descargar el documento automáticamente cuando preparan la audiencia preliminar para la imposición de medida de aseguramiento. (Torres Abril, Silva Vásquez & Gomez Simijaca, 2022)

## **Funcionamiento y características**

Base de datos masiva: Prisma reúne información de más de 5,7 millones de personas registradas en el sistema penal oral acusatorio, cruzando bases de datos de la Policía Nacional, la Fiscalía y el INPEC.

**Modelo de aprendizaje supervisado:** El sistema aplica algoritmos supervisados para examinar antecedentes penales, las particularidades del último delito y el contexto socioeconómico del individuo, generando un puntaje de riesgo de reincidencia en tiempo real.

**Soporte a la decisión judicial:** Prisma produce un informe digital que el fiscal puede descargar y presentar ante el juez de control de garantías durante la audiencia de solicitud de medida de aseguramiento.

**Optimización de recursos carcelarios:** Al canalizar los cupos hacia quienes muestran un alto riesgo de reincidencia, Prisma permite que personas que no representan un peligro significativo accedan a medidas menos gravosas.

## **Impacto y resultados**

Según la Fiscalía, el uso de Prisma podría disminuir en un 25% los eventos de reincidencia criminal manteniendo un número constante de personas con medida de aseguramiento o reducir en 36% el total de privados de libertad registrando el mismo nivel de reincidencia.

La herramienta colabora en la disminución de los errores de tipo I y II—es decir, encarcelar a quienes no volverían a delinquir y liberar a quienes sí lo harán—y, al hacerlo, favorece decisiones más objetivas y fundamentadas en datos.

## **Limitaciones y controversias**

Torres Abril, Silva Vásquez y Gomez Simijaca (2022) advierten que, si bien el programa PRISMA proyecta una imagen prometedora sobre el papel, su aplicación práctica enfrenta desafíos considerables. La falta de infraestructura tecnológica en muchas regiones del país—como una conectividad deficiente o la carencia de equipos— limita seriamente el acceso al sistema. A esto se suma la desconfianza generalizada hacia la justicia digital y el desconocimiento técnico por parte de usuarios y funcionarios judiciales. Las autoras también señalan riesgos asociados al uso de algoritmos: errores, sesgos y la opacidad en su funcionamiento pueden generar decisiones desiguales entre personas en condiciones similares, afectando de forma más dura a poblaciones vulnerables. Asimismo, destacan que PRISMA no incorpora adecuadamente factores sociales, económicos o psicológicos, y agrupa distintos delitos sin un análisis contextual, lo que distorsiona el concepto de reincidencia. Todo esto plantea una preocupación de fondo: para que estas herramientas sean realmente justas, su transparencia debe ir más allá de lo técnico, y alcanzar también lo ético y lo comprensible para cualquier ciudadano.

**PretorIA** es un programa que usa inteligencia artificial para ayudar a los jueces en Colombia a tomar decisiones más rápidas y consistentes, especialmente en casos repetitivos o muy similares. Funciona como una especie de asistente virtual que analiza muchas sentencias pasadas y sugiere cómo podría resolverse un caso nuevo, basándose en lo que ya

se ha decidido antes. No reemplaza al juez, pero le da una herramienta más para tomar decisiones informadas y coherentes.

La idea es agilizar el trabajo en los juzgados y dar más seguridad jurídica, aunque también hay preocupaciones sobre la transparencia y el posible sesgo de los datos que usa.

PretorIA es como un buscador inteligente de sentencias judiciales, parecido a Google pero enfocado solo en fallos de la Corte. Usa inteligencia artificial para clasificar las decisiones en categorías y permite buscarlas con filtros o palabras clave, lo que hace que encontrar información sea mucho más rápido. A diferencia de una búsqueda simple, también puede identificar sentencias relacionadas con temas específicos. Todo lo que se consulta queda guardado en el sistema, lo que permite rastrear cómo funciona y qué decisiones ha ayudado a tomar, lo que da más confianza en su uso. (Torres Abril, Silva Vásquez & Gomez Simijaca, 2022).

### **Impacto y beneficios**

**Transparencia y trazabilidad:** Pretoria refuerza la seguridad jurídica y la tutela de derechos fundamentales porque sistematiza los criterios de selección de casos y permite auditar cada etapa del proceso.

**Modernización judicial:** La herramienta forma parte del ecosistema digital que integra la remisión electrónica de expedientes y continúa creando alianzas con la academia y el sector privado para avanzar tecnológicamente.

**No reemplazo de funciones decisorias:** Pretoria amplía la capacidad analítica del juez, pero nunca toma la decisión final, lo que preserva íntegramente el principio de independencia judicial.

### **Limitaciones y debates**

Su uso ha suscitado debates sobre la autonomía del poder judicial, el riesgo de sesgos algorítmicos y la posible dependencia excesiva de estadísticas a costa del análisis individual. No obstante, la plataforma es un asistente inteligente que acelera la fase de estudio preliminar, sin sustituir en ningún momento al juez en la adopción de decisiones.

También han surgido preocupaciones importantes respecto a la falta de claridad en el funcionamiento de los algoritmos, la fiabilidad de los datos que sustentan sus recomendaciones y la urgente necesidad de contar con un marco normativo preciso que garantice la protección de los derechos fundamentales dentro del sistema judicial.

### **Reflexión Final**

Tanto PRISMA como el programa PRETORIA representan iniciativas importantes del ordenamiento colombiano para incorporar la tecnología en el despacho de procesos, elevar la calidad de las decisiones y aumentar la transparencia del sistema. Sin embargo, la puesta en marcha de ambos sistemas requiere capacitación continua para jueces y funcionarios, una evaluación ética rigurosa y un andamiaje normativo sólido que asegure su empleo responsable y no lesione la independencia judicial, la igualdad ante la ley ni el derecho al debido proceso.

## **Capítulo 4. Propuestas para una regulación ética y funcional**

El diseño de una inteligencia artificial en Colombia requieren, de manera urgente, un marco normativo que impulse la innovación, al tiempo que proteja efectivamente los derechos fundamentales, mantenga el control democrático sobre los algoritmos y asegure decisiones justas. En este sentido, el capítulo que sigue presenta un conjunto de propuestas jurídicas, técnicas y éticas, diseñadas para construir una gobernanza responsable de la IA acorde a las realidades colombianas.

