



DIRECTRICES PARA EL

USO DE SISTEMAS DE IA EN CORTES Y TRIBUNALES

Publicado en 2026 por la Organización de las Naciones Unidas para la Educación,
la Ciencia y la Cultura, 7, place de Fontenoy, 75352 París 07 SP, Francia

© UNESCO 2026

ISBN 9789233002746

DOI: 10.58338/UKER1775



Esta publicación está disponible en acceso abierto bajo la licencia Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) (<https://creativecommons.org/licenses/by-sa/3.0/igo/>). Al utilizar el contenido de la presente publicación, los usuarios aceptan las condiciones de utilización del Repositorio UNESCO de acceso abierto (<https://www.unesco.org/es/open-access/cc-sa>)

Título original: Guidelines for the Use of AI Systems in Courts and Tribunals

Publicado en 2025 por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura

Los términos empleados en esta publicación y la presentación de los datos que en ella aparecen no implican toma alguna de posición de parte de la UNESCO en cuanto al estatuto jurídico de los países, territorios, ciudades o regiones ni respecto de sus autoridades, fronteras o límites.

Las ideas y opiniones expresadas en esta obra son las de los autores y no reflejan necesariamente el punto de vista de la UNESCO ni comprometen a la Organización.

Las Directrices fueron preparadas por:

Juan David Gutiérrez, PhD, Profesor Asociado, Universidad de los Andes, Colombia.

Revisores de la UNESCO:

Cedric Wachholz, Elvis Michel Kenmoe, Guilherme Canela De Souza, Ikran Abdirahman, Jaco Du Toit, Mehdi Benchelah, Prateek Sibal, Kamel El Hilali.

Una lista de revisores externos y consultas está disponible en el anexo (p.31).

Gestión y coordinación de proyectos:

Prateek Sibal, Especialista de Programas, Políticas Digitales y Transformación Digital, UNESCO.

Ikran Abdirahman, consultor, Políticas Digitales y Transformación Digital, UNESCO.

Gustavo Fonseca Ribeiro, Consultor, Políticas Digitales y Transformación Digital, UNESCO.

Mei Jie Chen, Becaria, Políticas Digitales y Transformación Digital, UNESCO.

Dr. Kamel El Hilali, Especialista en IA y Estado de Derecho, Políticas Digitales y Transformación Digital, UNESCO.

Traducción : Juan David Gutiérrez

Fotografía de cubierta / Diseño gráfico / Diseño de cubierta / Ilustraciones / Maquetación:

Abracadabra Estudio de Diseño, México

Impresión: UNESCO

Impreso en Francia

Cómo citar: UNESCO, Directrices para el Uso de Sistemas de IA en Cortes y Tribunales, 2026



**Financiado por
la Unión Europea**

Las Directrices se han desarrollado como parte del proyecto financiado por la Comisión Europea "Apoyar a los Estados Miembros en la Implementación de la Recomendación de la UNESCO sobre la Ética de la IA mediante herramientas innovadoras."

RESUMEN BREVE

INTELIGENCIA ARTIFICIAL EN LOS TRIBUNALES: DE LA IDEA A LA REALIDAD

La inteligencia artificial está transformando los sistemas de justicia en todo el mundo, ofreciendo nuevas posibilidades para mejorar el acceso a la justicia, agilizar la gestión de casos y apoyar la toma de decisiones judiciales. Sin embargo, estas innovaciones también plantean complejos desafíos éticos y de derechos humanos. Las Directrices de la UNESCO para el Uso de Sistemas de IA en Cortes y Tribunales presentan el primer marco ético y operativo global para garantizar que la IA sirva a la justicia mientras defiende el Estado de derecho y los derechos fundamentales.

Construidas en torno a quince principios universales —desde la transparencia, la rendición de cuentas y la supervisión humana hasta la protección de los derechos humanos y la gobernanza multiactor—, las Directrices ofrecen una orientación práctica para jueces, administradores judiciales y responsables políticos que exploran la adopción de la IA. Defienden la IA como una herramienta asistencial, no sustitutiva, utilizada de forma responsable y siempre bajo supervisión humana significativa.

Complementadas por un conjunto de herramientas para el desarrollo de capacidades, las Directrices se basan en consultas con expertos de más de 160 países, involucrando a más de 36.000 actores judiciales. A medida que la IA se convierte en parte integral de los sistemas judiciales, esta publicación ofrece orientación oportuna para garantizar que la tecnología fortalezca la justicia, la transparencia y la confianza, construyendo las defensas de la paz en la mente de hombres y mujeres.

15

principios universales
para guiar la adopción,
gobernanza y supervisión
de la IA en el poder judicial.

DIRECTRICES PARA EL

USO DE SISTEMAS DE IA EN CORTES Y TRIBUNALES

Índice

Las Directrices en una página	6
Introducción.....	7
Objetivos de las Directrices	11
Metodología.....	14
Directrices para el uso de sistemas de IA en Cortes y Tribunales .	17
1. Principios	17
2. Directrices específicas para organizaciones judiciales	21
2.1. Sobre el desarrollo, adquisición, uso y evaluación de sistemas de IA.....	21
2.2. Talento humano	25
2.3. Uso de sistemas de IA generativa	26
3. Orientación específica para los miembros individuales del poder judicial	27
3.1. Antes de usar sistemas de IA	27
3.2. Uso de sistemas de IA.....	27
3.3. Uso de sistemas de IA generativa	28
Conclusión.....	32
Anexo	33
Notas finales.....	35

Las Directrices en una página

“Las Directrices establecen 15 principios universales para orientar la adopción, la gobernanza y la supervisión de IA en el sistema judicial”

La Inteligencia Artificial (IA) está cada vez más integrada en los sistemas judiciales de todo el mundo, creando oportunidades para mejorar el acceso a la justicia, mejorar la gestión de casos y apoyar la toma de decisiones judiciales. Sin embargo, su uso también plantea importantes desafíos éticos, legales y de derechos humanos.

