



Guía de uso ético y responsable de la IA para defensores públicos

Lineamientos prácticos con enfoque
de derechos humanos

www.defensoria.gov.co



#BuenFuturoHoy

DEFENSORÍA DEL PUEBLO

IRIS MARÍN ORTIZ
Defensora del Pueblo

PAULA CRISTINA APONTE URDANETA
Vicedefensora del Pueblo

MARGARITA BARRAQUER SOURDIS
Secretaria General

CAROLINA VERGEL TOVAR
Directora Nacional de Promoción y Divulgación de DD. HH.

Autores

Dirección Nacional de Defensoría Pública
Defensoría Delegada para la Protección de Derechos en
Ambientes Digitales y Libertad de Expresión

Corrección de estilo

Alejandra Restrepo Franco

Diseño y diagramación

Felipe Casas Santofimio

Obra de distribución gratuita.
El presente texto se puede reproducir,
fotocopiar o replicar, total o parcialmente,
citando la fuente.

2026

© Defensoría del Pueblo de Colombia
Calle 55 N.º 10-32
Bogotá, D.C.
www.defensoria.gov.co

#BuenFuturoHoy

Tabla de contenido

I. Introducción 3

Nota aclaratoria:	3
¿Por qué se necesita una guía sobre el uso ético y responsable de la inteligencia artificial para las y los defensores públicos?	3
¿Cuál es la relevancia de esta guía en el contexto actual de la defensa pública en Colombia?	3
¿Cuál es el propósito de esta guía?	4
Resultados de aprendizaje	4
¿Qué significa hacer un uso ético y responsable de la IA en la labor defensorial?	5
¿Qué encontrarán en esta guía?	6
¿Esta guía sustituye otras normas?	6

II. ¿Qué entendemos por...? 7

Modelo de IA	8
Herramienta de IA generativa	8
Resultado o salida	8
Anonimización y riesgo de reidentificación	8
Supervisión humana significativa	8

III. Principios básicos para un uso ético y responsable de la IA en la labor defensorial 9

Primacía de los derechos humanos y los derechos fundamentales	9
Apoyo y no sustitución	10
Supervisión humana	10
Criterio profesional y responsabilidad personal	10
Proporcionalidad	11

Tabla de contenido

Uso de la IA desde una perspectiva centrada en las personas	11
Enfoque diferencial e interseccional	11
Confidencialidad y protección de la información	12
Transparencia	12

IV. Usos comunes de la inteligencia artificial en la labor defensorial

13

Elaboración de resúmenes de expedientes o documentos jurídicos	14
Búsqueda y análisis de jurisprudencia o doctrina	16
Redacción de oficios, solicitudes, escritos jurídicos o comparación de textos	18
Preparación de audiencias y alegatos iniciales	20
Escenarios diferenciales de aplicación	22

V. Transparencia y escalas de uso

23

Semáforo de usos permitidos o no permitidos	24
Antes de usar IA: ruta de decisión	24
Ciclo de uso seguro y responsable	25
Prohibiciones expresas	27
Etiquetas y reconocimiento del uso de IA	27
Kit de supervivencia para un uso seguro y responsable de la IA	29
Tips concretos para la creación de <i>prompts</i>	32
Copilot y otras herramientas de IA disponibles en los servicios institucionales	36



Tabla de contenido



VI. Anexo práctico: <i>prompts</i> seguros y lista de chequeo	37
Ejemplos según nivel de riesgo	39
Lista de chequeo previa	40
VII. Gobernanza, actualización y control de versiones	41
VIII. Recursos adicionales	42
IX. Referencias	44
Normativa y política pública colombiana	44
Estándares internacionales y accesibilidad	44
Literatura académica	45

Presentación

Estimadas defensoras públicas y estimados defensores públicos:

En los últimos años, diferentes tipos de tecnologías —en particular, la inteligencia artificial (IA)— han irrumpido en nuestras labores diarias. La IA representa una oportunidad y también un desafío para el sistema de justicia y para la labor defensorial. Sabemos que, en determinados casos, puede convertirse en una herramienta valiosa para apoyar algunas de las tareas que ustedes realizan. Sin embargo, su uso debe caracterizarse ante todo por ser ético, responsable y consciente de las limitaciones que aún tiene este tipo de tecnología, un uso que complementa, pero nunca sustituya el juicio humano.

En coherencia con el firme compromiso de la Defensoría del Pueblo con la promoción y protección de los derechos humanos y en consonancia con los valores que orientan nuestro Decálogo Institucional, me complace presentarles la Guía de uso ético y responsable de la inteligencia artificial para las y los defensores públicos. Esta iniciativa se inscribe también en la visión del Derecho al Buen Futuro, que impulsa el uso justo, transparente y humano de las tecnologías para fortalecer la dignidad y las oportunidades de las generaciones presentes y futuras.

Esta Guía ha sido elaborada pensando en ustedes. Su propósito es ofrecer una orientación clara, prudente y accesible para que quienes decidan incorporar la IA como apoyo en el cumplimiento de sus funciones, puedan hacerlo garantizando la pulcritud, el rigor y la humanidad que distinguen el servicio que prestan.

La ética, la independencia, la empatía y la transparencia —principios recogidos en el Decálogo de la Defensoría del Pueblo— inspiran este esfuerzo institucional por fortalecer la confianza ciudadana y asegurar que el uso de la tecnología esté siempre al servicio de las personas, nunca en su contra.

La Defensoría del Pueblo reconoce y valora profundamente el trabajo que realizan, día a día, las y los defensores públicos en todo el territorio nacional. Su labor, muchas veces desarrollada en contextos de alta complejidad, presión institucional y limitaciones materiales, constituye una garantía esencial para el acceso a la justicia y la protección efectiva de los derechos fundamentales de las personas en mayor situación de vulnerabilidad. El compromiso, la vocación de servicio y el criterio profesional con el que ejercen su función son el fundamento sobre el cual se construyen la legitimidad y la confianza en la defensa pública.

Este énfasis responde a los retos sustantivos que plantea el uso de sistemas de inteligencia artificial en el ejercicio de la función defensorial. La IA puede reproducir sesgos existentes, amplificar desigualdades estructurales, afectar la autonomía profesional y generar opacidades que dificulten la rendición de cuentas. Su uso acrítico también puede comprometer la protección de datos personales, la confidencialidad, el secreto profesional y la relación de confianza con las personas usuarias. Por ello, la adopción de estas herramientas exige supervisión humana significativa, transparencia y salvaguardas proporcionales al riesgo.

Presentación

Los principios aquí contenidos se inspiran, entre otros, en los estándares de la UNESCO y en las directrices del Consejo Superior de la Judicatura, adaptados al contexto de la defensa pública colombiana. De esta manera, la Guía busca traducir los valores de la Defensoría en prácticas concretas de innovación ética, que promuevan una justicia más cercana, eficiente y respetuosa de los derechos digitales de la ciudadanía.

Esperamos que esta Guía se convierta en una herramienta útil para fortalecer su misión en la defensa de quienes más lo necesitan. Recuerden que, si bien la tecnología puede ser un gran apoyo, nunca debe reemplazar el criterio profesional, la empatía ni el compromiso humano que constituyen el corazón de su labor defensorial. Su trabajo, guiado por los principios del Buen Futuro, seguirá siendo un pilar esencial para construir una sociedad más justa, solidaria y humana.

El Derecho al Buen Futuro nos invita a mirar la tecnología no solo como un instrumento del presente, sino como una responsabilidad hacia el porvenir. Significa garantizar que los avances tecnológicos estén alineados con los valores que dan sentido a la defensa de los derechos humanos: la dignidad, la equidad, la sostenibilidad y la libertad. Incorporar este principio en nuestra labor diaria implica asegurar que cada decisión institucional contribuya a dejar un legado positivo, donde la innovación se traduzca en bienestar colectivo y justicia intergeneracional.

Con esta Guía, la Defensoría del Pueblo reafirma su compromiso de liderar una transformación digital con enfoque de derechos humanos. Ustedes, defensoras y defensores públicos, son el rostro humano de esa transformación. Con su criterio, su sensibilidad y su compromiso ético son quienes hacen posible que la tecnología sirva a las personas y no al revés. Juntos, seguiremos construyendo un futuro digital más justo, más humano y verdaderamente digno para todas y todos.

Este compromiso no se agota en la inteligencia artificial. Se extiende al manejo responsable de información reservada, la protección de datos personales, la seguridad digital y la adopción ética de nuevas tecnologías. En esa medida, esta Guía constituye un primer paso dentro de un proceso institucional más amplio y continuo que deberá traducirse en formación, protocolos, herramientas autorizadas, mecanismos de seguimiento y actualizaciones periódicas para acompañar a las y los defensores públicos.

Con aprecio y reconocimiento por su servicio,

IRIS MARÍN ORTIZ
Defensora del Pueblo



I. Introducción

Nota aclaratoria:

Este documento incorpora lineamientos en su versión inicial, sujetos a revisión ordinaria anual y a actualización extraordinaria cuando existan cambios normativos, jurisprudenciales, tecnológicos, institucionales o incidentes de seguridad relevantes. La Defensoría del Pueblo mantendrá un proceso abierto para recibir observaciones y evidencias de aplicación. Cada modificación deberá registrarse mediante control de versiones, con indicación de fecha, cambio realizado, fundamento y dependencias participantes. Esta naturaleza evolutiva no reduce la vigencia de los deberes legales, éticos, disciplinarios y de confidencialidad aplicables a la defensa pública.

¿Por qué se necesita una guía sobre el uso ético y responsable de la inteligencia artificial para las y los defensores públicos?

La inteligencia artificial (IA) ya está transformando el ejercicio del derecho, incluida la labor de las y los defensores públicos en Colombia. El uso de la IA sin una orientación clara puede generar riesgos éticos, técnicos, jurídicos y posibles vulneraciones a los derechos de las personas que buscan el acompañamiento de la Defensoría del Pueblo. De ahí que, esta guía pretende prevenir esos riesgos y ofrecer criterios claros para que, cuando se utilice, la IA sea un apoyo complementario a la labor defensorial, siempre bajo la supervisión humana, la ética profesional y el respeto por la dignidad de cada persona.



¿Cuál es la relevancia de esta guía en el contexto actual de la defensa pública en Colombia?