### **4.1. Estándares internacionales como fundamento orientador**

Las normas que rigen el uso de la inteligencia artificial en cada país deberán cimentarse en los principios orientadores que han elaborado organismos mundiales, entre los que destacan la OCDE, la UNESCO y la Unión Europea. En este sentido, la Recomendación sobre la ética de la Inteligencia Artificial que la UNESCO aprobó en 2021 indica que:

los sistemas de IA plantean nuevos tipos de cuestiones éticas que incluyen, aunque no exclusivamente, su impacto en la adopción de decisiones, el empleo y el trabajo, la interacción social, la atención de la salud, la educación, los medios de comunicación, el acceso a la información, la brecha digital, la protección del consumidor y de los datos personales, el medio ambiente, la democracia, el estado de derecho, la seguridad y el mantenimiento del orden, el doble uso y los derechos humanos y las libertades fundamentales, incluidas la libertad de expresión, la privacidad y la no discriminación. Además, surgen nuevos desafíos éticos por el potencial de los algoritmos de la IA para reproducir y reforzar los sesgos existentes, lo que puede exacerbar las formas ya existentes de discriminación, los prejuicios y los estereotipos. Algunas de estas cuestiones tienen que ver con la capacidad de los sistemas de IA para realizar tareas que antes solo podían hacer los seres vivos y que, en algunos casos, incluso se limitaban solo a los seres humanos. Estas características otorgan a los sistemas de IA una función nueva y determinante en las prácticas y la sociedad humanas, así como en su relación con el medio ambiente y los ecosistemas, creando un nuevo contexto para que los niños y los jóvenes crezcan, desarrollen una comprensión del mundo y de sí mismos, comprendan críticamente los medios de comunicación y la información y aprendan a tomar decisiones. A largo plazo, los sistemas de IA podrían disputar al ser humano el sentido especial de la experiencia y la capacidad de actuar que le son propios, lo que plantearía nuevas inquietudes sobre la autocomprensión, la interacción social, cultural y ambiental, la autonomía, la capacidad de actuar, el valor y la dignidad del ser humano, entre otras. (UNESCO, 2021, p. 10)

Este primer texto intergubernamental aprobado de forma consensuada a nivel mundial establece diez principios rectores entre los cuales destacan la transparencia y explicabilidad algorítmica, la proporcionalidad y la seguridad, la no discriminación, la privacidad y la responsabilidad. Su anexo operativo obliga a los Estados a crear marcos de evaluación de impacto ético, a mantener registros de sistemas catalogados como de alto riesgo y a establecer mecanismos efectivos de reparación. La Corte Constitucional colombiana evocó ya este documento al examinar el empleo de ChatGPT como elemento argumental en una

sentencia sobre salud infantil y recordó que toda herramienta de inteligencia artificial debe operar dentro de los límites de los derechos humanos y del debido proceso.

La tabla que sigue presenta una síntesis operativa de los once ámbitos de acción política establecidos por la UNESCO en su *Recomendación sobre la Ética de la Inteligencia Artificial* (2021). Para cada ámbito se expone la idea central y se señalan las principales implicaciones jurídicas y políticas, con el objetivo de ofrecer a los responsables de la formulación de políticas públicas un mapa claro de áreas prioritarias (evaluaciones éticas, gobernanza, datos, medio ambiente, género, educación, entre otras) que deben considerarse al diseñar marcos regulatorios y programas de implementación a nivel nacional.

**Tabla 3**

**Ámbitos de acción política de la UNESCO en inteligencia artificial**

Ámbito de acción	Idea principal	Implicaciones jurídicas/políticas
1. Evaluación del impacto ético	Establecer marcos de evaluación ética para garantizar que los sistemas de IA respeten derechos humanos y principios éticos.	Requiere marcos normativos de impacto ético vinculados a derechos humanos y protocolos de transparencia.
2. Gobernanza y administración éticas	Crear mecanismos regulatorios y de supervisión que aseguren responsabilidad, transparencia y rendición de cuentas.	Implica leyes de responsabilidad algorítmica, auditorías obligatorias y reparación de daños.
3. Política de datos	Promover gestión de datos justa, segura y responsable, asegurando calidad, accesibilidad y privacidad.	Demanda regulación en privacidad, protección de datos y estándares de interoperabilidad.
4. Desarrollo y cooperación internacional	Fomentar colaboración internacional para cerrar brechas tecnológicas y compartir buenas prácticas.	Conlleva tratados y acuerdos multilaterales para cooperación y transferencia tecnológica.
5. Medio ambiente y ecosistemas	Garantizar un desarrollo sostenible de la IA, minimizando huella ecológica y protegiendo el medio ambiente.	Obliga a incluir criterios ambientales en regulaciones de IA y políticas de sostenibilidad.
6. Género	Asegurar igualdad de	Necesita normas

	género en el diseño y uso de la IA, evitando sesgos y promoviendo inclusión femenina.	antidiscriminación y políticas de inclusión en innovación digital.
7. Cultura	Proteger y promover diversidad cultural y lingüística frente a riesgos de homogeneización tecnológica.	Exige marcos que protejan patrimonio cultural, diversidad lingüística y pluralismo digital.
8. Educación e investigación	Integrar alfabetización digital y ética de IA en educación, reforzando investigación interdisciplinaria responsable.	Requiere políticas públicas de educación digital y financiación de investigación ética en IA.
9. Comunicación e información	Impulsar acceso equitativo a información, combatir desinformación y garantizar transparencia en contenidos IA.	Implica regulación contra desinformación y transparencia en algoritmos de difusión de contenidos.
10. Economía y trabajo	Preparar sociedades para cambios laborales y económicos, asegurando condiciones justas y evitando desigualdad.	Demanda legislación laboral adaptada a automatización y políticas redistributivas frente a desigualdad.
11. Salud y bienestar social	Usar IA para mejorar salud pública y bienestar, asegurando aplicaciones seguras, justas y accesibles.	Exige normas sanitarias para IA médica y regulaciones de equidad en acceso a tecnologías de salud.

Nota. Adaptado de \*Recomendación sobre la ética de la inteligencia artificial\*, por (UNESCO, 2021).

De otro lado, los **Principios de la OCDE** sobre inteligencia artificial subrayan que los sistemas inteligentes deben ser transparentes, sólidos, susceptibles de explicación y orientados al bienestar humano. Estas pautas pueden tomarse como punto de partida para elaborar normas internas que se alineen con el conjunto de derechos fundamentales y principios constitucionales de Colombia.

Los cinco principios de la OCDE fueron adoptados por 46 gobiernos, incluidos varios que no son miembros, entre ellos Colombia, y subrayan que la inteligencia artificial debe ser humana, sólida, segura, transparente, inclusiva y sujeta a rendición de cuentas. El anexo del documento agrega orientaciones sobre estándares abiertos, innovación responsable y colaboración internacional en el uso de datos. Estas pautas sirvieron de referente a la Rama Judicial al redactar su borrador de “Plan Justicia Digital”, y también guiaron a la Unidad de

Transformación Digital e Informática al emitir sus instrucciones sobre el registro de proyectos de inteligencia artificial.

Además de fijar principios generales, la OCDE formuló un conjunto de **recomendaciones dirigidas a los responsables de políticas públicas**, con el propósito de orientar la creación de entornos normativos y regulatorios que favorezcan el desarrollo de una IA confiable, transparente y centrada en el ser humano. Dichas recomendaciones abarcan aspectos como la inversión en investigación e innovación, la consolidación de ecosistemas inclusivos, la adopción de marcos regulatorios ágiles, la capacitación y protección de los trabajadores, y la promoción de la cooperación internacional. Su valor radica en que ofrecen **parámetros prácticos y adaptables**, que pueden ser integrados por Estados como Colombia en el diseño de una política pública nacional de inteligencia artificial.