Para abordar estos temas, la UNESCO ha desarrollado las Directrices para el Uso de Sistemas de IA en Cortes y Tribunales, el primer marco ético y operativo orientado globalmente para la IA en contextos judiciales. Las Directrices tienen como objetivo garantizar que las aplicaciones de IA sirvan a la causa de la justicia, al tiempo que salvaguardan el estado de derecho y los derechos fundamentales.

Se fundamentan en quince principios universalmente aplicables que guían la adopción, gobernanza y supervisión de la IA en el poder judicial: la protección de los derechos humanos, la proporcionalidad, la viabilidad de los beneficios, la seguridad, la seguridad de la información, la precisión y fiabilidad, la explicabilidad, la auditabilidad, la transparencia, la conciencia y el uso informado, la responsabilidad, la rendición de cuentas (accountability) y la contestabilidad, la supervisión y toma de decisiones humanas, el diseño participativo y centrado en el humano, y gobernanza multiactor. En conjunto, estos principios proporcionan una dirección práctica para los líderes judiciales, jueces y personal judicial que consideran soluciones de IA. La IA no se concibe como un sustituto del razonamiento judicial, sino como una herramienta de apoyo para ampliar el acceso a la información, mejorar la eficiencia y apoyar la justicia equitativa, siempre bajo supervisión humana significativa y revisión ética.

Las Directrices se complementan con un conjunto completo de herramientas para el desarrollo de capacidades. A través del programa de la UNESCO para el poder judicial sobre la libertad de expresión y la seguridad de los periodistas, y el programa IA y Estado de Derecho, más de 36.000 actores judiciales en más de 160 países han participado, reforzando la conciencia tanto sobre los posibles beneficios como sobre los riesgos de la IA. El desarrollo de las Directrices se basó en consultas globales con las partes interesadas, asegurando que sus recomendaciones reflejen diversas tradiciones legales, capacidades tecnológicas y necesidades sociales.

A medida que la IA gana protagonismo en los tribunales, los sistemas legales se encuentran en un momento crucial. Las decisiones tomadas hoy sobre su despliegue moldearán la administración de justicia durante las próximas décadas. Ya seas un profesional judicial, un académico jurídico, un desarrollador tecnológico o un ciudadano comprometido, las Directrices ofrecen un recurso valioso para informar y apoyar la integración responsable de la IA en los sistemas de justicia de todo el mundo.

Introducción

La transformación digital está transformando la forma en que operan todas las ramas del gobierno. Los gobiernos de todo el mundo están adoptando la inteligencia artificial (IA) con distintos grados de intensidad y enfoque.¹ Por ejemplo, organismos públicos en América Latina y el Caribe están implementando o pilotando alrededor de 700 herramientas de IA para llevar a cabo diversas funciones gubernamentales.² El despliegue de sistemas de IA puede ayudar a mejorar la prestación de servicios públicos y automatizar procesos administrativos, potenciar la participación ciudadana, monitorear la infraestructura pública, optimizar la asignación de recursos y apoyar decisiones de política pública basadas en la evidencia en los sectores de la educación, la sanidad, la movilidad, la agricultura, la seguridad y el medio ambiente.

La adopción de herramientas de IA en el sector público exige marcos de gobernanza que equilibren la innovación con la protección de los derechos humanos, asegurando que el avance tecnológico sirva a los valores democráticos y contribuya al desarrollo sostenible.³

Las cortes y tribunales de todo el mundo están cada vez más interesados en adoptar herramientas de IA (Recuadro nº 1)⁴ para apoyar una variedad de tareas judiciales, como la búsqueda, revisión y descubrimiento de documentos a gran escala; resolución automatizada de disputas en línea; predicción de eventos futuros (por ejemplo, resultados de litigios); toma de decisiones y procesos de resolución; anonimización y pseudo-anonimización de decisiones judiciales; clasificación de documentos presentados; transcripción (voz y texto) y traducción (tanto voz como texto).⁵

Recuadro nº 1

CASOS DE USO DE LA IA POR PARTE DE TRIBUNALES Y DESPACHOS DE ABOGADOS

- El Tribunal Supremo brasileño implementó VICTOR, un sistema que tramita miles de recursos presentados ante el tribunal y facilita la identificación de casos que cumplen con el requisito de "repercusiones generales".
- El Tribunal Supremo de la India desplegó SUVAS, un software que traduce miles de documentos del inglés a diez idiomas locales y viceversa.
- En la Unión Europea (UE), los despachos de abogados y los jueces utilizan herramientas de procesamiento del lenguaje natural para predecir decisiones judiciales (por ejemplo, del Tribunal de Justicia de la UE).
- En Sudáfrica y Zimbabue, los despachos adoptaron herramientas de IA para la revisión de contratos, gestión e investigación de casos.

Fuentes: Aletras et al. (2016), Aneja y Mathew (2023), CEPEJ (2024), Kufakwababa (2021), Medvedeva et al. (2020), Ministro do Superior Tribunal de Justiça (2020) y UNESCO (2023).