La relevancia de esta guía radica en que brinda principios claros, orientaciones y ejemplos concretos de buenas prácticas y herramientas para que las y los defensores públicos que decidan incorporar sistemas de IA para apoyar su labor lo hagan de manera ética, responsable, segura y consciente de sus limitaciones y riesgos. Adicionalmente, esta guía promueve un uso de la IA que respalde el trabajo defensorial sin reemplazar el criterio jurídico ni el compromiso humano que caracterizan la defensa pública.

¿Cuál es el propósito de esta guía?

Este documento busca orientar a las y los defensores públicos para que, si así deciden hacerlo, utilicen herramientas de IA como apoyo a su labor defensorial de forma ética y responsable, conscientes de las potencialidades y limitaciones de este tipo de herramientas. Adicionalmente, esta guía pretende:



1. Brindar criterios prácticos para evaluar en qué casos podrían o no hacer uso de herramientas de IA como apoyo a su labor.



2. Prevenir posibles afectaciones a los derechos de las personas que buscan el acompañamiento de la Defensoría del Pueblo, cuando se haga uso de este tipo de sistemas de IA.



3. Fomentar un aprendizaje continuo sobre las oportunidades y riesgos de la IA en la labor defensorial.



4. Promover el uso transparente, responsable y trazable de las herramientas tecnológicas, para garantizar la rendición de cuentas y la preservación del criterio profesional en todas las actuaciones defensoriales.



5. Integrar el enfoque de derechos humanos, el Decálogo Institucional y el Derecho al Buen Futuro como principios orientadores de toda práctica digital en la defensa pública y asegurar que la innovación tecnológica contribuya al fortalecimiento de la dignidad, la equidad y la justicia social.



Resultados de aprendizaje

Al finalizar la lectura y aplicación de esta Guía, las y los defensores públicos estarán en capacidad de:

1. Distinguir los usos permitidos, condicionados y no permitidos de las herramientas de IA en la labor defensorial.

2. Aplicar supervisión humana, verificación jurídica y consulta de fuentes oficiales a todo resultado generado con apoyo de IA.

3. Proteger los datos personales, la información reservada, la confidencialidad y el secreto profesional durante todo el ciclo de uso.

4. Identificar sesgos e impactos diferenciados sobre personas y grupos de especial protección constitucional.

5. Documentar de manera proporcional y segura la herramienta, la finalidad, las verificaciones y las correcciones realizadas.

¿Qué significa hacer un uso ético y responsable de la IA en la labor defensorial?

Un uso ético y responsable de la inteligencia artificial (IA) en la labor defensorial implica comprender que estas herramientas son un medio de apoyo, no un reemplazo del criterio profesional ni del contacto humano con las personas usuarias. La función de la IA debe entenderse como complementaria a la labor jurídica, nunca sustitutiva de la reflexión, la empatía y el juicio humano que caracterizan el ejercicio de la defensa pública. Las decisiones que afectan derechos no pueden delegarse a sistemas automatizados; la responsabilidad última recae siempre en quien ejerce la defensa.

El uso responsable también exige valorar de manera rigurosa los riesgos asociados a la implementación de estas tecnologías. Antes de incorporar una herramienta de IA, las defensoras y los defensores públicos deben considerar su fiabilidad, su procedencia, los sesgos que pueda contener y los posibles impactos que su empleo pueda generar sobre las personas que buscan el acompañamiento de la Defensoría. Esta valoración cuidadosa materializa el principio de precaución y asegura que el interés superior de las personas —y no la mera eficiencia técnica— sea el criterio rector de toda decisión.

Asimismo, la ética en el uso de la IA exige una protección estricta de la confidencialidad. Aunque lo deseable sería que las defensoras y los defensores públicos verificaran plenamente cómo funcionan las plataformas y si cumplen a cabalidad las normas de protección de datos, lo cierto es que muchos sistemas de IA operan como tecnologías opacas y no permiten conocer en detalle su procesamiento interno ni la forma en que manejan la información.

Por esta razón, el deber se entiende como una obligación de diligencia reforzada: (i) utilizar únicamente herramientas autorizadas por la entidad; (ii) evitar el ingreso de información sensible o que permita identificar a las personas y (iii) verificar que las plataformas ofrezcan garantías razonables de seguridad digital. La confianza de las personas en la Defensoría depende, en buena medida, de este manejo prudente y cuidadoso de la información.

Finalmente, garantizar que las decisiones finales sean siempre humanas significa preservar la autonomía de la defensa pública frente a los sistemas automatizados. La IA puede ofrecer sugerencias, clasificar información o facilitar tareas operativas, pero la interpretación del derecho, la valoración de las pruebas y la estrategia jurídica deben permanecer en manos de quienes tienen el deber de proteger los derechos humanos. Solo así se asegura que la tecnología sea una aliada de la justicia y no un factor que la desplace o debilite.



¿Qué encontrarán en esta guía?

Esta guía, además del prólogo y la introducción, se organiza en nueve secciones. Incluye un glosario básico; principios para el uso ético y responsable; casos de uso y escenarios diferenciales; escalas de riesgo, transparencia y ayudas visuales; un anexo de *prompts* seguros; reglas de gobernanza y actualización; recursos adicionales y una bibliografía consolidada. La versión de revisión incorpora, además, una matriz de trazabilidad editorial para facilitar la validación de los ajustes.

¿Esta guía sustituye otras normas?

Esta guía ofrece orientaciones generales para el uso ético, seguro y responsable de la inteligencia artificial en el ejercicio de la labor defensorial. Su propósito es servir como un marco de referencia para fortalecer la autonomía profesional, la prudencia técnica y el compromiso con los derechos humanos en la aplicación de herramientas digitales.

No obstante, su contenido no sustituye el criterio jurídico ni el juicio ético que corresponde a cada defensora y defensor público en el cumplimiento de sus funciones. Tampoco reemplaza las normas legales, disciplinarias o reglamentarias que rigen la actuación de las y los servidores de la Defensoría del Pueblo.

En consecuencia, cada profesional conserva plena responsabilidad por las decisiones, actuaciones y resultados que deriven del uso de herramientas de inteligencia artificial. Esto incluye la obligación de verificar la confiabilidad de la información generada, garantizar la protección de los datos personales y evitar cualquier dependencia mecánica o delegación automática en los sistemas digitales. El uso de la inteligencia artificial debe entenderse, por tanto, como un apoyo complementario al trabajo humano y no como un sustituto del análisis jurídico, la deliberación ética ni la valoración probatoria o argumentativa que sustentan la función defensorial.



II. ¿Qué entendemos por...?

Algoritmo: “Serie de instrucciones para realizar cálculos u otras tareas, ya sea en matemáticas o en informática. En el caso de la inteligencia artificial, un algoritmo proporciona las reglas o instrucciones que permiten a un computador aprender a aprender del entorno y realizar un conjunto de tareas”¹.

Alucinaciones y desinformación: “Resultados imprecisos o inexactos que pueden generar los sistemas de IA generativa, especialmente los basados en modelos de lenguaje de gran tamaño (LLM)”².

Modelo de lenguaje de gran tamaño (LLM por su sigla en inglés): “Es un tipo de algoritmo de inteligencia artificial que utiliza técnicas de aprendizaje profundo y conjuntos de datos muy grandes para comprender, resumir, generar y predecir nuevo contenido. El término IA generativa también está estrechamente relacionado con los LLM, que son, de hecho, un tipo de IA generativa que se ha diseñado específicamente para ayudar a generar contenido”³.

Prompt: “Instrucciones realizadas por el usuario de la IA generativa, para que esta produzca un resultado o salida”⁴.

Sesgo de IA: “El sesgo de IA es una diferencia sistemática en el tratamiento de ciertos objetos, personas o grupos (por ejemplo, estereotipos, prejuicios o favoritismo) en comparación con otros mediante algoritmos de IA. Este tipo de sesgos resultan de los datos usados para entrenar el algoritmo, lo que puede reforzar prejuicios de raza, género, sexualidad, etnia o discapacidad”⁵. Por ejemplo, un sistema puede recomendar de forma reiterada jurisprudencia en la que la víctima sea un hombre, debido a que fue entrenado con un mayor número de casos de ese tipo, lo que puede dejar por fuera sentencias relevantes relacionadas con mujeres o con poblaciones diversas.

Sistema de IA: “Sistemas computacionales que pueden ‘procesar datos e información de un modo que se asemeja a un comportamiento inteligente y que típicamente incluye aspectos de razonamiento, aprendizaje, percepción, predicción, planificación o control”⁶.



1. Consejo Superior de la Judicatura. Acuerdo PCSJA24-12243 del 16 de diciembre de 2024.

2. Ibíd.

3. Ibíd.

4. Ibíd.

5. Ibíd.

6. UNESCO (2024). Directrices de la UNESCO para el uso de sistemas de inteligencia artificial en juzgados y tribunales.

II. ¿Qué entendemos por...?

Sistemas o herramientas de inteligencia artificial generativa: “Sistemas computacionales que se comunican en lenguaje natural, es decir, se comunican de manera similar a como lo harían los humanos, que son capaces de dar respuestas a preguntas relativamente complejas y pueden crear contenidos, proporcionar un texto, imagen o sonido, siguiendo una pregunta formulada o instrucciones (*prompt*). La IA generativa, en lugar de conservar las páginas web existentes, genera nuevos contenidos de forma automática en respuesta a instrucciones en interfaces conversacionales de lenguaje natural (*prompts*). Normalmente, las aplicaciones informáticas conocidas como chatbots operan a partir de estos sistemas”⁷.

Modelo de IA:

Componente matemático o computacional entrenado para producir inferencias, predicciones o contenidos. Un modelo no equivale por sí solo a la totalidad del sistema de IA.

Herramienta de IA generativa:

Aplicación que permite a una persona interactuar con uno o varios modelos para generar texto, imágenes, audio, código u otros contenidos a partir de instrucciones.

Resultado o salida:

Contenido producido por la herramienta en respuesta a un *prompt*. Debe considerarse un insumo provisional sujeto a revisión y verificación humana.

Anonimización y riesgo de reidentificación:

La anonimización busca impedir que una persona pueda ser identificada. Retirar nombres no siempre es suficiente: la combinación de territorio, fechas,

hechos, cargos, condiciones de salud u otros datos puede permitir la reidentificación.