**Tabla 4**

**Recomendaciones para los responsables de políticas**

Recomendación	Idea principal	Justificación
1. Inversión en I+D en IA	Apoyar la investigación e innovación interdisciplinaria en IA confiable. Impulsar inversiones en datos abiertos, representativos y seguros.	Fomenta el desarrollo de soluciones inclusivas, seguras y libres de sesgos, fortaleciendo capacidades nacionales.
2. Ecosistema inclusivo para IA	Construir un ecosistema digital dinámico, accesible, sostenible e interoperable: datos, infraestructura y conocimiento compartido.	Garantiza acceso equitativo a recursos tecnológicos y facilita la participación de diversos actores, incluidas PYMEs y sector público.
3. Gobernanza ágil e interoperable	Crear un marco político adaptable con experimentación (sandboxes), cooperación interjurisdiccional y regulación efectiva de IA.	Permite políticas flexibles que se ajusten a los avances tecnológicos y fomenta el alineamiento normativo entre jurisdicciones.
4. Formación humana y preparación laboral	Equipar a la población con habilidades digitales y asegurar una transición justa mediante capacitación continua y protección social.	Reduce riesgos de desplazamiento laboral y contribuye a una distribución más justa de los beneficios de la IA.
5. Cooperación internacional	Fomentar colaboración multilateral y multisectorial, compartiendo conocimiento, estándares y métricas comparables en	Potencia el alineamiento global, facilita el acceso a experiencias internacionales y promueve gobernanza responsable compartida.

	IA.	
--	-----	--

Nota. Adaptado de \*Recomendación del Consejo sobre inteligencia artificial\*, por (OCDE, 2019)

Otro de los avances más llamativos en la regulación global de la inteligencia artificial ha llegado de Europa, con la reciente adopción del Reglamento sobre Inteligencia Artificial, conocido como AI Act, que fue formalmente aprobado en abril de 2024. Con este texto, la Unión Europea se erige en pionera al ofrecer el primer marco legal completo que orienta el diseño, el despliegue y el uso cotidiano de sistemas de IA, anclándose en un análisis de riesgos, en la protección de derechos y en principios de ética digital. El eje central del AI Act es el **enfoque basado en riesgos**, que constituye la lógica estructural de toda la norma. En lugar de regular de manera uniforme cualquier sistema de IA, el reglamento establece una clasificación jerárquica en función del nivel de riesgo que cada aplicación implica para los derechos fundamentales y la seguridad de las personas. De este modo, se prohíben aquellas prácticas que atentan gravemente contra la dignidad humana —como la puntuación social estatal o la manipulación subliminal—, se imponen exigencias estrictas a los sistemas de **alto riesgo** en sectores sensibles (justicia, salud, empleo, educación, infraestructuras), se asignan obligaciones de transparencia a las aplicaciones de **riesgo limitado**, y se otorga mayor libertad a los usos de **riesgo mínimo**. Este modelo regulatorio apunta a equilibrar el fomento de la innovación tecnológica con la protección efectiva de los derechos fundamentales, marcando un precedente internacional en la gobernanza de la inteligencia artificial (Unión Europea, 2024).

El futuro marco de la Unión Europea sobre la inteligencia artificial no solo establece normas generales, sino que también asigna tareas concretas a quienes diseñan y emplean estos sistemas y quienes distribuyen y hacen funcionar estos sistemas: comprobación de conformidad antes de lanzar un producto al mercado, archivo de una documentación técnica accesible, anotación en el registro europeo y supervisión humana efectiva, especialmente en los llamados sistemas de alto riesgo. A esa carga se suman exigencias sobre la claridad del funcionamiento algorítmico, el respeto a la privacidad, la prevención de sesgos discriminatorios y la solidez técnica, un conjunto de principios que ha sido alineado deliberadamente con los valores democráticos centrales de la gobernanza europea. Por esa razón numerosos analistas consideran ya el proyecto del AI Act como un modelo exportable, capaz de proporcionar normas que otras regiones puedan copiar, de estimular reformas nacionales en todo el mundo y de equilibrar de un modo razonable el impulso a la innovación con la defensa de la dignidad humana. Su estructura resulta particularmente útil para naciones como Colombia porque ofrece un mapa práctico sobre cómo anticiparse y protegerse frente a los riesgos de las tecnologías disruptivas sin cerrar la puerta al progreso.

García San José (2022) destaca que la Carta de las Naciones Unidas sigue siendo, hoy más que nunca, un faro normativo y ético en los debates sobre la gobernanza global de tecnologías emergentes como la inteligencia artificial. Lejos de ser un vestigio del pasado, su estructura axiológica ha demostrado una asombrosa capacidad de adaptación al siglo XXI. Su carácter inclusivo no se limita a los Estados fundadores, sino que ha incorporado a las naciones nacidas del proceso de descolonización, consolidando un sentido colectivo de pertenencia jurídica. Su fuerza universal no reside únicamente en el número de países que la respaldan, sino en la visión de una comunidad internacional que valora el interés común por encima de agendas particulares. Y en cuanto a su atemporalidad, esta se expresa en la

forma en que sus principios se reinterpretan y se proyectan hacia los dilemas éticos actuales: la automatización de decisiones, el diseño de algoritmos y la conducción de sistemas inteligentes. Así entendida, la Carta no solo ofrece una base jurídica estable, sino también una brújula moral para guiar la formulación de normas globales que aseguren que la inteligencia artificial esté al servicio de la dignidad humana y del bien común.

#### **4.2. El Estado de Derecho como Marco Regulatorio Democrático**

El Estado de derecho es una pieza clave para asegurarse de que el desarrollo y uso de tecnologías como la inteligencia artificial no se desvíen de los valores que sostienen a las democracias. No se trata solo de dar seguridad jurídica, sino de poner reglas claras y controles que garanticen el respeto por los derechos humanos, la dignidad de las personas, la responsabilidad de quienes diseñan estas herramientas y la equidad en un mundo cada vez más digital.

En una democracia moderna, vivir en un Estado de derecho significa que nadie está por encima de la ley: ni los ciudadanos, ni las empresas, ni el mismo gobierno. Todos, sin excepción, deben respetar unas reglas claras, que han sido creadas de forma pública y que se aplican por igual para todos. Pero no se trata solo de tener leyes escritas; lo verdaderamente importante es que esas normas sostienen la base del sistema democrático y sirven para proteger nuestros derechos más esenciales como personas.