Estas aplicaciones de sistemas de IA en el poder judicial tienen el potencial de mejorar la eficiencia, la eficacia y el acceso a la justicia. Además, el uso de herramientas de IA fiables podría reducir el “ruido” en el sistema judicial, que se refiere a la variabilidad no deseada en sentencias que deberían ser idénticas, tal como la definieron el difunto ganador del Nobel Daniel Kahneman y sus colaboradores Olivier Sibony y Cass R. Sunstein.⁶

Más recientemente, jueces, personal de apoyo judicial, fiscales y abogados de todo el mundo han empezado a utilizar chatbots impulsados por Modelos de Lenguaje de Gran Escala (LLMs) para redactar documentos legales, dictar decisiones judiciales y elaborar argumentos en audiencias (*Recuadro n° 2*).⁷ Además, estos chatbots han sido utilizados por jueces y tribunales en diversas áreas del derecho: constitucional, administrativa, de privacidad, penal y familiar, entre otras.⁸

Recuadro n° 2 CASOS DE USO DE LA IA GENERATIVA POR PARTE DE JUECES Y ABOGADOS

- En Argentina, los jueces han utilizado los LLM para resumir sus decisiones en un lenguaje sencillo y accesible.
- Despachos de abogados, empresas de servicios legales y universidades han desarrollado sistemas de IA generativa basados en LLMs, de forma independiente o con empresas tecnológicas, para realizar investigaciones legales y trabajos de litigio, añadir contexto a un caso, resumir textos legales y redactar correos electrónicos y contratos. Estas herramientas no se limitan a los LLM comerciales de uso general; jueces y abogados también pueden utilizar LLMs diseñados exclusivamente para realizar actividades legales, así como otros LLMs de código abierto.
- No obstante, se han reportado casos en Australia, Brasil, Canadá, Sudáfrica, España y Estados Unidos sobre jueces y abogados que emitieron decisiones judiciales o presentaron documentos legales con referencias a sentencias inexistentes, debido al uso inadecuado de chatbots de IA.

Fuentes: Adams (2023), Ambrogi (2023), Bennetts et al. (2023), CEPEJ (2024), Gutiérrez (2024), Harden (2024), LexisNexis (2023), Merken (2025), Morcillo (2025), Taylor (2025), Terzidou (2025), Weiser (2023), Weiser y Bromwich (2023) y Witten (2023).

Sin embargo, la orientación formal dirigida a individuos u organizaciones sobre el uso adecuado de herramientas de IA en el sector de la justicia es escasa. Solo unas pocas autoridades judiciales han publicado principios, normas o directrices sobre el uso ético y responsable de las herramientas de IA.

Entre ellas se encuentran Argentina (2024), Australia (2024), Brasil (2020, 2025), Canadá (2023, 2024), Colombia (2024), Nueva Zelanda (2023), Singapur (2024), Reino Unido de Gran Bretaña e Irlanda del Norte (2023) y Estados Unidos (2024).⁹ Además, varios tribunales han emitido sentencias que establecen principios y normas para el uso de herramientas de IA por parte del poder judicial (*Recuadro n° 3*).

Recuadro n° 3

SENTENCIA DE LA CORTE CONSTITUCIONAL COLOMBIANA SOBRE EL USO DE SISTEMAS DE IA POR PARTE DEL PODER JUDICIAL

- En agosto de 2024, la Corte Constitucional de Colombia decidió sobre la constitucionalidad de una sentencia de segunda instancia relativa al derecho fundamental a la salud de un menor. La Corte examinó si el juez de segunda instancia había vulnerado el derecho fundamental al debido proceso al incorporar texto producido con ChatGPT en su fallo. El contenido creado con el chatbot consistió en respuestas a cuatro preguntas legales que el juez transcribió como parte de la motivación de su decisión.
- En su sentencia, la Corte citó la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial, describió sus valores y principios, y reconoció que su marco ético es “uno de los más consultados y difundidos en materia de IA a nivel mundial”. La sentencia también citó el Kit Global de Herramientas de la UNESCO sobre IA y Estado de Derecho para el Poder Judicial y reconoció que “suministra elementos de gran valor a través de los cuales orienta sobre las leyes, principios, normas y jurisprudencia internacional de derechos humanos relevantes en el uso ético de la IA, a efectos de mitigar sus posibles riesgos para los derechos humanos.”
- La Corte dictaminó que el uso de sistemas de IA por parte del poder judicial es constitucional siempre que la máquina no sustituya al juicio humano, especialmente en la toma de

decisiones judiciales. Los sistemas de IA pueden ser utilizados por el poder judicial para apoyar tareas de gestión (por ejemplo, toma de notas e informes), tareas documentales (por ejemplo, clasificación de documentos), funciones judiciales (por ejemplo, identificar la jurisprudencia pertinente) y corregir y sintetizar textos bajo la supervisión del usuario.

Además, la Corte ordenó que los miembros del poder judicial que desearan utilizar sistemas de IA para apoyar sus actividades judiciales debían seguir "principios de (i) transparencia, (ii) responsabilidad, (iii) privacidad, (iv) no sustitución de la racionalidad humana, (v) seriedad y verificación, (vi) prevención de riesgos, (vii) igualdad y equidad, (viii) control humano, (ix) regulación ética, (x) adecuación a buenas prácticas y estándares colectivos, (xi) seguimiento continuo y adaptación y (xii) idoneidad." (Fuente: Colombia, Corte Constitucional, Sentencia T – 323 / 2024).

Tras esta sentencia, Colombia decidió implementar el Borrador de Directrices de la UNESCO sobre el uso de la IA en tribunales y tribunales. Las Directrices de Colombia para el Uso Responsable y Seguro de la IA Generativa en el Poder Judicial fueron adoptadas en diciembre de 2024.

Las nuevas leyes sobre IA aumentan la necesidad de orientación en esta materia. Por ejemplo, el Reglamento de IA de la Unión Europea clasifica como "alto riesgo" aquellos sistemas de IA que están "destinados a ser utilizados por una autoridad judicial, o en su nombre, para ayudar a una autoridad judicial en la investigación e interpretación de hechos y de la ley, así como en la garantía del cumplimiento del Derecho a un conjunto concreto de hechos, o a ser utilizados de forma similar en una resolución alternativa de litigios".¹⁰ Esta clasificación desencadena obligaciones para los responsables del despliegue de estos sistemas de IA, como implementar un sistema de gestión de riesgos y garantizar la supervisión humana de las herramientas.¹¹

Además, los sistemas de IA integrados en los procesos de toma de decisiones pueden tener efectos en cascada en todo el sistema judicial. Por un lado, tales impactos pueden constituir externalidades negativas y violaciones de los derechos humanos. Por

ejemplo, el despliegue podría contribuir a discriminar sistemáticamente a individuos o grupos vulnerables o marginados, o a restringir el acceso al ecosistema digital y a los dispositivos pertinentes.¹² Por otro lado, los sistemas de IA para la gestión de los tribunales del poder judicial pueden beneficiar a los usuarios judiciales al garantizar que los juicios se celebren en un plazo razonable. Estos sistemas de IA pueden apoyar actividades previas al juicio (por ejemplo, automatizar el sistema de presentación de los tribunales), las audiencias judiciales (por ejemplo, la traducción automática) y los procedimientos posteriores a la sentencia (por ejemplo, anonimizar las decisiones judiciales).¹³ Por ello, se requieren nuevas medidas a nivel organizacional e individual para prevenir y mitigar los efectos adversos y mejorar la administración de justicia.