Supervisión humana significativa:

Intervención real y competente de una persona con capacidad para comprender el contexto, cuestionar el resultado, corregirlo, descartarlo y asumir la responsabilidad final.

Regla terminológica: esta Guía utiliza “inteligencia artificial (IA)” para la categoría general; “sistema de IA” para el conjunto de modelos, datos, software, infraestructura y procesos; “modelo de IA” para el componente computacional; y “herramienta de IA generativa” para la aplicación con la que interactúa la persona usuaria.

Sistema de IA (convención terminológica de esta Guía):

En este documento se utilizan tres términos relacionados que conviene distinguir: (1) ‘inteligencia artificial (IA)’ es el término genérico que designa el campo tecnológico y el conjunto de sistemas que lo componen; (2) ‘sistema de IA’ es el término técnico-jurídico de referencia —adoptado por el Reglamento (UE) 2024/1689 y el CONPES 3975 de 2019— y se emplea en secciones de principios y prohibiciones; (3) ‘herramienta de IA’ es su equivalente coloquial, admisible en orientaciones prácticas y ejemplos; (4) ‘IA generativa’ o ‘sistema de IA generativa’ designa el subtipo más relevante para la labor defensorial —chatbots, asistentes de escritura, buscadores potenciados— y se usa cuando la orientación aplica exclusivamente a estos sistemas. En ningún caso deben intercambiarse estos términos en secciones normativas. (Reglamento UE 2024/1689, art. 3; CONPES 3975 de 2019; Decreto 1263 de 2022).

7. Supra nota 1.

III. Principios básicos para un uso ético y responsable de la IA en la labor defensorial

Primacía de los derechos humanos y los derechos fundamentales

Todo uso de la inteligencia artificial (IA) como apoyo a la labor de las y los defensores públicos debe garantizar, proteger y promover los derechos humanos y los derechos fundamentales. Este principio implica que la tecnología nunca puede anteponerse a la dignidad, la libertad y la igualdad de las personas. La IA debe servir como herramienta para fortalecer la justicia y ampliar el acceso a los derechos, no para restringirlos ni condicionar su ejercicio. En consecuencia, cualquier uso de estas herramientas debe evaluarse con base en su impacto sobre las personas y asegurar que toda innovación tecnológica esté al servicio de la protección efectiva, la no discriminación y el respeto integral de los derechos humanos.

La IA no opera en un contexto social vacío. Los sistemas de aprendizaje automático reproducen los patrones de los datos con los que fueron entrenados y, en muchos casos, esos datos reflejan desigualdades históricas y estructurales.

La evidencia internacional muestra que las herramientas de IA generativa pueden amplificar sesgos de género, raza y origen étnico: sistemas entrenados con datos predominantemente masculinos tienden a subrepresentar a las mujeres en roles de autoridad jurídica; las herramientas de reconocimiento biométrico presentan tasas de error significativamente más altas en personas racializadas y los modelos de análisis de riesgo tienen mayor probabilidad de producir resultados desfavorables para personas con bajos ingresos o perfiles migrantes (ONU Mujeres, 2025; Lousada-Arochena, 2024; ONU-DH Colombia, 2024).

En la labor defensorial, esta realidad adquiere una dimensión crítica: cuando una herramienta de IA sugiere jurisprudencia, argumentos o borradores de escritos, puede estar invisibilizando precisamente los precedentes que protegen a las poblaciones que más necesitan el acompañamiento de la Defensoría del Pueblo. Por ello, el uso de la IA en la defensa pública exige vigilancia activa frente a estos sesgos, revisión crítica de los resultados y la aplicación permanente del enfoque diferencial e interseccional que orienta la misión institucional.





Apoyo y no sustitución

La inteligencia artificial puede ser una aliada valiosa para apoyar ciertas tareas en la labor defensorial, siempre que su uso se mantenga dentro de los límites que preservan la autonomía profesional y el juicio humano. Su papel es complementar, no reemplazar, la decisión y el criterio de la o el defensor público, para garantizar que la tecnología nunca sustituya el razonamiento jurídico, la empatía ni la sensibilidad que caracterizan la atención a las personas usuarias del servicio. En suma, la IA ayuda, pero no decide.

Supervisión humana

Todo uso de la inteligencia artificial requiere revisión y control humano. Estos sistemas pueden equivocarse, generar sesgos, producir información inexacta o presentar resultados no verificables. Por ello, cada texto, análisis, recomendación o producto elaborado con apoyo de la IA debe ser cuidadosamente revisado y validado antes de ser utilizado o compartido, para garantizar la responsabilidad profesional y la fiabilidad de la información.

Criterio profesional y responsabilidad personal

Cada defensora o defensor público es responsable del contenido que presenta, incluso si fue asistido por IA. Toda la información generada por estas herramientas debe verificarse, corregirse y adaptarse antes de ser utilizada. El compromiso ético y profesional exige que el uso de la IA no debilite la calidad, la veracidad ni la independencia del trabajo jurídico.

Las personas que utilicen herramientas de inteligencia artificial deben asumir la responsabilidad de comprender sus limitaciones, potencialidades y características antes de incorporarlas en cualquier proceso de trabajo, creación o toma de decisiones. Este principio exige un uso informado y consciente que permita dimensionar los alcances reales de la tecnología, prevenir errores derivados de expectativas equivocadas y garantizar que la intervención humana siga siendo un elemento central en la valoración, validación y supervisión de los resultados generados por sistemas de IA.



Proporcionalidad

La inteligencia artificial debe utilizarse únicamente cuando resulte verdaderamente útil y pertinente para la labor defensorial. La tecnología es un medio, no un fin en sí misma, por lo que su aplicación debe responder siempre a un propósito claro y a una justificación adecuada. Cada defensora y cada defensor público tiene la responsabilidad de valorar si la herramienta aporta un beneficio real al caso y si su uso es proporcional a los fines de la defensa y al respeto de los derechos involucrados.

Existen tareas que se resuelven con una lectura directa, solicitudes simples que requieren una respuesta inmediata o gestiones que exigen consultas obligatorias en bases oficiales, por lo que la intervención de la IA no aporta valor y puede generar demoras o inconsistencias. De igual forma, ciertos documentos requieren una interpretación jurídica estrictamente humana, especialmente cuando contienen matices constitucionales, elementos probatorios sensibles o valoraciones que forman parte del juicio profesional del defensor público. Estas orientaciones permiten evitar la dependencia tecnológica y promover un uso racional y pertinente de la IA.

Uso de la IA desde una perspectiva centrada en las personas

Para la Defensoría del Pueblo, en su labor constitucional como institución nacional de derechos humanos, la ciudadanía constituye el eje central de su misión. Desde esta premisa, los sistemas de inteligencia artificial deben emplearse como herramientas para acercar la justicia, facilitar la comprensión y mejorar la atención a las personas, pero nunca para sustituir el trato humano, la empatía ni la escucha activa que dan sentido a la labor defensorial.

Enfoque diferencial e interseccional

Los sistemas de IA pueden producir impactos distintos según el género, la orientación sexual, la identidad de género, la edad, la discapacidad, la pertenencia étnica, la nacionalidad, el territorio, la situación socioeconómica o la privación de la libertad en las personas. La aparente neutralidad técnica no garantiza resultados equitativos. La revisión humana deberá identificar si el sistema invisibiliza circunstancias relevantes, reproduce estereotipos, utiliza categorías inadecuadas o amplifica desigualdades preexistentes (UNESCO, 2021; Buolamwini & Gebru, 2018).

El enfoque diferencial no autoriza a pedir a la herramienta que infiera atributos sensibles. Su finalidad es que la defensora o el defensor reconozca las circunstancias concretas de la persona, revise críticamente la salida y adopte salvaguardas. Cuando el riesgo no pueda mitigarse o el sistema pueda afectar la igualdad, la dignidad, la defensa técnica o el debido proceso, deberá abstenerse de utilizarlo.

Población o condición	Riesgo posible	Salvaguarda mínima
Mujeres y víctimas de violencias basadas en género	Minimización del contexto, culpabilización o uso de estereotipos.	Revisar contexto de poder, lenguaje, riesgo y enfoque de género.
Personas racializadas y comunidades étnicas	Asociaciones entre identidad, territorio y peligrosidad; invisibilización de autoridades o normas propias.	Evitar inferencias; verificar contexto cultural, enfoque étnico y fuentes especializadas.
Personas LGBTQ+	Categorías binarias, nombre o género incorrecto, exposición innecesaria de datos sensibles.	Respetar autodenominación e identidad; minimizar datos y revisar lenguaje no discriminatorio.
Personas migrantes	Errores de traducción, confusión normativa o perfilamiento por nacionalidad.	Verificar traducción y régimen aplicable; acudir a fuentes oficiales y apoyo humano.
Personas con discapacidad	Inaccesibilidad, interpretación errónea de formas de comunicación o valoración indebida de credibilidad.	Aplicar ajustes razonables y no convertir transcripciones automáticas en juicios de credibilidad.
Niñas, niños y adolescentes	Ausencia de enfoque de curso de vida, privacidad o autonomía progresiva.	Aplicar interés superior, protección reforzada y lenguaje adecuado.
Personas privadas de la libertad	Reproducción acrítica de antecedentes o perfiles de riesgo.	Contrastar fuentes, individualizar el análisis y preservar confidencialidad.

La información del cuadro es orientadora y debe aplicarse de manera contextual, para evitar clasificaciones automáticas de personas o grupos

Confidencialidad y protección de la información

Las y los defensores públicos deben proteger los datos personales, sensibles, reservados y la información amparada por confidencialidad o secreto profesional. Solo podrán emplear herramientas expresamente evaluadas, habilitadas o autorizadas por la Defensoría del Pueblo para la finalidad y la clasificación de información correspondientes.

La anonimización no habilita automáticamente el uso: antes de ingresar cualquier contenido, deberá evaluarse si la combinación de hechos, fechas, territorio, cargos u otras características permite reidentificar a una persona. Cuando subsista un riesgo razonable, la información no deberá ser ingresada.

Transparencia

El uso de la inteligencia artificial en la labor defensorial debe ser transparente, comprensible y verificable. Cuando sea pertinente, se debe informar que se empleó un sistema de IA como apoyo en el proceso. Los resultados obtenidos deben poder explicarse y sustentarse con claridad, para garantizar la trazabilidad y la rendición de cuentas. La transparencia en el uso de estas herramientas fortalece la confianza ciudadana y consolida la credibilidad institucional de la Defensoría del Pueblo.