En el contexto del orden constitucional colombiano, la Constitución Política consagra de manera explícita que:

Colombia es un Estado social de derecho, organizado en forma de República unitaria, descentralizada, con autonomía de sus entidades territoriales, democrática, participativa y pluralista, fundada en el respeto de la dignidad humana, en el trabajo y la solidaridad de las personas que la integran y en la prevalencia del interés general. (Constitución Política de Colombia, 1991, art. 1)

Esta definición incorpora elementos tanto del Estado de derecho clásico como del Estado social, creando un marco normativo integral.

La inteligencia artificial, entendida como la habilidad de las máquinas para aprender y resolver problemas de forma similar a los humanos, introduce retos que los marcos legales tradicionales no estaban preparados para enfrentar. A diferencia de otras tecnologías, la IA puede operar de manera autónoma, adaptarse con el tiempo y tomar decisiones sin intervención humana directa. Esto la hace especialmente compleja cuando se aplica en ámbitos sensibles como la justicia, la salud o la educación, donde sus errores pueden generar consecuencias graves. Entre los riesgos más preocupantes están la falta de transparencia en cómo funcionan los algoritmos, la reproducción de prejuicios, la toma de decisiones sin justificación clara y el impacto directo en los derechos de las personas. Por ello, es fundamental que el Estado de Derecho actúe como una guía firme para asegurar que el uso de esta tecnología no debilite los pilares democráticos ni las libertades básicas.

#### **Fortalezas del Enfoque Jurídico-Constitucional**

El Estado de derecho ofrece varios elementos que lo posicionan favorablemente como marco regulatorio para la IA:

## 1. Supremacía de la Dignidad Humana

El principio de dignidad humana, como fundamento del Estado social de derecho, proporciona un criterio axiológico central para evaluar el desarrollo y despliegue de sistemas de IA. La Constitución colombiana establece este principio como base fundamental, lo que requiere que cualquier aplicación tecnológica respete la condición inherente del ser humano.

La dignidad inviolable e intrínseca de cada ser humano constituye la base del sistema universal, indivisible, inalienable, interdependiente e interrelacionado de derechos humanos y libertades fundamentales. Por consiguiente, el respeto, la protección y la promoción de la dignidad humana y de los derechos establecidos por el derecho internacional, en particular el derecho internacional de los derechos humanos, son esenciales a lo largo del ciclo de vida de los sistemas de IA. La dignidad humana tiene que ver con el reconocimiento del valor intrínseco e igual de cada ser humano, con independencia de su raza, color, ascendencia, género, edad, idioma, religión, opiniones políticas, origen nacional, étnico o social, condición económica o social de nacimiento, discapacidad o cualquier otro motivo. (UNESCO, 2021, p. 18)

Cuando hablamos de dignidad humana en el contexto de la inteligencia artificial, es clave recordar que una persona no puede reducirse a un simple conjunto de algoritmos. Las decisiones que afectan derechos fundamentales deben respetar siempre la capacidad de cada individuo para decidir por sí mismo y reconocerlo como un ser humano completo, con historia, emociones y valor propio.

## 2. Control de Constitucionalidad

En Colombia, el sistema de control de constitucionalidad es bastante sólido y es considerado mixto: por un lado, está la Corte Constitucional, que se encarga del control concentrado frente a las leyes, y por otro, todos los jueces pueden hacer control difuso al verificar si existe o no vulneración de derechos fundamentales. Gracias a esa estructura, el país cuenta con herramientas fuertes para revisar si el uso de la inteligencia artificial respeta los valores y principios que defiende la Constitución.

La **Sentencia T-323/24 (2024)** de la Corte Constitucional colombiana representa un precedente significativo al establecer criterios orientadores para el uso de IA en el ámbito judicial. La Corte determinó principios como:

- **Transparencia:** Obligación de evidenciar claramente el uso de herramientas de IA
- **Responsabilidad:** El usuario debe comprender los impactos del uso de estas tecnologías
- **No sustitución de la racionalidad humana:** Imposibilidad ética y jurídica de reemplazar la acción humana
- **Privacidad:** Protección de datos personales y sensibles

## 3. El rol democrático y participativo

El Estado de Derecho no solo es un principio jurídico, sino también un marco político. En sociedades democráticas, las normas que rigen la IA deben ser elaboradas con participación ciudadana, deliberación pública y control institucional, no solo impuestas por actores tecnológicos o económicos. Esto contribuye a la legitimidad normativa y a que las soluciones tecnológicas reflejen el interés general.

Además, la regulación bajo el Estado de Derecho debe contemplar garantías procesales para los ciudadanos, como el derecho a impugnar decisiones automatizadas, ser informado sobre los criterios usados por los algoritmos y obtener reparación en caso de afectaciones indebidas.

Según el análisis de Wisner Glusko (2022), a partir de los principios formulados por el Grupo de Expertos de Alto Nivel sobre IA (CAHAI), la Propuesta de Reglamento de la IA presentada por la Comisión Europea y la Resolución del Parlamento Europeo sobre la IA en la era digital, se desprenden varias orientaciones clave para asegurar que el uso de sistemas de inteligencia artificial no debilite, sino que fortalezca el Estado de Derecho. En primer lugar, es fundamental que la dignidad humana siga siendo el centro de todo desarrollo normativo y aplicación práctica de estas tecnologías, tanto en el ámbito público como privado. Esto significa que las decisiones impulsadas por la IA no deben deshumanizar ni reducir a las personas a simples datos. En segundo lugar, los derechos humanos deben convertirse en el eje rector del diseño y uso de la IA, con especial énfasis en evitar cualquier forma de discriminación derivada del uso de algoritmos o del tratamiento automatizado de información. Debido al escaso entendimiento sobre los efectos reales de la IA en los derechos fundamentales, se vuelve urgente promover una mayor conciencia en todos los sectores, desde los ciudadanos hasta los responsables políticos y empresariales. En tercer lugar, el principio de legalidad debe traducirse en mecanismos concretos de rendición de cuentas y reparación ante posibles daños, lo que incluye garantizar el acceso a procedimientos –judiciales y extrajudiciales– para impugnar decisiones automatizadas que afecten la vida o el patrimonio de las personas. También es crucial la existencia de una supervisión humana real y activa, capaz de evaluar los efectos económicos, sociales, jurídicos y éticos de las decisiones tomadas por sistemas de IA, con posibilidad de intervenir, modificar o anular dichas decisiones si así se requiere. Esta supervisión no solo protege a las personas, sino también a la democracia misma frente a posibles manipulaciones o usos indebidos. A esto se suma la necesidad de que los sistemas sean transparentes, auditables, trazables y explicables, pues solo así se puede ejercer un control efectivo sobre su funcionamiento y evitar que actúen como cajas negras impenetrables. En esa misma línea, el respeto a la privacidad y el adecuado manejo de datos personales durante todo el ciclo de vida de los sistemas de IA es imprescindible, en conformidad con las normas de protección vigentes. Finalmente, se debe considerar el impacto social que estas tecnologías pueden tener a corto, mediano y largo plazo, anticipando riesgos y adoptando medidas de prevención que aseguren que su implementación contribuya de forma real al bienestar colectivo. En conjunto, estas directrices permiten pensar en una IA que funcione como aliada del interés público y no como una amenaza para los principios democráticos.