Además, los sistemas de IA son una pluralidad de tecnologías que pueden transformar y trastocar la práctica del derecho y la resolución judicial.¹⁴ Los cambios resultantes en el sector de la justicia no deben poner en riesgo los derechos humanos. La Resolución 78/265 adoptada por la Asamblea General de las Naciones Unidas (ONU) el 21 de marzo de 2024 enfatizó que "los derechos humanos y las libertades fundamentales deben ser respetados, protegidos y promovidos a lo largo del ciclo de vida de los sistemas de inteligencia artificial". Por lo tanto, el poder judicial, y más en general los Estados miembros de la ONU, "que se abstengan o dejen de usar sistemas de inteligencia artificial que sean imposibles de operar en consonancia con el derecho internacional o que supongan riesgos indebidos para el disfrute de los derechos humanos, en especial de quienes se encuentran en situaciones vulnerables, y reafirma que los derechos de las personas también deben estar protegidos en Internet, también durante el ciclo de vida de los sistemas de inteligencia artificial".¹⁵

Por último, el uso de sistemas de IA no debe obstaculizar la competencia, independencia e imparcialidad del poder judicial porque, como se establece en los considerados de los Principios de Conducta Judicial de Bangalore, estos atributos son esenciales para proteger los derechos humanos y porque "la aplicación de todos los demás derechos depende en último término de la correcta administración de la justicia."¹⁶



Las Directrices proponen principios y recomendaciones para adoptar, desplegar y utilizar herramientas de IA para apoyar la administración de justicia respetando los derechos humanos.

Objetivos de las Directrices

Las *Directrices de la UNESCO para el Uso de Sistemas de IA en Cortes y Tribunales* (en adelante, *Directrices*) ofrecen orientación sobre las medidas que el sector judicial podría considerar para mejorar sus capacidades en el contexto de la transformación digital. Las Directrices proponen principios y recomendaciones para adoptar, desplegar y utilizar herramientas de IA para apoyar la administración de justicia. Las Directrices proporcionan orientación tanto a nivel organizativo (tribunal) como a nivel individual (jueces, funcionarios judiciales) (Nota: a partir de ahora, nos referiremos colectivamente a estas organizaciones e individuos como el poder judicial).

Aunque las Directrices están diseñadas específicamente para orientar al poder judicial, podrían resultar pertinentes para la profesión jurídica en general. En este sentido, la mayoría de los principios y recomendaciones propuestos podrían adaptarse a fiscales, árbitros,¹⁷ abogados, colegios de abogados, funcionarios públicos e investigadores. En este sentido, organizaciones de la sociedad civil, como asociaciones profesionales y universidades, han publicado recientemente directrices y recomendaciones similares.¹⁸

Recuadro nº 4

DEFINICIONES CLAVE

- Sistema de IA: Sistemas computacionales "capaces de procesar datos e información de una manera que se asemeja a un comportamiento inteligente, y abarca generalmente aspectos de razonamiento, aprendizaje, percepción, predicción, planificación o control" (UNESCO 2022, 10).
- Sistema de IA generativa: Sistemas computacionales "que se comunican en lenguaje natural, capaces de responder a preguntas relativamente complejas y pueden crear contenido (proporcionar texto, imagen o sonido) tras una pregunta o instrucciones formuladas (prompt)" (CEPEJ-GT-CYBERJUST 2024, 2).
- Miembros individuales del poder judicial: magistrados, jueces, magistrados, funcionarios judiciales y personal de apoyo judicial.
- Operador judicial: Un profesional jurídico que participa en la administración de justicia o que tiene un papel legal activo en un proceso judicial, como jueces, personal de apoyo judicial, fiscales y abogados.
- Modelos de lenguaje (LM): Modelos de IA generativa "que procesan entradas textuales, conocidas como prompts, y generan salidas de texto basadas en ellas. Sus entradas, así como sus salidas, pueden estar en diferentes formatos de texto como lenguaje natural, texto tabulado o incluso código de programa" (BSI 2024, 7).
- Ciclo de vida de los sistemas de IA: Ciclo que incluye las etapas de pre-diseño, diseño, desarrollo, evaluación, prueba, despliegue,

uso, venta, adquisición, operación y desmantelamiento (Asamblea General de la ONU, resolución 78/265).

- Aprendizaje Automático (ML): "un conjunto de técnicas que permite a las máquinas aprender automáticamente usando patrones y deducciones en lugar de instrucciones directas de una persona" (UNESCO 2023, 21).
- Procesamiento de Lenguaje Natural (PLN): "una técnica de aprendizaje automático que analiza grandes cantidades de datos de texto o voz humana (transcritos o acústicos) en busca de propiedades específicas, como significado, contenido, intención, actitud y contexto" (UNESCO 2023, 21).
- Organizaciones del poder judicial: Órganos que gobiernan el poder judicial, los tribunales y los tribunales.
- Sistemas de IA seguros, protegidos y confiables: Sistemas de IA que "están centrados en las personas, son fiables, se pueden explicar, son éticos e inclusivos, respetan plenamente la promoción y la protección de los derechos humanos y el derecho internacional, mantienen la privacidad, están orientados al desarrollo sostenible y son responsables" (Asamblea General de la ONU, resolución 78/265).