IV. Usos comunes de la inteligencia artificial en la labor defensorial



El uso de la inteligencia artificial en la labor defensorial debe orientarse siempre por los principios básicos antes mencionados, en especial los de primacía de los derechos humanos, responsabilidad, transparencia y proporcionalidad. Esta sección busca traducir dichos principios en ejemplos concretos que muestren cómo las y los defensores públicos pueden aprovechar las herramientas de IA como apoyo técnico y estratégico, sin poner en riesgo su criterio jurídico, su independencia ni el enfoque humano que caracteriza la defensa pública.

Los casos y ejemplos presentados pretenden ilustrar de manera práctica las oportunidades que la IA ofrece para optimizar la gestión de la información, fortalecer la argumentación jurídica y mejorar la calidad del servicio, sin desconocer los límites éticos y jurídicos que deben regir su uso. Se trata de mostrar cómo la tecnología puede servir para ampliar el acceso a la justicia, agilizar procesos y facilitar el análisis de casos complejos, siempre que se mantenga un control humano efectivo y se protejan los derechos de las personas usuarias.

Asimismo, esta sección busca establecer un estándar común de buenas prácticas institucionales que oriente el uso responsable y consciente de estas herramientas. Promover un lenguaje y una comprensión compartida sobre el uso de la IA en la defensa pública contribuye a evitar riesgos de sesgo, pérdida de confidencialidad o dependencia tecnológica y refuerza la rendición de cuentas dentro del ejercicio defensorial.

Elaboración de resúmenes de expedientes o documentos jurídicos

Esta actividad consiste en revisar, ordenar y sintetizar la información contenida en los expedientes, denuncias, pruebas o documentos relevantes de un caso. Su propósito es facilitar la comprensión de los hechos, los antecedentes y las actuaciones procesales, al permitir que la defensora o defensor público tenga una visión clara del caso antes de diseñar su estrategia jurídica.



Permitido: utilizar IA para resumir textos extensos, agrupar información por temas o generar listados de hechos clave. También puede ayudar a crear esquemas o líneas de tiempo.



Evitar: subir documentos con nombres, números de identificación, audios o datos personales sin anonimizar. No usar plataformas abiertas (como versiones gratuitas) ni almacenar allí la información del caso.



Supervisión humana: revisar siempre que los resúmenes sean fieles al contenido original. Comparar los resultados con el expediente físico o digital. Corregir errores, omisiones o interpretaciones inexactas. La defensora o el defensor público debe validar personalmente la información antes de utilizarla.



Sugerencias específicas:

-  El primer paso para comprender un caso es revisar personalmente el expediente completo. Ningún resumen generado por IA reemplaza el conocimiento directo del contenido original. La lectura inicial permite a la defensora o al defensor público identificar la naturaleza de los hechos, su secuencia temporal, la relevancia de las pruebas y la coherencia entre los documentos. Una vez familiarizado con el caso, el profesional puede comparar los resúmenes generados por la IA con los hechos reales descritos en el expediente, detectar posibles omisiones o interpretaciones erróneas y corregirlas.
-  Anonimizar cualquier documento antes de utilizar herramientas de inteligencia artificial. Antes de ingresar información en una herramienta de IA, es indispensable eliminar nombres, números de identificación, direcciones, datos médicos, grabaciones, fotografías y cualquier otro elemento que permita reconocer a las personas involucradas. Esta depuración previa protege la confidencialidad del caso y evita vulneraciones al derecho a la privacidad de las personas usuarias.
-  Para ello, se recomienda sustituir los nombres propios por referencias neutras como “Persona A”, “Persona B” o “Servidor Público 01” y reemplazar los números de identificación por códigos internos consecutivos. También es necesario suprimir teléfonos, correos electrónicos y cualquier dato que permita la identificación directa o indirecta de una persona. Cuando el documento incluya información altamente sensible —como datos sobre salud, infancia, violencia basada en género o condición étnica—, se debe retirar dicha información y, de ser pertinente, utilizar descriptores generales que no comprometan la identidad de nadie.
-  Además, en los expedientes que involucren a varias personas, se recomienda asignar un código único a cada una y utilizarlo de manera consistente en todo el documento. Antes de cargar el texto a una herramienta de IA, se debe verificar que no permanezca ningún dato personal ni información capaz de revelar la identidad de alguien. En caso de duda es apropiado solicitar orientación a la Defensoría Delegada para la Protección de Derechos Digitales y Libertad de Expresión.
-  Los resultados producidos por una herramienta de IA deben entenderse como un punto de partida, no como un producto final. El defensor o la defensora pública debe enriquecer los resúmenes con notas personales, interpretaciones jurídicas y observaciones derivadas de su experiencia profesional.
-  Los sistemas de IA carecen de criterio jurídico y de sensibilidad frente al contexto social y humano de cada caso. Por ello no debe usarse para determinar qué hechos son jurídicamente relevantes, ni para establecer su impacto en la estrategia de defensa. Esa tarea corresponde exclusivamente a la o el defensor público.



Para recordar: el resumen es una ayuda, no reemplaza la lectura del expediente. La precisión y comprensión del caso dependen de la o el defensor público.

Búsqueda y análisis de jurisprudencia o doctrina

En esta actividad, la defensora o el defensor público identifica normas, sentencias, conceptos y doctrinas que pueden sustentar los argumentos de su defensa. Esto requiere localizar fuentes confiables, comparar criterios judiciales y evaluar su pertinencia para el caso. Esta etapa es importante para garantizar la solidez jurídica y la coherencia argumentativa de la estrategia defensorial.



Permitido: solicitar a la IA ayuda para identificar normas o sentencias relevantes o para resumir criterios jurisprudenciales ya conocidos. Puede servir para ordenar información o para identificar temas relacionados.



Evitar: confiar única y exclusivamente en la información generada sin verificar las fuentes o citar textos sin comprobación. En otras palabras, copiar y citar directamente los textos generados por IA sin verificar su existencia o vigencia. No usarla como sustituto de bases de datos oficiales.



Supervisión humana: la IA puede alucinar y generar información errónea, por ejemplo, citar jurisprudencia que no existe, que está desactualizada o que no corresponde a tribunales colombianos. Por ello se debe contrastar cada resultado con fuentes primarias, como las páginas web de la Corte Constitucional, el Consejo de Estado, la Corte Suprema y las bases institucionales. Verificar que las sentencias u otros documentos citados son reales, aplican en cada caso y están actualizados.

Sugerencias específicas:



Las herramientas de IA pueden resultar útiles para obtener una visión general de un asunto. Sin embargo, en muchos casos, es preferible emplear buscadores especializados de jurisprudencia o bases de datos oficiales, como las de la Corte Constitucional, que pueden ser una mejor opción para buscar jurisprudencia.



Cada vez que una herramienta de IA proporcione una cita, sentencia o doctrina, la defensora o el defensor público debe realizar un proceso sistemático de verificación. Esto implica corroborar que: (i) la sentencia o norma realmente existe; (ii) su contenido es jurídicamente relevante y aplicable al caso en cuestión; (iii) pertenece al ordenamiento jurídico colombiano y (iv) el texto reproducido coincide fielmente con la fuente original.



Las herramientas de IA no comprenden el contexto normativo, doctrinal ni ético de los fallos judiciales y, por tanto, no están capacitadas para interpretar su significado jurídico ni su aplicación a un caso concreto. Cualquier intento de dejar que la IA “explique” una sentencia puede conducir a conclusiones equivocadas o incompletas.



Para recordar: la IA puede sugerir sentencias falsas o incompletas. Siempre es necesario y obligatorio verificar la fuente en portales oficiales antes de citarla.

Redacción de oficios, solicitudes, escritos jurídicos o comparación de textos

Esta actividad comprende la elaboración de documentos formales como peticiones, oficios, memoriales, solicitudes de pruebas y comunicaciones con otras entidades e incluye, además, el cruce de información entre distintos tipos de documentos para identificar posibles inconsistencias.



Permitido: usar IA para generar borradores de textos, mejorar la redacción o la estructura, revisar la ortografía del documento o adaptar el tono de la comunicación o elaborar formatos para producción esquematizada de documentos. Puede ser útil para reformular párrafos y hacerlos más claros.

Evitar: presentar documentos generados íntegramente por IA sin ningún tipo de revisión, porque esto puede conllevar a que se envíe información errónea, falsa o imprecisa.



Supervisión humana: revisar cuidadosamente todo texto antes de firmarlo o enviarlo. Para ello es importante confirmar que las citas, nombres, fechas y datos sean correctos y que la versión final debe ser siempre autoría de la o el defensor público.



Sugerencias específicas:



Con el fin de garantizar la transparencia en la labor defensorial, cuando se haya utilizado una herramienta de inteligencia artificial para elaborar un borrador o revisar un texto, se recomienda incluir una breve nota aclaratoria en un pie de página (ver sección V de este documento con ejemplos).



Verificar que el lenguaje sea inclusivo, respetuoso y comprensible.



Evitar automatizar las respuestas a los usuarios; la comunicación debe ser personalizada.



La IA puede resultar útil para reformular textos jurídicos extensos o densos en un lenguaje más claro y accesible, con el fin de facilitar su comprensión por parte de las o los usuarios. Sin embargo, esta simplificación debe hacerse con el máximo cuidado: se debe conservar el sentido jurídico de las normas, la exactitud de los hechos y la coherencia procesal.



Para recordar: la IA puede ayudar a redactar, pero la responsabilidad del contenido del texto que se envía y de la firma es siempre de quien remite dicho correo, comunicación, oficio o documento.



Preparación de audiencias y alegatos iniciales

Esta etapa incluye la organización de la estrategia oral, la preparación de argumentos, la revisión de pruebas, la elaboración de guiones y la anticipación de preguntas o de escenarios posibles. En la preparación de audiencias y alegatos iniciales se busca fortalecer la intervención de la defensora o el defensor público en las diligencias judiciales.



Permitido: usar IA para estructurar argumentos, ensayar exposiciones, generar ideas, posibles preguntas o simular escenarios de audiencia.