## **Conclusión**

El **Estado de derecho** no solo puede sino que debe servir como marco regulatorio democrático para la inteligencia artificial. La experiencia colombiana demuestra que es posible adaptar los principios constitucionales fundamentales para abordar los desafíos de la era digital sin comprometer los valores democráticos.

La clave del éxito radica en mantener la **dignidad humana** como principio rector, garantizar la **transparencia algorítmica**, promover la **participación ciudadana** informada y asegurar una **supervisión humana efectiva** en todas las aplicaciones de IA que afecten derechos fundamentales.

Colombia está bien posicionada para convertirse en un **modelo regional** de regulación democrática de IA, combinando innovación tecnológica con protección robusta de derechos humanos. Este enfoque integral requiere un compromiso sostenido de todas las ramas del poder público, la sociedad civil y el sector privado para construir un futuro donde la tecnología sirva al florecimiento humano y al fortalecimiento de la democracia.

### **4.3 Recomendaciones para una política pública en Colombia**

Según el Departamento Nacional de Planeación (2025), la inteligencia artificial se está convirtiendo en una tecnología con impacto transversal en diversas áreas sociales y económicas. Su avance global ha sido acelerado gracias al acceso masivo a datos y a nuevas tecnologías complementarias. En Colombia, se reconoce su potencial para enfrentar retos estructurales como la pobreza, la inseguridad alimentaria y la transición hacia una economía basada en el conocimiento. A partir del Documento CONPES 3975 de 2019, se han fortalecido condiciones institucionales y técnicas para integrar herramientas digitales. Sin embargo, el país aún enfrenta obstáculos importantes para investigar, desarrollar e implementar sistemas de IA de manera ética, sostenible y orientada al bienestar colectivo.

Según Sánchez Bravo (2022), a medida que las expectativas de la población se multiplican, también crece la dimensión y la complejidad de las funciones del sector público. Hoy en día, los ciudadanos no se conforman con recibir servicios: exigen transparencia en los procesos de toma de decisiones e incluso desean involucrarse activamente en ellos. De este modo, la administración, que tradicionalmente lidera las transformaciones sociales, ve ahora a sus usuarios reclamando un papel protagonista en esas mismas transformaciones, y son las tecnologías emergentes las que facilitan estas nuevas formas de participación ciudadana.

En Colombia, el avance de la inteligencia artificial ha llevado a que el Estado impulse iniciativas importantes, como el Documento CONPES 4144 de 2025, que traza el camino hacia una política nacional sobre el tema. Aunque esa estrategia toca muchos sectores, es clave aterrizar sus principios al campo jurídico, porque cuando la IA entra en juego en decisiones públicas o judiciales, lo que está en riesgo son derechos fundamentales. La idea es construir una política sólida, ética y útil, que permita aprovechar la IA dentro del derecho sin poner en juego la confianza ciudadana ni las garantías que debe proteger el Estado. (Departamento Nacional de Planeación, 2025)

#### **1. Fortalecimiento de una gobernanza institucional especializada en IA y derecho.**

Una política pública robusta debe contemplar la creación de una autoridad reguladora interinstitucional especializada en IA y derecho, con capacidad jurídica, técnica y

presupuestal. Esta entidad deberá coordinar con el Ministerio de Tecnologías de la Información y las Comunicaciones (MinTIC), el Ministerio de Justicia y el Consejo Superior de la Judicatura, y tener funciones específicas como:

- Emitir lineamientos vinculantes sobre el uso ético y legal de IA en el sistema judicial y administrativo.
- Supervisar el cumplimiento de los principios éticos y de derechos fundamentales en sistemas algorítmicos utilizados en decisiones jurídicas o administrativas.
- Establecer estándares técnicos y legales para la auditoría de algoritmos, con criterios de transparencia, explicabilidad y no discriminación.

El Departamento Nacional de Planeación (2025) anunció que distintas entidades del Gobierno —como el DAPRE, el Ministerio TIC, el Ministerio de Ciencia y la CRC— trabajarán en 2025 para crear normas, guías éticas y buenas prácticas sobre el uso de inteligencia artificial en el sector público. También diseñarán orientaciones dirigidas al sector privado, la academia y centros de investigación. Más adelante, entre 2028 y 2030, se planea extender esta estrategia a los gobiernos locales, buscando alinear sus planes de desarrollo con principios éticos y actualizar la política nacional de IA de forma colaborativa.

## **2. Diseño de un marco normativo gradual y sectorialmente adaptado**

Si bien un marco ético es esencial como base, resulta indispensable avanzar hacia una regulación jurídica progresiva, que combine normas generales con desarrollos sectoriales específicos. Se recomienda:

- Reformar y/o adicionar leyes existentes, como el Código Contencioso Administrativo y la Ley Estatutaria de Protección de Datos Personales (Ley 1581 de 2012), para incluir el tratamiento automatizado de datos y decisiones automatizadas en el sector público.
- Formular una Ley de inteligencia artificial aplicada al sector público, inspirada en los principios de legalidad algorítmica, debido proceso digital, y responsabilidad diferenciada por el uso de sistemas automatizados.
- Establecer un marco para sandbox regulatorios jurídicos, donde se permita la experimentación controlada de IA en procesos judiciales o administrativos, sin renunciar al cumplimiento de estándares constitucionales.

Como parte de una visión estratégica orientada a consolidar un entorno digital más justo y garantista, Sánchez Bravo (2022) plantea como primer objetivo específico la creación de una Carta de Derechos Digitales. Esta iniciativa busca establecer con claridad los derechos tanto de los ciudadanos como de los actores que operan en el entorno digital, ofreciendo un marco interpretativo sólido para ciertos principios jurídicos y reconociendo, al mismo tiempo, derechos emergentes en respuesta a los desafíos propios de las nuevas tecnologías. Además de brindar mayor seguridad jurídica, esta Carta pretende ser un instrumento eficaz para cerrar las brechas digitales que, en los últimos años, se han profundizado a causa de desigualdades sociales, económicas, de género, edad o ubicación geográfica.

### **3. Implementación de auditorías éticas y jurídicas de algoritmos**

Una política pública efectiva debe institucionalizar mecanismos permanentes de auditoría algorítmica en todos los sistemas que se utilicen para toma de decisiones en el ámbito legal. Para ello, se recomienda:

- Establecer un registro público de algoritmos aplicados al derecho, donde consten los sistemas utilizados por entidades públicas, su finalidad, datos entrenados y procedimientos de validación.
- Establecer una metodología nacional para evaluaciones de impacto algorítmico (AIA), obligatoria para todo sistema que tenga efectos jurídicos sobre derechos fundamentales, inspirada en la Evaluación de Impacto en Protección de Datos (DPIA) del Reglamento General de Protección de Datos de la UE.