Sin embargo, el uso indebido de los sistemas de IA puede minar la confianza de la sociedad en el sistema judicial. Los posibles impactos negativos requieren que el poder judicial implemente marcos de rendición de cuentas antes y después de desplegar sistemas de IA.

Por ello, las *Directrices* proponen orientaciones específicas para garantizar la integración por parte del poder judicial de sistemas de IA éticos y respetuosos de los derechos humanos.

Además, las *Directrices* incluyen recomendaciones específicas para sistemas de IA generativa. Estas herramientas pueden ayudar a mejorar la calidad del trabajo judicial facilitando la gestión judicial, la redacción de documentos, la automatización de tareas y el apoyo a los procesos de toma de decisiones. Además, los LLM pueden facilitar la interacción entre ordenadores y usuarios mediante el lenguaje natural.



Las Directrices son un documento vivo que requerirá actualizaciones periódicas para abordar el rápido avance de la tecnología de IA y sus crecientes aplicaciones en los sistemas judiciales.

Metodología

Las *Directrices* se basan en el extenso trabajo de la UNESCO con los operadores judiciales en la última década. También se basan en documentos de antecedentes elaborados por la UNESCO, en un proceso de consulta pública y, posteriormente, en un proceso de revisión por pares de los borradores de directrices por expertos en IA jurídica de todo el mundo, así como en fuentes secundarias y literatura especializada.

Primero, desde 2013, la UNESCO ha formado a más de 36.000 operadores judiciales en 160 países sobre temas diversos como la libertad de expresión, el acceso a la información y la aplicación e impacto de los sistemas de justicia con IA. Estos actores judiciales recibieron formación a través de una “serie de Cursos Masivos Abiertos en Línea (MOOCs), formación y talleres sobre el terreno para jueces del Tribunal Supremo y la publicación de varios kits de herramientas y directrices.”¹⁹

Además, más de 5.000 personas participaron en el MOOC global de la UNESCO sobre IA y Estado de Derecho, que se ofreció para reforzar “las capacidades de los operadores judiciales en todo el mundo para proteger los derechos humanos, los principios democráticos y el Estado de derecho en relación con el uso de la IA.”²⁰ Además, a finales de 2023 se realizó una encuesta global sobre el uso de sistemas de IA por parte de los operadores judiciales. Los resultados ofrecieron perspectivas sobre cómo los operadores judiciales acceden a y utilizan los sistemas de IA, y destacaron sus necesidades en cuanto a formación y orientación.²¹

En conjunto, esta experiencia práctica reveló las expectativas y percepciones de los jueces respecto a la gobernanza de los sistemas de IA en las salas de juicio, lo que aportó sustancialmente a la redacción de las Directrices.

En segundo lugar, los documentos publicados por la UNESCO sobre la ética de la IA también fueron importantes para establecer los principios, valores y prácticas principales incluidos en las Directrices. Este es el caso de la UNESCO:

“Recomendación sobre la ética de la inteligencia artificial”²²

“Análisis de Impacto Ético: Una herramienta de la Recomendación sobre la Ética de la Inteligencia Artificial”;²³ y

“Kit de herramientas global sobre IA y el estado de derecho para el poder judicial”,²⁴ que proporcionó un marco sobre las aplicaciones y riesgos de los sistemas de IA en contextos judiciales.

En tercer lugar, la consulta pública de los proyectos de directrices en agosto y septiembre de 2024 permitió a la UNESCO reunir más de cien contribuciones, provenientes de individuos y organizaciones de 41 países, junto con comentarios colectivos de eventos organizados con agentes judiciales en diferentes regiones.²⁵

Cuarto, veinte revisores externos participaron en la revisión del borrador de las directrices. Este panel global diverso incluía jueces, juristas, abogados y especialistas en tecnología de África, Asia, Europa, Norte de África y Oriente Medio, América del Norte y Latinoamérica. Su experiencia complementaria —que abarca la práctica en tribunales, la investigación en tecnología jurídica, la gobernanza de la IA y la transformación digital de los sistemas legales— aseguró que las directrices fueran evaluadas críticamente desde múltiples perspectivas, lo que mejoró su precisión y practicidad.

Quinto, las Directrices consideraron los Principios de Conducta Judicial de Bangalore, diversas fuentes secundarias y literatura especializada, citadas en las notas finales, entre ellas principios y directrices publicados por autoridades judiciales, tribunales y organizaciones de la sociedad civil de diversas jurisdicciones, así como colegios y sociedades de abogados.

Por último, las Directrices son un documento vivo que requerirá actualizaciones periódicas para abordar el rápido avance de la tecnología de IA y sus crecientes aplicaciones en los sistemas judiciales.



Las Directrices tienen como objetivo promover el respeto universal a la justicia y al Estado de derecho, asegurando, al mismo tiempo, la protección y la promoción de los derechos humanos y las libertades fundamentales.

Directrices para el uso de sistemas de IA en Cortes y Tribunales

Las *Directrices* tienen como objetivo promover el respeto universal a la justicia y al Estado de derecho, asegurando, al mismo tiempo, la protección y la promoción de los derechos humanos y las libertades fundamentales.

1. Principios

Las Directrices ofrecen principios y recomendaciones específicas para organizaciones judiciales e individuos que implementan sistemas de IA, incluida la IA generativa. En este sentido, los “principios” se refieren a prescripciones generales destinadas a orientar a organizaciones e individuos en el desarrollo, la adquisición y el uso de sistemas de IA de manera ética y respetuosa con los derechos humanos. Las Directrices desarrollan 15 principios para alcanzar este objetivo (Figura nº 1).