Evitar: confiar en respuestas automáticas para definir estrategias de defensa o en argumentos sin sustento jurídico. No usar IA durante audiencias reales. Está expresamente prohibido recrear audiencias reales en las cuales se mencionen nombres de partes o se incluya información sensible durante estos ejercicios.



Supervisión humana: verificar que los argumentos propuestos sean jurídicamente válidos y éticamente apropiados. Ajustar los guiones a las características de cada caso, de la persona usuaria y del contexto procesal. Corregir posibles sesgos o información irrelevante.

Sugerencias específicas:

-  La IA puede ser útil para estructurar esquemas argumentativos, organizar los hechos del caso o identificar posibles líneas de exposición. Sin embargo, la definición de la táctica procesal (qué pruebas presentar, qué argumentos priorizar o qué postura asumir frente al tribunal) corresponde exclusivamente al criterio de la defensora o del defensor público.
-  Antes de incorporar frases, argumentos o formulaciones propuestas por una herramienta de IA, se recomienda verificar que el lenguaje utilizado sea jurídico, respetuoso y adecuado al contexto procesal.
-  Practicar audiencias con simuladores solo en entornos seguros.
-  No almacenar grabaciones ni estrategias reales en plataformas externas, ya que ello representa un grave riesgo para la confidencialidad, la privacidad y el secreto profesional.



Para recordar: La IA puede ayudar a prepararse, pero no a representar. La experiencia, el conocimiento y la empatía de la defensora o el defensor público son insustituibles.



Escenarios diferenciales de aplicación

Los siguientes casos simulados muestran por qué la revisión humana debe incorporar el contexto de las personas y no limitarse a comprobar la corrección gramatical del resultado:

Escenario	Riesgo de la herramienta	Respuesta defensorial
Víctima de violencia basada en género	El resumen presenta agresiones reiteradas como “conflictos de pareja” y omite patrones de control.	Revisar el expediente completo, recuperar el contexto de violencia y corregir el lenguaje.
Persona migrante	La traducción altera una expresión cultural o confunde una categoría jurídica del país de origen.	Usar la traducción solo como apoyo y obtener interpretación humana cuando sea necesaria.
Persona con discapacidad	La transcripción interpreta pausas o formas de comunicación como contradicciones.	No convertir la transcripción en valoración de credibilidad y aplicar ajustes razonables.
Persona LGBTQ+	El borrador utiliza un nombre o género distinto al indicado por la persona.	Corregir, respetar la identidad y retirar información sensible innecesaria.
Comunidad étnica	La respuesta desconoce autoridades propias, contexto territorial o jurisprudencia específica.	Consultar fuentes especializadas y aplicar enfoque étnico y cultural.
Persona privada de la libertad	La herramienta prioriza antecedentes y omite condiciones actuales del caso.	Individualizar el análisis, verificar expediente y evitar perfiles automáticos.

Estos ejemplos no deben utilizarse con información real en herramientas no autorizadas.

V. Transparencia y escalas de uso

En esta sección se presentan una serie de recomendaciones y herramientas prácticas orientadas a promover un uso ético, responsable y transparente de la inteligencia artificial en la labor defensorial. Estas orientaciones buscan servir de guía para que las y los defensores públicos incorporen la tecnología de manera consciente, para asegurar que su aplicación fortalezca la protección de los derechos humanos y no la comprometa.

Además, ofrecen criterios claros para identificar buenas prácticas, prevenir riesgos y garantizar que el uso de la IA se mantenga siempre bajo supervisión humana, con criterios de justicia, proporcionalidad y respeto por la dignidad de las personas. De este modo, la tecnología se entiende como un apoyo a la misión constitucional de la Defensoría del Pueblo y no como un sustituto de su labor humana, empática y garantista.



Semáforo de usos permitidos o no permitidos

Antes de usar IA: ruta de decisión

La ruta siguiente debe aplicarse antes de formular cualquier *prompt*. Si una respuesta conduce a “no usar”, la persona defensora deberá detener el proceso y recurrir a una alternativa institucional segura. La existencia de una licencia paga no sustituye la autorización institucional ni garantiza privacidad.



Figura 1. Ruta de decisión previa al uso de inteligencia artificial.

Ciclo de uso seguro y responsable



Figura 2. Ciclo de uso seguro, responsable y trazable.

El siguiente semáforo orienta a las y los defensores públicos sobre los distintos niveles de riesgo y conveniencia en el uso de herramientas de inteligencia artificial, proporcionando una guía visual sencilla para tomar decisiones informadas.

El nivel del semáforo depende de factores como la sensibilidad de la información involucrada, el propósito del uso de la tecnología y el grado de supervisión humana que se ejerza sobre los resultados obtenidos.

Su objetivo es ayudar a identificar cuándo el uso de la IA puede considerarse seguro y apropiado, cuándo requiere precauciones adicionales y cuándo debe evitarse por completo.

Esta herramienta busca fomentar un uso consciente, ético y proporcional de la tecnología y recordar que el juicio humano sigue siendo esencial para garantizar la protección de los derechos y la integridad del proceso defensorial.



Tarea o actividad	Uso permitido
Ingresar o compartir información personal, sensible, reservada o confidencial en herramientas no autorizadas.	No permitido.
Pedir a la IA que adopte decisiones, valore pruebas, resuelva el problema jurídico o redacte un concepto oficial definitivo.	No permitido.
Generar alegatos, escritos o conceptos finales listos para presentar con información real del caso.	No permitido.
Definir la teoría del caso, la estrategia de defensa o la táctica procesal.	No permitido.



Tarea o actividad	Uso permitido
Transcripción de entrevistas.	Solo en herramienta autorizada, con consentimiento cuando aplique, minimización de datos y revisión integral.
Resumen de documentos y elaboración de cronologías.	Sí, con información anonimizada, control de reidentificación y verificación contra el original.
Traducción de documentos.	Sí, con herramienta autorizada y revisión humana; no sustituye intérprete cuando sea necesario.
Apoyo en estructuras o borradores jurídicos preliminares.	Sí, sin delegar valoración, aplicación normativa o estrategia; requiere revisión exhaustiva.
Búsqueda de normas o sentencias.	Sí, solo como exploración; toda referencia debe verificarse en fuente oficial.



Tarea o actividad	Uso permitido
Mejorar ortografía, gramática o claridad de apartes sin datos personales, sensibles, reservados o del caso.	Sí, con revisión de la persona defensora.
Adaptar lenguaje, crear listas de chequeo, formatos vacíos o borradores de comunicaciones generales.	Sí, con revisión humana y sin información protegida.



Prohibiciones expresas

Está prohibido ingresar en herramientas no autorizadas información confidencial, personal, sensible, reservada o relacionada con víctimas, delitos, niños, niñas y adolescentes, personas en situación de vulnerabilidad o estrategias de defensa.

Tampoco podrán delegarse a un sistema de IA la valoración probatoria, la interpretación final del caso, la decisión jurídica ni la definición de la estrategia. Incluso en herramientas autorizadas deberán respetarse la finalidad habilitada, la clasificación de la información, la minimización de datos y las políticas institucionales aplicables.

Etiquetas y reconocimiento del uso de IA

La transparencia debe aplicarse de manera significativa y proporcional. Todo uso sustantivo de IA relacionado con un caso deberá contar con trazabilidad interna segura.

La divulgación externa será necesaria cuando la herramienta haya incidido materialmente en el contenido, cuando sea relevante para la comprensión o contradicción del producto, o cuando una norma o instrucción institucional lo exija. Las correcciones puramente ortográficas o de formato no requieren, por sí solas, una etiqueta pública, sin perjuicio del registro interno que corresponda.

Nivel	Regla	Ejemplo
Trazabilidad interna	Registrar herramienta, finalidad, <i>prompt</i> depurado, verificación y correcciones, sin duplicar datos protegidos.	Resumen o borrador relacionado con un caso.
Divulgación externa	Informar el uso cuando la contribución sea sustantiva o relevante para control y contradicción.	Documento cuyo contenido o estructura dependió materialmente de IA.
Sin etiqueta externa automática	No exigir nota pública por correcciones menores de ortografía o formato.	Ajuste de puntuación en un texto ya elaborado por la persona defensora.

Ejemplos de posibles etiquetas para agregar como pie de página de los documentos

- Este documento fue elaborado con apoyo parcial de herramientas de inteligencia artificial (nombre del sistema utilizado), bajo la supervisión y validación de la o el defensor público responsable (nombre del funcionario o dependencia), conforme con los principios de ética, transparencia y rendición de cuentas de la Defensoría del Pueblo.
- La redacción de este texto contó con la asistencia de inteligencia artificial (nombre del sistema utilizado) para tareas de síntesis y estructuración de la información. El contenido fue revisado, ajustado y aprobado por la o el defensor público encargado (nombre del funcionario o dependencia).
- Cronología generada con apoyo de herramientas de inteligencia artificial (nombre del sistema o herramienta utilizada), utilizada como apoyo técnico para la organización de la información, bajo supervisión humana y revisión final de la o el defensor público responsable.
- Este informe fue elaborado con asistencia tecnológica basada en inteligencia artificial (nombre del sistema utilizado), empleada para el procesamiento y análisis de información. Los resultados fueron verificados y contextualizados por la o el defensor público encargado (nombre del funcionario o dependencia).
- Documento preparado con apoyo de inteligencia artificial (nombre del sistema utilizado), utilizada exclusivamente como herramienta de apoyo técnico. El contenido, las conclusiones y las recomendaciones reflejan el criterio profesional de la o el defensor público responsable (nombre del funcionario o dependencia).
- Producto comunicativo elaborado con asistencia parcial de inteligencia artificial (nombre del sistema utilizado), en cumplimiento de los lineamientos institucionales sobre uso ético y responsable de la tecnología. El texto final fue revisado y validado por la o el defensor público responsable (nombre del funcionario o dependencia).
- Este texto fue desarrollado con el apoyo de inteligencia artificial (nombre del sistema utilizado) para la organización de ideas y edición preliminar. Todo su contenido ha sido revisado y aprobado por la o el defensor público responsable (nombre del funcionario o dependencia), para garantizar la fidelidad y la integridad de la información.