Desde una perspectiva orientada al fortalecimiento de los mecanismos de control y responsabilidad en el uso de tecnologías avanzadas, Guío Español (2020) destaca la auditoría de algoritmos como una propuesta clave promovida por organizaciones de la sociedad civil y respaldada, de manera notable, por el gobierno francés. En el marco de su estrategia nacional de inteligencia artificial, Francia ha contemplado la creación de una plataforma pública encargada de examinar los algoritmos utilizados por el Estado, aunque dicha propuesta aún se encuentra en discusión parlamentaria. Esta forma de supervisión técnica busca identificar comportamientos algorítmicos que puedan resultar socialmente perjudiciales, como la discriminación por razón de género, a través de métodos cuantitativos y pruebas experimentales. No obstante, hasta ahora, este tipo de auditorías han sido iniciativas puntuales, realizadas por actores externos —como periodistas o académicos— que frecuentemente se enfrentan a obstáculos legales al investigar plataformas privadas sin autorización, contraviniendo los términos de uso. Frente a esta realidad, se vuelve imperativo establecer mecanismos de auditoría que sean sostenibles, jurídicamente viables y sistemáticos. Además, se subraya la necesidad de avanzar hacia el diseño ético de algoritmos, un enfoque cada vez más relevante dentro de la comunidad investigativa, que propone identificar desde el inicio los riesgos éticos asociados a un sistema algorítmico y crear soluciones que los eviten estructuralmente, en lugar de corregirlos a posteriori.

### **5. Formación y apropiación del conocimiento ético-jurídico en IA**

Para garantizar una aplicación efectiva y legítima de la IA en el campo del derecho, es necesario promover una cultura ética y crítica en todos los niveles del ecosistema jurídico. Se recomienda:

- Incluir contenidos de ética algorítmica, derechos digitales y legalidad tecnológica en programas universitarios de derecho, administración pública, ingeniería y ciencias de datos.
- Fomentar la formación continua de jueces, fiscales, conciliadores, defensores públicos y servidores administrativos sobre los usos y límites de la IA en su quehacer jurídico.

- Impulsar estrategias de apropiación social del conocimiento, dirigidas a la ciudadanía, con el fin de que puedan comprender e impugnar decisiones tomadas o influenciadas por IA.

El desarrollo de competencias digitales en la ciudadanía se ha convertido en una necesidad fundamental para permitir una participación activa, segura y eficiente en la sociedad digital, facilitando la comunicación, el acceso a la información, las transacciones económicas y la interacción con las administraciones públicas. Más allá de estas habilidades básicas, se requieren competencias digitales avanzadas para desempeñar actividades profesionales complejas, especialmente en un entorno laboral cada vez más digitalizado, donde numerosos sectores demandan conocimientos técnicos específicos relacionados con herramientas digitales sofisticadas. En contraste, los trabajadores con habilidades digitales limitadas enfrentan un mayor riesgo de exclusión laboral, profundizándose así la brecha digital. Paralelamente, emergen sectores estratégicos orientados al diseño, implementación y mantenimiento de sistemas digitales, que representan oportunidades de empleo altamente cualificado y bien remunerado. En este contexto, resulta imprescindible fortalecer el sistema educativo y promover la formación continua a lo largo de la vida, en línea con los lineamientos del Plan Europeo de Educación Digital. (Sánchez Bravo, 2022).

Según el Departamento Nacional de Planeación (2025), entre 2025 y 2028 el DANE, con apoyo de varias entidades del Gobierno, incorporará en sus estadísticas aspectos relacionados con la apropiación social de la inteligencia artificial, bajo un enfoque sensible a las diferencias culturales y territoriales. Paralelamente, se lanzará una estrategia nacional liderada por el Ministerio de Ciencia para acercar la IA a la ciudadanía, con campañas pedagógicas, acciones de divulgación y herramientas para medir su comprensión en distintos sectores. En este mismo marco, entre 2025 y 2027, el DNP y el Ministerio TIC desarrollarán una guía práctica para fortalecer los procesos educativos y formativos en torno al uso y análisis de esta tecnología. Por su parte, el Ministerio de Trabajo investigará cómo la transformación digital está impactando el empleo y, en alianza con el SENA y otras entidades, implementará entre 2025 y 2030 una estrategia para mitigar los efectos de la automatización, que incluirá programas de reconversión laboral y la promoción de derechos en el entorno digital.

## **6. Desarrollo de capacidades y talento humano especializado en IA jurídica**

La política pública nacional debe contemplar inversiones específicas para cerrar la brecha de talento en el cruce entre derecho y tecnología. Se recomienda:

- Financiar becas e incentivos para la formación de abogados, jueces y funcionarios públicos en ciencia de datos e IA aplicada al derecho.
- Estimular programas interdisciplinarios de posgrado en universidades públicas que articulen derecho, filosofía, informática y ética.
- Crear centros de innovación legal donde se desarrollen soluciones tecnológicas abiertas y auditables, desde el Estado.

A nivel internacional, algunas iniciativas educativas están abriendo camino en la integración crítica de la inteligencia artificial en las aulas. Un caso ejemplar es el proyecto del MIT, que, según Guío Español (2020), promueve la enseñanza de la IA en el

bachillerato desde una perspectiva ética. Más que enfocarse solo en lo técnico, esta propuesta busca que los estudiantes comprendan que los algoritmos son construcciones humanas, con sesgos y consecuencias sociales, y que pueden —y deben— ser cuestionados y rediseñados. A través de un currículo de código abierto, el programa combina aprendizajes sobre el funcionamiento de la IA con ejercicios de análisis, reflexión y rediseño de sistemas, incentivando a los jóvenes a convertirse en creadores críticos de tecnología y no solo en usuarios pasivos.

Con una mirada a largo plazo, el Departamento Nacional de Planeación (2025) propone una estrategia educativa que, entre 2025 y 2030, impulse la transformación digital en las escuelas del país. En el corazón de este plan está el programa *Computadores para Educar*, que dotará a sedes oficiales con equipos, contenidos pedagógicos y herramientas alineadas con la política educativa y la ampliación de la conectividad escolar. A la par, el Ministerio de Educación incluirá un componente sobre inteligencia artificial dentro de su política de uso seguro de las TIC, abordando temas técnicos, éticos y formativos. Esta línea se reforzará con iniciativas conjuntas con el Ministerio de Ciencia para promover el uso responsable de la IA en las aulas y fortalecer las habilidades digitales de docentes y estudiantes. Además, el Ministerio TIC, junto al de Educación, liderará la “Misión para la Transformación Digital de la Educación”, un modelo integral para integrar tecnologías emergentes en todos los niveles del sistema educativo. Finalmente, entre 2025 y 2026, se trabajará en actualizar las pruebas de competencias digitales —como SABER— con base en la experiencia de la Prueba TI, adecuándolas a los retos de la inteligencia artificial en la formación de ciudadanos del siglo XXI.