Figura nº 1. Principios para el desarrollo, adquisición, adopción, despliegue y uso de sistemas de IA por parte del poder judicial

Las organizaciones y los miembros individuales del poder judicial deben seguir estos principios a lo largo de todo el ciclo de vida de los sistemas de IA:



1.1. Protección de los derechos humanos: Toma las medidas necesarias para respetar, proteger y promover los derechos humanos y el estado de derecho a lo largo de todo el ciclo de vida de los sistemas de IA.²⁶ Deben existir salvaguardas eficaces respecto de los sistemas de IA que puedan representar un alto riesgo para los derechos humanos, como la implementación de las evaluaciones de rendición de cuentas descritas en la sección 2 de las Directrices. El despliegue de sistemas de IA debe proteger especialmente los derechos de las mujeres y las niñas, las personas con discapacidad, los niños, los refugiados, los migrantes, los apátridas, las minorías y otras poblaciones en situación de marginación y/o vulnerabilidad. Aunque todos los derechos humanos pueden verse potencialmente afectados por ciertos sistemas de IA a lo largo del ciclo de vida de estas herramientas, los siguientes son especialmente importantes en el contexto de los procedimientos judiciales:

a. No discriminación: Adoptar, desplegar y utilizar sistemas de IA que busquen alcanzar objetivos mediante procesos que salvaguarden la equidad y garanticen un acceso inclusivo a la tecnología²⁷. Prevenir el desarrollo y las aplicaciones sesgadas de sistemas y resultados de IA que puedan agravar, reproducir, reforzar o perpetuar la discriminación basada en raza, color, ascendencia, género, edad, idioma, religión, opinión política, condición nacional, étnica, social o económica, discapacidad y cualquier otro motivo.²⁸

b. Igualdad: Garantizar el acceso y el trato igualitario para todos ante los tribunales, independientemente de sus capacidades digitales.²⁹ Por ejemplo, si las herramientas de IA median el acceso a la justicia, pueden surgir barreras adicionales para quienes carecen de recursos o habilidades digitales. Por lo tanto, la implementación de herramientas de IA en lo judicial debería tener en cuenta a quienes no pueden permitirse estas tecnologías, carecen de alfabetización digital o enfrentan otras barreras tecnológicas.³⁰ Para ello, es pertinente seguir el principio de No Dejar a Nadie Atrás (LNOB), desarrollado como parte de la Agenda de las Naciones Unidas para el Desarrollo Sostenible³¹, para abordar la exclusión digital.

c. Equidad procesal: Evaluar las implicaciones de los sistemas de IA para garantizar la equidad procedimental a lo largo de todo su ciclo de vida y evitar despliegues que vulneren este derecho.

d. Derecho a la privacidad y protección de datos personales: Adoptar, desplegar y utilizar sistemas de IA que protejan la privacidad y los datos personales tratados en el marco de la administración de justicia. El poder judicial debería evitar utilizar herramientas de IA de forma que generen riesgos de divulgación de dichos datos o permitan el acceso no autorizado por parte de terceros.³²

e. Libertad y seguridad: Garantizar que las personas no sean detenidas basándose en decisiones asistidas por IA de las que no tienen conocimiento, que no puedan entender cómo el sistema de IA tomó su decisión o que no puedan impugnar o apelar el proceso de toma de decisiones o la decisión en sí misma.



1.2. Proporcionalidad: Adoptar, desplegar y utilizar sistemas de IA que tengan como objetivo lograr fines legítimos y proporcionales en el contexto de su uso.³³



1.3. Viabilidad de los beneficios: Antes de adoptar sistemas de IA, el poder judicial debe evaluar tanto los beneficios potenciales específicos (para el poder judicial, los usuarios y el público) como su capacidad institucional para generarlos.³⁴



1.4. Seguridad: Adoptar, desplegar y utilizar sistemas de IA que eviten, aborden, prevengan, mitiguen o eliminen daños no deseados para las partes, los jueces, los tribunales a los que pertenecen, la reputación y autoridad del poder judicial, y para el público.³⁵



1.5. Seguridad de la información: Adoptar, desplegar y utilizar sistemas de IA que protejan la información confidencial conforme a los estándares internacionales de acceso a la información, establecer salvaguardas frente a amenazas cibernéticas y garantizar el control continuo del poder judicial sobre sus datos.³⁶



1.6. Precisión y fiabilidad: Adoptar, desplegar y utilizar sistemas de IA precisos, es decir, sistemas de IA que puedan proporcionar información útil y relevante y producir resultados y predicciones correctas,³⁷ y sistemas de IA fiables, es decir, sistemas de IA que funcionen “correctamente con una variedad de entradas y en una variedad de situaciones”.³⁸



1.7. Explicabilidad: Adoptar, desplegar y utilizar sistemas de IA que puedan explicar la razón de ser detrás de sus salidas y decisiones y cómo emplean las entradas para tales fines. La explicabilidad se refiere a hacer inteligible y proporcionar información –tanto a quienes despliegan la herramienta como a sus usuarios– sobre las entradas, razonamientos, resultados y funcionamiento de los sistemas de IA, haciendo que estos elementos sean comprensibles, interpretables y rastreables para los humanos.³⁹



1.8. Auditabilidad: Adoptar medidas administrativas, legales y humanas para asegurar y permitir que los sistemas de IA puedan ser auditados durante su despliegue.⁴⁰



1.9. Justicia transparente y abierta: Adoptar, desplegar y utilizar sistemas de IA que sean transparentes en cuanto a cómo se desarrollaron, cómo funcionan, sus datos de entrenamiento, sus limitaciones (incluido el margen de error), sus capacidades y el propósito de los sistemas.⁴¹ Informar de manera adecuada y oportuna cuándo y cómo se despliegan, adquieren, despliegan y utilizan los sistemas de IA y cómo funcionan estas herramientas, especialmente cuando las decisiones tomadas con o basadas en ellas puedan afectar los derechos y libertades de individuos o comunidades.⁴² Asegurarse de que las personas sean conscientes de que pueden impugnar dichas decisiones.