Kit de supervivencia para un uso seguro y responsable de la IA⁸

1. Recuerda que la IA no entiende ni valida su propio contenido

Las herramientas de IA generativa no comprenden el significado de lo que producen. No tienen conciencia, criterio jurídico ni capacidad para verificar si la información es correcta o aplicable a un caso real. Esto significa que pueden alucinar, generar respuestas falsas, inexactas o engañosas con apariencia de verdad. Por eso, cada texto, resumen o cita obtenido mediante IA debe ser verificado, revisado y contrastado con fuentes oficiales o documentos originales antes de utilizarlo en el trabajo defensorial.

2. Protege la confidencialidad: nunca ingreses información sensible o reservada

No se debe incluir en sistemas de inteligencia artificial —especialmente en servicios gratuitos, de acceso abierto o con políticas de privacidad desconocidas— ningún tipo de información que permita identificar a las personas usuarias. Esto incluye nombres, números de cédula, antecedentes judiciales, datos médicos, grabaciones, fotografías, direcciones, comunicaciones privadas o cualquier otro dato sensible. La protección de la identidad y la privacidad de las personas es una obligación legal y ética de la Defensoría del Pueblo. Ingresar este tipo de información en plataformas no seguras puede exponer a las personas a riesgos de vulneración de derechos, filtraciones o usos indebidos de sus datos. Por ello, las y los defensores públicos deben verificar siempre las condiciones de seguridad, confidencialidad y almacenamiento de cualquier herramienta tecnológica antes de emplearla en el tratamiento o análisis de casos.

3. Elige cuidadosamente las herramientas de IA

Existen herramientas de acceso libre, comercial, institucional o de código abierto, pero no todas ofrecen garantías equivalentes. Solo podrán emplearse herramientas expresamente evaluadas, habilitadas o autorizadas por la Defensoría del Pueblo para la finalidad y el tipo de información correspondientes.

⁸ Esta sección se construyó a partir de los estándares de UNESCO (2024), Guía para el uso de Inteligencia Artificial para Abogados del Laboratorio Fintech y Legaltech Unidad de Innovación y Transformación Digital del CPACF (2025); Lawyers and Generative AI Guidance (2024) y la UIA Guidelines on the use of artificial intelligence systems by lawyers (2025).

El carácter pago de una suscripción individual no constituye, por sí solo, garantía de seguridad, confidencialidad, exclusión del entrenamiento, auditoría o cumplimiento institucional.

Antes de utilizar cualquier sistema, es fundamental informarse sobre:

- Quién administra el sistema (proveedor, empresa o entidad responsable).
- Qué tipo de datos procesa y cómo los almacena o utiliza.
- Si cuenta con políticas de confidencialidad y seguridad claras.
- Si permite conservar y auditar el registro de los *prompts* (instrucciones) utilizados y sus resultados.

El uso de herramientas institucionalmente autorizadas garantiza que la tecnología esté al servicio de la función defensorial y no exponga la información ni los derechos de las personas usuarias.

4. Comprende que su uso implica riesgos

No se espera que las y los defensores públicos sean expertos en programación o en desarrollo tecnológico. Sin embargo, sí es fundamental que comprendan los riesgos asociados al uso de herramientas de inteligencia artificial antes de emplearlas, de modo que puedan anticiparlos y adoptar medidas razonables para mitigarlos. El conocimiento básico sobre cómo funcionan estos sistemas es una forma de protección para la Institución y para las personas atendidas.

Entre los principales riesgos del uso de la IA se encuentran la filtración o exposición no autorizada de datos sensibles, la generación de información errónea o inexacta, la presencia de sesgos que pueden reproducir o amplificar discriminaciones existentes, la posible vulneración de derechos fundamentales, y la pérdida de control o trazabilidad sobre los contenidos generados. Conocer estos riesgos permite decidir cuándo y cómo utilizar estas herramientas de manera ética, segura y conforme con los principios de la Defensoría del Pueblo.

5. Es fundamental mantener la trazabilidad y la seguridad de la información

Los usos sustantivos de IA deben dejar un rastro verificable y proporcional que permita reconstruir la finalidad, la herramienta utilizada, el *prompt* depurado, el resultado relevante, las fuentes consultadas y las correcciones humanas. La bitácora no debe reproducir datos personales, sensibles, reservados ni información de la estrategia. Su acceso, conservación y eliminación deberán sujetarse a las políticas institucionales de seguridad, gestión documental y protección de datos.

Registrar los *prompts* o comandos empleados, junto con las versiones o iteraciones de los resultados, no solo permite explicar cómo se llegó a una determinada conclusión, sino también identificar posibles errores, sesgos o inconsistencias generadas por el sistema. Mantener este registro constituye una buena práctica institucional, ya que protege a la persona defensora y a la Defensoría del Pueblo frente a eventuales cuestionamientos sobre la validez, el origen o la integridad de la información.

Además, la trazabilidad favorece la mejora continua en el uso de la tecnología, al permitir evaluar qué herramientas son más confiables y en qué contextos resultan útiles. De este modo, la inteligencia artificial puede incorporarse a la labor defensorial con criterios de control, responsabilidad y aprendizaje institucional, para asegurar que su uso fortalezca la confianza ciudadana y no comprometa la integridad de los procesos.

6. Evita la dependencia tecnológica y conserva tu autonomía profesional

La inteligencia artificial es un apoyo, no una solución automática. Ninguna herramienta sustituye el juicio jurídico, la empatía ni la sensibilidad humana que exige la defensa pública. Por ello, cada resultado debe someterse a una revisión crítica y, en todo caso, la decisión final debe adoptarse conforme con la ley, la ética y la protección de los derechos de la persona usuaria.

7. Actualizarse sobre normas y riesgos emergentes

La inteligencia artificial evoluciona constantemente. Por ello, las y los defensores públicos deben mantenerse informados sobre las leyes, reglamentos y estándares internacionales que regulan el desarrollo y uso de estas tecnologías, así como sobre los avances éticos, doctrinales y jurisprudenciales que inciden en su aplicación en el ámbito de la justicia. Comprender estos cambios es esencial para garantizar que el uso de la IA se mantenga dentro de los marcos legales y de derechos humanos que orientan la labor defensorial.

La Defensoría del Pueblo acompañará a las y los defensores públicos en este proceso de actualización permanente, para promover espacios de formación, divulgación y reflexión sobre el impacto de la tecnología en el acceso a la justicia y en la protección de los derechos fundamentales. De este modo, la institución reafirma su compromiso con un uso responsable, informado y humano de la inteligencia artificial.

8. Priorizar siempre los derechos humanos y los valores institucionales

Toda aplicación de inteligencia artificial en la Defensoría Pública debe alinearse con los principios constitucionales, el marco legal que regula la misión y funciones de la Defensoría del Pueblo, así como con el Decálogo Institucional. El uso de estas tecnologías debe desarrollarse dentro de los límites del Estado social y democrático de derecho y en coherencia con los mandatos de promoción, defensa y garantía de los derechos humanos. En ningún caso, la inteligencia artificial puede poner en riesgo los derechos de las personas usuarias, afectar la confidencialidad de la información o debilitar la confianza ciudadana en la justicia. Por el contrario, su incorporación debe fortalecer la transparencia, la equidad, la rendición de cuentas y la legitimidad de la acción defensorial.



Tips concretos para la creación de *prompts*

El uso adecuado de herramientas de inteligencia artificial depende, en gran medida, de la claridad con la que se formulan las instrucciones o *prompts*. Un *prompt* es el mensaje o conjunto de indicaciones que se entrega a la herramienta para que produzca una respuesta. Formularlo correctamente permite obtener resultados más precisos, útiles y ajustados al marco jurídico, ético y de transparencia institucional de la Defensoría del Pueblo.

Esta sección ofrece orientaciones prácticas para elaborar *prompts* efectivos, claros y seguros, que ayuden a aprovechar el potencial de la inteligencia artificial sin comprometer la confidencialidad de la información, la integridad de los datos ni el criterio profesional. Cada recomendación busca fortalecer el uso responsable y transparente de estas herramientas dentro de la labor defensorial, para garantizar siempre el respeto por los derechos humanos, la rendición de cuentas y la protección de los datos personales.

Además, todo uso de la inteligencia artificial debe guiarse por los principios definidos en esta Guía: la primacía de los derechos humanos y de los derechos fundamentales, el apoyo y no sustitución del juicio humano, la supervisión permanente, el criterio profesional y la responsabilidad personal, la proporcionalidad en su aplicación, el uso centrado en las personas, la confidencialidad y protección de la información y la transparencia en todas las etapas del proceso. Estos principios aseguran que la tecnología se emplee al servicio de la justicia y no en detrimento de los derechos, y se refuerce la legitimidad de la labor defensorial y la confianza de la ciudadanía en la Defensoría del Pueblo como institución nacional de derechos humanos.

Todo uso de la inteligencia artificial debe estar sujeto a supervisión humana, trazabilidad y transparencia verificable, de manera que cada resultado pueda ser revisado, explicado y justificado. La transparencia en el uso de estas tecnologías no solo refuerza la rendición de cuentas, sino que consolida la credibilidad de la Defensoría del Pueblo como institución nacional de derechos humanos.

Recordemos que la inteligencia artificial es un apoyo técnico, no un sustituto del razonamiento jurídico, la empatía ni la sensibilidad humana. Por ello, la calidad y precisión de los resultados dependerá directamente de la forma en que se formulen las solicitudes, del cuidado en su revisión y del compromiso con los principios éticos, humanos y de transparencia que orientan la función defensorial. A continuación, se presentan algunos consejos concretos para redactar *prompts* eficaces, supervisados y éticamente responsables.

Asimismo, se recomienda adoptar un formato orientador para registrar los *prompts* utilizados y los resultados obtenidos. Un esquema sencillo que incluya la fecha, la descripción de la tarea, el *prompt* empleado, la herramienta utilizada y las correcciones introducidas permitiría asegurar la trazabilidad del proceso, facilitar auditorías futuras y reforzar la transparencia en el uso institucional de la IA. Este registro también favorecería la revisión posterior de decisiones, la identificación de buenas prácticas y la mejora continua en el uso de estas tecnologías dentro de la labor defensorial. Por ejemplo, un defensor podría anotar: *prompt*: “Resumir en dos párrafos el marco normativo del delito de amenazas (sin datos personales)”. Correcciones humanas: ajuste de citas y contextualización de acuerdo con el caso real.