## **7. Participación multisectorial y construcción colectiva de estándares**

Finalmente, es fundamental que la política pública sobre IA en el derecho colombiano se construya a partir de una gobernanza participativa y una perspectiva diferencial, territorial y de género. Para ello:

- Promover espacios de consulta vinculante con organizaciones de la sociedad civil, academia, comunidades indígenas, afrocolombianas y campesinas, que puedan verse afectadas por decisiones algorítmicas.
- Fomentar la co-creación de estándares éticos y jurídicos junto a actores nacionales e internacionales, como la OCDE, UNESCO y BID, desde una mirada crítica y contextualizada.
- Incorporar una evaluación interseccional de riesgos algorítmicos, que considere desigualdades estructurales en su impacto.

Para enfrentar los desafíos de la inteligencia artificial, Colombia trazó una hoja de ruta ambiciosa hacia una infraestructura digital ética, segura y resiliente. Entre 2025 y 2030, el gobierno modernizará sus marcos de seguridad, actualizará modelos de protección de datos, y creará herramientas para guiar tanto al sector público como al privado en el uso responsable de la IA. La estrategia incluye desde la elaboración de catálogos institucionales y políticas de compras públicas orientadas a la ciberseguridad, hasta campañas contra la desinformación algorítmica y capacitaciones para proteger el entorno digital del Estado.

Todo este esfuerzo, liderado por entidades como el Ministerio TIC, el DNP y el DAPRE, apunta a lograr una gobernanza tecnológica que cuide los derechos sin frenar la innovación (Departamento Nacional de Planeación, 2025).

## **5. conclusion.**

Colombia está atravesando un momento decisivo en su relación con la inteligencia artificial dentro del sistema jurídico. Esta monografía ha puesto sobre la mesa una tensión constante: por un lado, el entusiasmo por la innovación tecnológica; por otro, la urgencia de blindar los derechos fundamentales. Aunque hoy existen herramientas como PROMETEA, PRISMA y PRETORIA funcionando en distintos frentes del aparato judicial, el marco legal sigue siendo escaso y fragmentado. Esta falta de regulación deja a la ciudadanía expuesta a riesgos muy reales, como decisiones opacas, sesgos automatizados y un debilitamiento de las garantías procesales. La Sentencia T-323 de 2024 marca un punto de referencia importante en la jurisprudencia nacional, pero también evidencia lo limitado que resulta depender solo de los jueces: lo que se necesita con urgencia es una normativa sólida, integral y anticipada.

A lo largo del análisis quedó claro que el uso de inteligencia artificial en la justicia colombiana ha generado efectos mixtos. Por un lado, PRETORIA ha contribuido positivamente a mejorar la gestión de expedientes y a reducir tiempos en los procesos. Por otro lado, casos como el de PRISMA generan alertas sobre sesgos heredados y decisiones automatizadas que afectan directamente la libertad personal. Este contraste encarna el dilema central del estudio: cómo encontrar el equilibrio justo entre impulsar la tecnología y salvaguardar los derechos fundamentales. No se trata solo de un debate técnico, sino ético y estructural, especialmente en un país con desigualdades tan arraigadas. Las lecciones internacionales, donde se han documentado fallos graves por sesgos raciales en tecnologías como el reconocimiento facial o la predicción criminal, son una advertencia clara de lo que Colombia debe evitar con una regulación seria y preventiva.

La propuesta regulatoria presentada aquí parte del Estado de Derecho y se nutre de estándares globales —como los formulados por la OCDE, la UNESCO y la Unión Europea—, adaptados al contexto constitucional colombiano. Las recomendaciones incluyen crear una institucionalidad especializada en gobernanza algorítmica, exigir auditorías éticas a los sistemas de alto riesgo, invertir en capacidades técnicas y jurídicas, y promover activamente la participación ciudadana en el diseño de marcos éticos. La idea es que la regulación no sea un obstáculo para la innovación, sino una herramienta que permita avanzar de manera responsable, asegurando transparencia, control humano y explicaciones claras en las decisiones automatizadas que afectan la vida de las personas.

Finalmente, esta investigación plantea que el verdadero éxito de la transformación digital del derecho en Colombia dependerá de qué tan bien logremos construir un ecosistema colaborativo. No basta con buenas intenciones o soluciones técnicas; se necesita una alianza real entre Estado, academia, empresas y ciudadanía. Solo así podremos aprovechar el potencial democratizador de la inteligencia artificial para mejorar el acceso a la justicia, sin dejar de lado los principios que sostienen el Estado de Derecho. Colombia tiene una oportunidad única de liderar con el ejemplo en América Latina y demostrar que sí es posible innovar sin renunciar a la dignidad, la equidad y los derechos humanos.

## Referencias

- Blog de Derecho de los Negocios – Universidad Externado de Colombia. (2019, 12 de marzo). ¡Promete, Prometea! Inteligencia artificial en la Corte Constitucional. <https://dernegocios.uexternado.edu.co/comercio-electronico/promete-prometea-inteligencia-artificial-en-la-corte-constitucional/>
- Cáceres, E. (2006). Inteligencia artificial, derecho y E JUSTICE (El Proyecto II] CONACYT). Boletín Mexicano de Derecho Comparado, 39(116), 593–611. [https://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S0041-86332006000200012](https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0041-86332006000200012)
- Carvajal Martínez, A. (2010). *Los periodistas y el derecho a la información en Colombia* (2ª ed.). Editorial Copy Net S.A.

- Cámara de Representantes de Colombia. (2024, 6 de agosto). Proyecto de Ley Estatutaria "Por la cual se define y regula la inteligencia artificial se ajusta a estándares de derechos humanos, se establecen límites frente a su desarrollo, uso e implementación se modifica parcialmente la ley 1581 de 2012 y se dictan otras disposiciones" [Documento inédito].
- Colombia. (2008). Ley 1266 de 2008: por la cual se dictan las disposiciones generales del hábeas data y se regula el manejo de la información contenida en bases de datos personales, en especial la financiera, crediticia, comercial, de servicios y la proveniente de terceros países y se dictan otras disposiciones. Artículo 4. [http://www.secretariassenado.gov.co/senado/basedoc/ley\\_1266\\_2008.html](http://www.secretariassenado.gov.co/senado/basedoc/ley_1266_2008.html)
- Colombia. (2012). Ley 1581 de 2012: Por la cual se dictan disposiciones generales para la protección de datos personales. Artículo 4. [http://www.secretariassenado.gov.co/senado/basedoc/ley\\_1581\\_2012.html](http://www.secretariassenado.gov.co/senado/basedoc/ley_1581_2012.html)
- Constitución Política de Colombia. (1991). Constitución Política de Colombia -1991. [http://www.secretariassenado.gov.co/senado/basedoc/constitucion\\_politica\\_1991.html](http://www.secretariassenado.gov.co/senado/basedoc/constitucion_politica_1991.html)
- Corte Constitucional. (2024, 12 de agosto). Sentencia T-323/24 (Juan Carlos Cortés González, M. P.). <https://www.corteconstitucional.gov.co/relatoria/2024/t-323-24.htm>
- Corte Constitucional. (2011, 6 de julio). *Sentencia C-748/11* (Luis Ernesto Vargas Silva, M. P.). <https://www.corteconstitucional.gov.co/relatoria/2011/c-748-11.htm>
- Departamento Nacional de Planeación [DNP]. (2019, 8 de noviembre). Política nacional para la transformación digital e inteligencia artificial (Documento CONPES 3975). <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/3975.pdf>
- Departamento Nacional de Planeación [DNP]. (2025). Política nacional de inteligencia artificial: Documento CONPES 4144. Consejo Nacional de Política Económica y Social. <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/4144.pdf>
- Galindo, F. (2019). ¿Inteligencia Artificial y Derecho? Sí, pero ¿cómo? Revista Democracia Digital e Governo Eletrônico, 1(18), 36–57. <https://buscalegis.ufsc.br/revistas/index.php/observatoriodoegov/article/view/310/206>
- García San José, D. (2022). Significado y alcance de los valores de la Carta de Naciones Unidas en la regulación internacional de la inteligencia artificial (IA). En F. H.