1.10. Concienciación y uso informado: Comprender las funcionalidades, tipos de uso, posibles impactos, limitaciones y riesgos de los sistemas de IA disponibles, para tomar decisiones informadas sobre su implementación; conocer el propósito previsto del uso de un sistema de IA específico para llevar a cabo actividades judiciales; y mantenerse al día de los avances tecnológicos.⁴³



1.11. Responsabilidad: Las organizaciones e individuos que desarrollen y utilicen sistemas de IA deben asumir la responsabilidad de las decisiones y acciones tomadas con el apoyo de herramientas de IA, sin perjuicio de la posible responsabilidad del proveedor en caso de que el sistema de IA sea defectuoso.⁴⁴



1.12. Rendición de cuentas y contestabilidad: Garantizar la rendición de cuentas, informando y explicando cómo y por qué ciertas herramientas de IA fueron adoptadas por el poder judicial, así como los resultados obtenidos (o no) con su despliegue. Además, introduce mecanismos que permitan a las partes afectadas impugnar y contrainterrogar cualquier resultado producido por un sistema de IA que pueda influir en el resultado del caso en cuestión, así como impugnar las decisiones y responsabilizar a los operadores judiciales por sus errores.⁴⁵



1.13. Supervisión y toma de decisiones humanas: Los jueces y magistrados no deben delegar ninguna parte de su mandato ni depender exclusivamente de los sistemas de IA para adoptar decisiones o automatizar procesos completos que puedan afectar negativamente a los derechos de individuos o comunidades. Deben preservar la posibilidad de intervención humana en todos los procesos de toma de decisiones en los que se desplieguen los sistemas de IA.⁴⁶ Las herramientas de IA no deben sustituir el análisis independiente de hechos, leyes y pruebas por parte del responsable de la toma de decisiones.



1.14. Diseño participativo y centrado en el ser humano: El desarrollo, despliegue y uso de sistemas de IA deben seguir principios de diseño centrado en el ser humano para complementar y ampliar las capacidades del poder judicial y respetar la dignidad y la autonomía humanas. El proceso de diseño debe incluir a miembros del poder judicial y al público para garantizar que el proceso sea inclusivo.⁴⁷



1.15. Gobernanza y colaboración multiactor: Las organizaciones judiciales deben consultar a diversos interesados, realizar consultas públicas e implementar enfoques de participación ciudadana a lo largo del ciclo de vida del sistema de IA, especialmente a quienes puedan verse afectados directa o indirectamente por su despliegue, como mujeres, personas con discapacidad, menores y niños. Incorporar sus comentarios en el desarrollo y uso de herramientas de IA que se emplean para emitir juicios o que tienen potencial para influir en cualquier otro asunto legal significativo.⁴⁸

2. Directrices específicas para organizaciones judiciales

Estas Directrices definen las “recomendaciones” como vías de acción sugeridas que derivan de los “principios” que ayudan tanto a las organizaciones a adquirir como a desarrollar sistemas de IA seguros, protegidos y fiables, como a los individuos a utilizar sistemas de IA de manera ética, responsable y respetuosa con los derechos humanos.

Las recomendaciones de esta sección se dividen en tres categorías (Figura nº 2) y están dirigidas a las autoridades judiciales, cortes y tribunales que pretendan analizar, desarrollar, adoptar, utilizar y evaluar sistemas de IA.

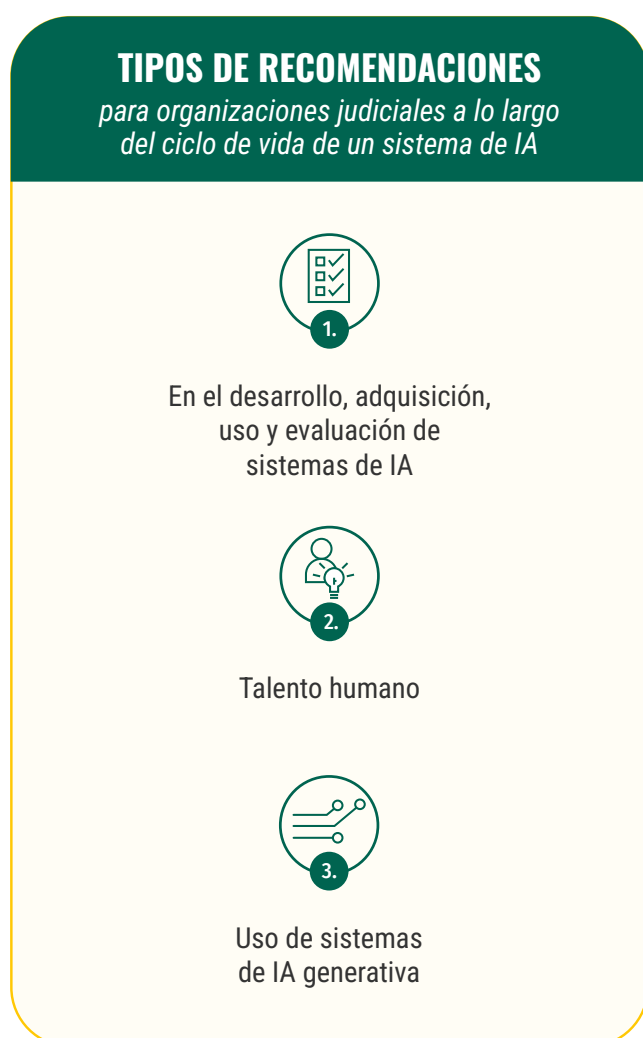


Figura nº 2. Tipos de recomendaciones para organizaciones judiciales a lo largo del ciclo de vida de un sistema de IA

2.1. Sobre el desarrollo, adquisición, uso y evaluación de sistemas de IA

Las siguientes recomendaciones ofrecen orientación sobre las acciones que deben llevar a cabo las organizaciones y autoridades judiciales en las diferentes etapas del ciclo de vida de la IA.