- **Comience con una acción clara y específica**

Evite formular preguntas vagas o ambiguas. Utilice verbos de acción como resumir, comparar, explicar, analizar, sintetizar o generar borrador para que el sistema entienda exactamente qué resultado espera.

Por ejemplo:

“Resumir los hechos del expediente en 150 palabras”.

Cuanto más precisa sea la instrucción, más útil y coherente será la respuesta.



- **Proporciona contexto pertinente y relevante**

Incluya información general del caso o antecedentes jurídicos, pero nunca datos personales o sensibles. Contextualizar permite que la herramienta comprenda mejor el propósito del texto y genere resultados más ajustados al marco legal.

Por ejemplo:

“Teniendo en cuenta la tipificación del delito de hurto contemplado en el artículo 239 del Código Penal colombiano, elaborar una síntesis jurídica del caso que identifique los elementos del tipo penal y la posible responsabilidad del procesado. Según el artículo 239 del Código Penal, modificado por el artículo 11 de la Ley 2197 de 2022, comete hurto quien se apodera de una cosa mueble ajena con el propósito de obtener provecho para sí o para otro, conducta sancionada con prisión de treinta y dos (32) a ciento ocho (108) meses. La norma establece además que la pena será de treinta y dos (32) a cuarenta y ocho (48) meses cuando la cuantía sea inferior a cuatro (4) salarios mínimos legales mensuales vigentes, y de cuarenta y ocho (48) a ciento ocho (108) meses cuando la cuantía sea igual o superior a dicho valor”.

Este tipo de contexto jurídico, sin incluir datos personales ni información sensible, permite que la inteligencia artificial genere respuestas más precisas, útiles y alineadas con el ordenamiento legal colombiano.

Cuando el contexto brindado es suficiente, se le debe indicar a la IA que se fije solo en dicha información y no establezca relaciones con otras búsquedas del propio modelo o en internet. De ser necesario también se puede indicar que haga búsquedas profundas en internet a partir de los asistentes de investigación:

“Con base exclusivamente en el texto del artículo 239 del Código Penal colombiano que se transcribe a continuación [...], organice una tabla con: (i) elementos normativos del tipo penal; (ii) preguntas que una persona defensora debería verificar en el expediente; y (iii) fuentes oficiales que deben consultarse. No determine responsabilidad, no valore pruebas, no agregue jurisprudencia ni complete información faltante”.

V. Transparencia y escalas de uso

- **Defina el rol desde el cual realiza la solicitud**

Indicar el rol profesional o institucional orienta la perspectiva de la respuesta. Las herramientas de IA adaptan el enfoque según el contexto.

Por ejemplo:

“Rol: defensor público de Colombia especializado en temas de familia”.

Esto permite que las sugerencias estén más alineadas con la función defensorial y el marco jurídico colombiano.

- **Precisa el formato o estructura deseada**

Indique claramente si espera un esquema, lista de viñetas, párrafo explicativo, modelo de memorial o texto argumentativo. Especificar la extensión o el tipo de lenguaje también mejora los resultados.

Por ejemplo:

“Elaborar un resumen en tres párrafos, en lenguaje claro y con enfoque de derechos, redactar un modelo de memorial breve con numeración y lenguaje jurídico técnico y presentar los argumentos principales en formato de lista con viñetas que destaquen los fundamentos constitucionales”.

Cuanto más precisas sean las instrucciones sobre la estructura y el estilo, más coherente y pertinente será la respuesta generada.

- **Haga iteraciones y ajustes según los resultados**

Si la primera respuesta no es adecuada, refine el *prompt*. Puede delimitar el enfoque, pedir ejemplos adicionales o solicitar una versión más breve o precisa. La interacción progresiva con la herramienta permite mejorar la calidad del resultado y mantener el control sobre la información generada.

- **Incluye instrucciones sobre el lenguaje o tono**

Indique si desea un lenguaje simple, técnico, jurídico o dirigido a una audiencia no experta. También puede aclarar el tono (formal, pedagógico, respetuoso, institucional). Esto ayuda a que la IA produzca textos acordes con el público y los fines comunicativos de la Defensoría del Pueblo.

- **Solicite fuentes y referencias verificables**

Cuando sea pertinente, pida que se incluyan citas completas —con número de sentencia, fecha y tribunal— para poder verificar la información posteriormente. Recuerde que las herramientas de IA pueden generar referencias inexactas o inexistentes, por lo que toda cita debe contrastarse con bases de datos oficiales o fuentes institucionales confiables

- **Mantenga trazabilidad de los *prompts* utilizados**

Conserve un registro de las instrucciones ingresadas (*prompts*) y de las respuestas obtenidas. Esto permite garantizar la transparencia del proceso, facilitar revisiones posteriores y cumplir con los estándares institucionales de rendición de cuentas y auditoría. Documentar el uso de la IA también contribuye a la mejora continua y al aprendizaje colectivo dentro de la Defensoría.

- **Considere el impacto ambiental de la tecnología**

Cada consulta a un sistema de inteligencia artificial implica un consumo significativo de energía y agua por el funcionamiento de los servidores y centros de datos. Antes de usar la IA, reflexione si la tarea lo justifica o si puede resolverse mediante un buscador tradicional, una base de datos jurídica o el propio análisis humano. Usar la IA con responsabilidad también significa cuidar el medioambiente: optimice su uso, evite preguntas repetitivas y valore siempre el criterio propio antes de acudir a la tecnología.



Copilot y otras herramientas de IA disponibles en los servicios institucionales

La Defensoría del Pueblo, como parte de su infraestructura tecnológica institucional, utiliza los servicios de Microsoft 365, los cuales incluyen, dentro de su entorno, el asistente virtual con inteligencia artificial Copilot. Este recurso forma parte de las herramientas tecnológicas provistas oficialmente por la Entidad para apoyar el cumplimiento de sus funciones misionales y administrativas.

Es importante precisar que la disponibilidad de esta función no implica la obligación de emplearla. Su uso constituye una posibilidad adicional orientada a facilitar tareas de redacción, sistematización o análisis de información, siempre bajo la supervisión de las y los servidores públicos y en estricto cumplimiento de los lineamientos éticos, técnicos y de seguridad definidos en esta guía.

La disponibilidad de una función dentro de Microsoft 365 o de otro servicio institucional no implica autorización para tratar cualquier tipo de información. Solo podrán utilizarse las versiones, funcionalidades y configuraciones expresamente habilitadas por la Defensoría del Pueblo, de acuerdo con el catálogo vigente de herramientas autorizadas y con la clasificación de la información.

Queda prohibido utilizar cuentas personales, suscripciones adquiridas individualmente, versiones gratuitas o herramientas no evaluadas para actividades institucionales. La información de casos solo podrá procesarse cuando la herramienta y la finalidad hayan sido autorizadas específicamente para esa clase de datos.

La eventual adopción de herramientas propias o especializadas deberá estar precedida por evaluación jurídica, técnica, de seguridad y de impacto en derechos humanos; definición de responsables; pruebas controladas; documentación; capacitación; monitoreo y mecanismos de reporte de incidentes. La referencia a iniciativas en fase piloto no constituye autorización general de uso hasta que la Entidad comunique formalmente su habilitación y alcance.

De manera paralela, la Defensoría del Pueblo avanza en el proceso de desarrollo e incubación de herramientas propias de inteligencia artificial, diseñadas específicamente para fortalecer la labor defensorial y optimizar los procesos internos de análisis y respuesta institucional. Estas iniciativas se encuentran actualmente en fase piloto, razón por la cual su implementación será gradual y acompañada de procesos de capacitación, monitoreo y evaluación continua.

A medida que la institución adopte la presente guía y consolide nuevas soluciones tecnológicas, los lineamientos aquí contenidos serán revisados y actualizados de manera periódica, con el propósito de incorporar las mejores prácticas en materia de uso seguro, responsable, ético y transparente de la inteligencia artificial en el ejercicio de la función pública defensorial.

VI. Anexo práctico: *prompts* seguros y lista de chequeo

Un *prompt* claro puede mejorar la utilidad de una salida, pero no elimina los riesgos de alucinación, sesgo, exposición de datos o dependencia. La siguiente plantilla solo debe utilizarse en herramientas autorizadas y con información que pueda tratarse conforme con la política institucional.

Componente	Instrucción recomendada
Tarea	Indique una acción concreta: resumir, comparar, ordenar, revisar claridad o crear un formato vacío.
Fuente	Limite la respuesta al texto suministrado o a fuentes oficiales expresamente indicadas.
Límites	Solicite no inventar normas, hechos, citas ni completar información faltante.
Datos	No incluya nombres, identificadores, información sensible, reservada ni elementos de la estrategia.
Formato	Defina extensión, estructura, audiencia y lenguaje.
Verificación	Solicite señalar afirmaciones que requieren contraste en una fuente oficial.

Otros ejemplos:

Área 1 — Resúmenes de expedientes [nivel AMARILLO — con verificación]

Prompt 1.1 — Resumen de hechos: [VERDE — Solo texto, sin datos personales] «Eres un asistente jurídico especializado en derecho penal colombiano. Resume en máximo 200 palabras los hechos del siguiente caso, en orden cronológico, donde destagues los elementos jurídicamente relevantes para la defensa. El texto no contiene datos personales: [PEGAR TEXTO ANONIMIZADO].»

Prompt 1.2 — Línea de tiempo: [AMARILLO — Verificar resultado vs. expediente físico] «A partir del siguiente relato de hechos [TEXTO ANONIMIZADO], elabora una línea de tiempo con las fechas y eventos principales, en formato de lista ordenada cronológicamente.»

Área 2 — Búsqueda de jurisprudencia [nivel AMARILLO — verificación en fuente primaria obligatoria]

Prompt 2.1 — Línea jurisprudencial: [AMARILLO — Verificar cada resultado en Relatoría Corte Constitucional] «Identifica las sentencias de la Corte Constitucional colombiana más relevantes sobre el derecho a la defensa técnica en procesos penales de menores de edad. Para cada sentencia indica: número, año, magistrado ponente y ratio decidendi principal.» Recuerde que debe verificar cada sentencia en la Relatoría oficial antes de citarla.