- Llano Alonso (Dir.), J. Garrido Martín & R. Valdivia Jiménez (Coords.), *Inteligencia artificial y filosofía del derecho* (pp. 431–448). Ediciones Laborum.
- Guío Español, A. (2020). Marco ético para la inteligencia artificial en Colombia: Documento para discusión. CAF - Banco de Desarrollo de América Latina. <https://www.usergioarboleda.edu.co/wp-content/uploads/2021/11/Marco-etico-para-la-inteligencia-artificial-en-Colombia-Maestria-en-Inteligencia-artificial.pdf>
- Instituto Colombiano de Derecho Procesal [ICDP]. (2024). Análisis y recomendaciones al Proyecto de Ley Estatutaria No. 154 de 2024: Hacia una regulación ética y pro-innovación de la inteligencia artificial en Colombia [Informe técnico]. [https://www.camara.gov.co/sites/default/files/2024-11/An%C3%A1lisis%20y%20Recomendaciones%20al%20Proyecto%20de%20Ley%20Estatutaria%20No.%20154%20de%202024\\_%20Hacia%20una%20Regulaci%C3%B3n%20%C3%89tica%20y%20Pro-Innovaci%C3%B3n%20de%20la%20Inteligencia%20Artificial%20en%20Colombia.pdf](https://www.camara.gov.co/sites/default/files/2024-11/An%C3%A1lisis%20y%20Recomendaciones%20al%20Proyecto%20de%20Ley%20Estatutaria%20No.%20154%20de%202024_%20Hacia%20una%20Regulaci%C3%B3n%20%C3%89tica%20y%20Pro-Innovaci%C3%B3n%20de%20la%20Inteligencia%20Artificial%20en%20Colombia.pdf)
- Ley 1431, 2009, art. 6. República de Colombia. [http://www.secretariassenado.gov.co/senado/basedoc/ley\\_1341\\_2009.html](http://www.secretariassenado.gov.co/senado/basedoc/ley_1341_2009.html)
- Ministerio de Educación Nacional. (2025, 7 de marzo). Educación superior inicia el camino para integrar la inteligencia artificial en el país. <https://www.mineducacion.gov.co/portal/salaprensa/Comunicados/423752:Educacion-superior-inicia-el-camino-para-integrar-la-inteligencia-artificial-en-el-pais>
- Montes, J. (2025, 7 de abril). Una juez colombiana lidera la transformación del sistema judicial con Copilot. Microsoft LATAM. <https://news.microsoft.com/source/latam/features/ia/una-juez-colombiana-lidera-la-transformacion-del-sistema-judicial-con-copilot/>
- Morales Cáceres, A. (2021). El impacto de la inteligencia artificial en el Derecho. *Advocatus*, 39, 39–71. <https://doi.org/10.26439/advocatus2021.n39.5117>
- Nieto Martínez, L. (2025, 6 de julio). Inteligencia artificial en Colombia: entre la autorregulación y la regulación foránea (I). *Ámbito Jurídico*. <https://ambitojuridico.com/noticias/derecho-publico/mercantil-propiedad-intelectual-y-arbitraje/inteligencia-artificial-en>
- Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2019). *Recomendación del Consejo sobre inteligencia artificial*. OECD Legal Instruments, OECD/LEGAL/0449. <https://www.oecd.org/en/topics/ai-principles.html>

Pérez-Ugena, M. (2024). La inteligencia artificial: definición, regulación y riesgos para los derechos fundamentales. *Estudios de Deusto*, 72(1), 307-338.  
<https://doi.org/10.18543/ed.3108>

Sánchez Bravo, Á. (2022). España Digital 2025. Estrategia Nacional de Inteligencia Artificial. En F. H. Llano Alonso (Dir.), J. Garrido Martín & R. Valdivia Jiménez (Coords.), *Inteligencia artificial y filosofía del derecho* (pp. 501–527). Ediciones Laborum.

Suárez Manrique, W. Y., De León Vargas, G. I., & Monsalve Pinto, J. J. (2023, 15 de diciembre). Un análisis sobre la aplicabilidad de la inteligencia artificial en el derecho colombiano. *Prolegómenos*, 26(52), 11–23.  
<https://doi.org/10.18359/prole.5625>

Tirant lo Blanch. (2024, 21 de octubre). Sof IA: La IA para abogados más eficiente. <https://prime.tirant.com/co/actualidad-prime/sof-ia-la-ia-para-abogados-mas-eficiente/>

Torres Abril, J. S., Silva Vásquez, Z. Y., & Gómez Simijaca, V. (2022). Herramientas de la inteligencia artificial dentro del sistema judicial colombiano: Estudio del caso PretorIA y PRISMA. *Revista Principia Iuris*, 19(40), 47–66.

UNESCO. (2021, 23 de noviembre). Recomendación sobre la ética de la inteligencia artificial.  
[https://unesdoc.unesco.org/ark:/48223/pf0000381137\\_spa/PDF/381137spa.pdf.m  
ulti](https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa/PDF/381137spa.pdf.multi)

Unión Europea. (2024). *Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo relativo a la inteligencia artificial (AI Act)*. Diario Oficial de la Unión Europea, L, 2024/1689. [https://eur-lex.europa.eu/legal-content/ES/TXT/?  
uri=CELEX%3A32024R1689](https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32024R1689)

Wisner Glusko, D. C. (2022). Breves reflexiones sobre la importancia del Estado de derecho en el desarrollo del marco legal sobre los sistemas de inteligencia artificial en la Unión Europea. En F. H. Llano Alonso (Dir.), J. Garrido Martín & R. Valdivia Jiménez (Coords.), *Inteligencia artificial y filosofía del derecho* (pp. 529–545). Ediciones Laborum.