Prompt 2.2 — Resumen de criterio: [AMARILLO — Verificar vigencia y aplicabilidad] «Resume en tres párrafos el criterio de la Corte Suprema de Justicia colombiana sobre el principio de congruencia en materia penal, a partir de la jurisprudencia disponible hasta 2024. Indica si existe algún hito jurisprudencial que haya modificado la posición.»

Área 3 — Redacción de escritos [niveles AMARILLO y VERDE según contenido]

Prompt 3.1 — Borrador de petición: [AMARILLO — Revisar exhaustivamente antes de firmar] «Actúa como defensor público colombiano especializado en derecho de familia. Elabora un borrador de derecho de petición dirigido al ICBF donde se solicite información sobre el proceso de restablecimiento de derechos al niño A (anonimizado). El escrito debe incluir: saludo formal, hechos en tres párrafos, fundamento jurídico (Ley 1098 de 2006), solicitud concreta y cierre formal.»

Prompt 3.2 — Mejora de redacción: [VERDE — Texto sin datos personales, revisión de forma] «El siguiente texto es un borrador de oficio institucional [TEXTO SIN DATOS PERSONALES]. Mejora la redacción para que sea más clara, formal y coherente, sin cambiar el sentido ni agregar información nueva.»

Área 4 — Preparación de audiencias [nivel AMARILLO — sin datos reales del caso]

Prompt 4.1 — Estructura argumentativa: [AMARILLO — Solo con caso hipotético o anonimizado] «Actúa como defensor público en derecho penal. Estructura los argumentos principales para sostener la atipicidad de la conducta de hurto en un caso hipotético donde el procesado actuó bajo estado de necesidad. Organiza los argumentos por relevancia jurídica e indica qué normas y sentencias podrían sustentarlos.»

Prompt 4.2 — Simulación de preguntas: [AMARILLO — Solo escenario ficticio] «Simula las cinco preguntas más probables que un juez penal haría a la defensa durante una audiencia de solicitud de libertad provisional. Incluye también las posibles contrapreguntas de la Fiscalía.»

Prompts PROHIBIDOS en todos los casos [nivel ROJO]

-  «Resume el expediente del señor [NOMBRE REAL], cédula [NÚM.].» — contiene datos personales identificables directos.
-  «Redacta la estrategia de defensa para el caso de [NOMBRE REAL].» — la estrategia de defensa corresponde exclusivamente al criterio profesional del defensor público.
-  «¿Debo pedir la preclusión o ir a juicio en el caso de [NOMBRE REAL]?» — delegación indebida de decisión jurídica a la IA.
-  «Analiza estas fotos del expediente para identificar al acusado.» — datos biométricos e información visual sensible

Ejemplos según nivel de riesgo

Nivel	Ejemplo
Uso de menor riesgo	“Mejore la claridad y la ortografía del siguiente párrafo institucional, sin alterar su significado. El texto no contiene información personal ni reservada. Señale los cambios realizados”.
Uso condicionado	“Con base exclusivamente en los artículos transcritos, organice un cuadro comparativo. No agregue normas ni jurisprudencia externa. Indique cualquier ambigüedad que requiera revisión humana”.
Uso no permitido	“Analice este expediente completo, determine qué testigo no es creíble y diseñe la estrategia de defensa”. Este uso es inadmisibles, porque involucra información confidencial, valoración probatoria y delegación de la estrategia.

Lista de chequeo previa

1. ¿La herramienta y la funcionalidad están autorizadas para esta finalidad?
2. ¿La información está libre de datos personales, sensibles, reservados y elementos de la estrategia?
3. ¿Se evaluó el riesgo de reidentificación?
4. ¿La tarea puede recibir apoyo sin delegar valoración, decisión o estrategia?
5. ¿El resultado puede revisarse integralmente y contrastarse con fuentes oficiales?
6. ¿Se documentará el uso sin reproducir información protegida?
7. ¿El uso es necesario, proporcional y más adecuado que una alternativa convencional?



VII. Gobernanza, actualización y control de versiones

La Guía será objeto de revisión ordinaria anual. También deberá revisarse de manera extraordinaria cuando se expida una norma o sentencia relevante, cambien las herramientas autorizadas, se modifiquen términos de uso o políticas de tratamiento, ocurra un incidente de seguridad, se identifique un riesgo significativo o el pilotaje evidencie dificultades de aplicación.

La coordinación sustantiva corresponderá a la DADLE, con participación de la Dirección Nacional de Defensoría Pública, la Dirección Nacional de Promoción y Divulgación, las áreas o dependencias con competencias en temas de tecnología, seguridad y privacidad de la información, protección de datos y gestión documental, así como representantes del público objetivo.

Las actualizaciones deberán incluir validación de comprensión y aplicabilidad con defensores públicos de diferentes regiones y áreas de práctica.

Versión	Fecha	Cambio principal	Dependencias que revisaron
1.0	2026	Publicación inicial consolidada; incorporación de enfoque diferencial, ruta de decisión, anexo de <i>prompts</i> y protocolo de actualización.	DADLE, DNDP, Dirección Nacional de Promoción y Divulgación y dependencias técnicas competentes.

Canal para observaciones y reportes: ambientes.digitales@defensoria.gov.co



VIII. Recursos adicionales

El desarrollo ético y seguro de la inteligencia artificial en el ámbito jurídico exige un proceso continuo de formación y actualización. Esta sección pone a disposición un conjunto de recursos normativos, técnicos y pedagógicos para que las y los defensores públicos profundicen en principios, estándares y herramientas que regulan el uso de la IA en la administración de justicia. Estos materiales complementan la guía y constituyen referencias esenciales para fortalecer capacidades institucionales y personales con enfoque de derechos humanos, transparencia y responsabilidad profesional.

Cómo utilizar estos recursos

1. Diagnosticar necesidades: identificar casos de uso, riesgos y brechas de conocimiento en la unidad o despacho.
2. Aplicar el marco normativo: contrastar cada caso con los estándares aquí listados, definir salvaguardas y controles.
3. Capacitar y documentar: incorporar los recursos en planes de formación, registrar decisiones y evaluar resultados.

Recursos normativos y jurisprudenciales

- Ley 1581 de 2012. Por la cual se dictan disposiciones generales para la protección de datos personales.
- Directrices de la UNESCO para el uso de sistemas de IA en juzgados y tribunales (2024). Principios internacionales sobre ética, transparencia, debida diligencia y respeto por los derechos humanos en contextos judiciales; incluyen pautas para evaluación de impacto y explicabilidad.
- Acuerdo PCSJA24-12243 del Consejo Superior de la Judicatura (2024). Lineamientos oficiales para un uso ético y responsable de la IA en la Rama Judicial; define roles, controles, trazabilidad y criterios de rendición de cuentas.
- Sentencia T-323 de 2024 (Corte Constitucional). Criterios orientadores sobre uso de IA generativa en actuaciones judiciales y administrativas; refuerza exigencias de motivación suficiente, control humano significativo y respeto por el debido proceso.
- Sentencia T-067 de 2025 (Corte Constitucional). Desarrollo de estándares de transparencia algorítmica y acceso a la información frente a sistemas automatizados utilizados por entidades públicas; útil para sustentar obligaciones de explicabilidad y entrega de insumos relevantes.

Nota de cumplimiento: toda utilización de herramientas de IA en funciones defensoriales debe mantenerse desde supervisión humana, con motivación suficiente, registro de decisiones y respeto por los principios de legalidad, necesidad, proporcionalidad, no discriminación, transparencia y rendición de cuentas. La Defensoría del Pueblo revisará y actualizará esta guía de manera periódica para reflejar mejores prácticas y nuevas disposiciones aplicables.

Contacto: para preguntas o solicitudes, escribir a la Delegada para la Protección de Derechos en Ambientes Digitales y Libertad de Expresión: ambientes.digitales@Defensoria.gov.co .



IX. Referencias

Normativa y política pública colombiana

Congreso de la República de Colombia (2012). Ley 1581 de 2012, por la cual se dictan disposiciones generales para la protección de datos personales.

Consejo Nacional de Política Económica y Social (2025). Documento CONPES 4144: Política Nacional de Inteligencia Artificial.

Consejo Superior de la Judicatura (2024). Acuerdo PCSJA24-12243 del 16 de diciembre de 2024, por el cual se adoptan lineamientos para el uso responsable, seguro y ético de la IA en la Rama Judicial.

Corte Constitucional de Colombia (2024). Sentencia T-323 de 2024.

Corte Constitucional de Colombia (2025). Sentencia T-067 de 2025.

Defensoría del Pueblo y Procuraduría General de la Nación (2025). Directiva Conjunta 007 de 2025: Estándares mínimos sobre transparencia algorítmica.

Presidencia de la República, Ministerio de Tecnologías de la Información y las Comunicaciones, DNP, Ministerio de Ciencia, Tecnología e Innovación, Comisión de Regulación y las Comunicaciones (2026). Guía ética para la implementación, desarrollo y uso de sistemas de inteligencia artificial en entidades públicas de Colombia.

Estándares internacionales y accesibilidad

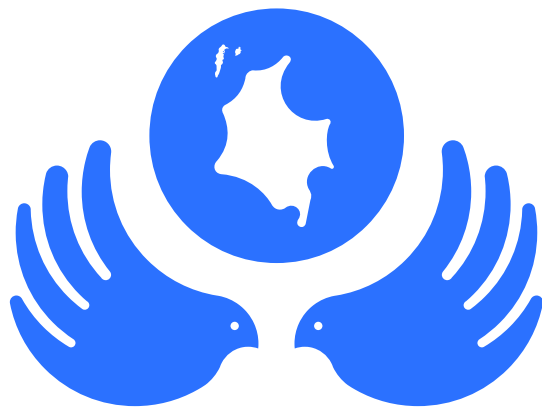
Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2025). Guidelines for the use of AI systems in courts and tribunals [Directrices para el uso de sistemas de IA en juzgados y tribunales]. UNESCO.

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2021). Recomendación sobre la ética de la inteligencia artificial. UNESCO.

World Wide Web Consortium (2024). Web Content Accessibility Guidelines (WCAG) 2.2.

Literatura académica

- Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 77–91.
- Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances*, 4(1), eaao5580.
- Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1), 121–127.
- Vinuesa, R., Azizpour, H., Leite, I., et al. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications*, 11, 233.



**Defensoría
del Pueblo**

C O L O M B I A

y la Naturaleza