2.1.1. Proteger la independencia e imparcialidad del poder judicial.

Garantizar que el poder judicial tenga un papel de supervisión y responsabilidad en el desarrollo, despliegue y uso de herramientas de IA y que existan regulaciones para proteger su independencia.⁴⁹ Las herramientas de IA no deberían alterar ni anular esencialmente las características fundamentales de la imparcialidad de los jueces y los tribunales.⁵⁰

2.1.2. Considerar las especificidades culturales y contextuales

de la población afectada por el despliegue de sistemas de IA, así como las barreras para acceder a la justicia de individuos sin el conocimiento o las habilidades necesarios para interactuar con el sistema legal, por ejemplo, poblaciones mayores, subrepresentadas y desatendidas.

2.1.3. Establecer una gobernanza multiactor de las herramientas de IA.

Redactar una política o estrategia de IA que diagnostique las principales barreras para la adopción efectiva de sistemas de IA, establezca objetivos y métricas y formule planes de acción. Además, explorar la creación de un consejo asesor multiactor para la revisión continua y las actualizaciones de la política o estrategia de IA. Se debería invitar a integrantes del poder judicial a participar en la mejora continua de la política de IA y de cualquier otro instrumento adoptado por la organización, por ejemplo, proporcionando retroalimentación basada en sus experiencias con los sistemas de IA y las normas y orientaciones establecidas para su uso. Además, antes de adoptar sistemas de IA utilizados para la toma de decisiones o herramientas que apoyen procesos de gestión y que puedan afectar significativamente a derechos fundamentales, como el acceso a la justicia, la organización debería consultar con los interesados

relevantes sobre las oportunidades, los riesgos potenciales e impactos de la introducción de dichas herramientas. El poder judicial debe implicar de forma proactiva a las comunidades marginadas en los procesos de desarrollo, asegurando que se tengan en cuenta sus perspectivas.

2.1.4. Adoptar medidas algorítmicas de transparencia y explicabilidad desde el principio.

Empezar por decidir si se debe optar por sistemas de código abierto o propietarios. Antes de comprometerse a adoptar una herramienta de IA, la organización judicial debería evaluar las consecuencias de optar por cualquiera de las dos opciones respecto del acceso a la información del sistema, la explicabilidad y la auditabilidad.⁵¹ Luego, considerar medidas de transparencia en las primeras etapas del desarrollo y la adquisición de herramientas de IA. Por ejemplo, desarrollar, adquirir y⁵² obtener información de los desarrolladores y proveedores del sistema de IA sobre cómo se desarrolló el modelo (incluidos los datos utilizados para entrenarlo) y cómo funciona, los usos adecuados de la herramienta, así como los límites y riesgos asociados a su adopción.^{54 55}

2.1.5. Necesidad, proporcionalidad, idoneidad y alineación. La necesidad, proporcionalidad e idoneidad de utilizar un sistema de IA para realizar tareas deben evaluarse y establecerse desde el principio, alineándose con los objetivos de la organización que utiliza el sistema de IA (por ejemplo, que proteja los derechos humanos y garantice la independencia y la autonomía del poder judicial).

2.1.6. Justicia por diseño. Cuando la organización judicial encargue a un tercero el desarrollo o la adquisición de herramientas de IA, se deben tomar medidas para garantizar que los desarrolladores o proveedores eviten sesgos en los datos y en el diseño de los algoritmos. Esto último implica que los equipos de desarrollo y de pruebas sean multidisciplinarios y diversos.

2.1.7. Implementación por fases de herramientas de IA. Considerar la posibilidad de probar nuevas herramientas de IA antes de su

despliegue a gran escala y planificar una implementación por fases. La fase piloto permite evaluar la funcionalidad, el impacto potencial y los posibles riesgos de la IA en un entorno controlado. Una estrategia de implementación por fases permitiría ajustes basados en los hallazgos iniciales y en la retroalimentación de los interesados, promoviendo así un proceso de despliegue más informado, responsable y éticamente sólido.

2.1.8. Sostenibilidad e impacto ambiental de los sistemas de IA. El desarrollo, despliegue y uso de ciertos sistemas de IA pueden tener una huella ambiental sustancial⁵⁶ que el poder judicial debería tener en cuenta.

2.1.9. Establecer normas y directrices internas para el desarrollo, despliegue y uso de herramientas de IA basadas en el marco de derechos humanos.⁵⁷ Determinar qué sistemas de IA pueden utilizarse y cuáles usos están aceptados y prohibidos para procesos y tareas específicas. Además, las directrices deberían incluir protocolos para la notificación de incidentes e información clara sobre la responsabilidad de cada actor dentro de la organización. Las directrices deberían enfatizar que los miembros del poder judicial asuman la responsabilidad por cualquier material que produzcan utilizando IA. Para ello, las directrices podrían incluir un análisis de las posibles responsabilidades legales y los mecanismos correspondientes para prevenirlas. Además, las directrices deberían ofrecer escenarios detallados y ejemplos para orientar a los miembros del poder judicial sobre qué constituye un uso negligente y los pasos a seguir cuando no cumplen con los protocolos establecidos para el uso de IA.

2.1.10. Asegurar la intervención humana. Asegúrate de que el sistema siempre permita la intervención humana. En otras palabras, el control y supervisión humanos, o al menos la monitorización, deberían producirse durante todas las etapas de implementación y uso del sistema de IA.

2.1.11. Implementar marcos de rendición de cuentas antes de desplegar sistemas de IA.

Antes de desplegar un sistema de IA para apoyar los procesos de toma de decisiones, definir un marco de rendición de cuentas⁵⁸ para prevenir y mitigar los impactos y riesgos negativos y potenciar los positivos. El poder judicial puede implementar un “análisis de impacto algorítmico”⁵⁹ para identificar los efectos probables (positivos y negativos), un “análisis de riesgos algorítmicos”⁶⁰ para mapear los riesgos potenciales (como fallos del sistema) o un “análisis de impacto en derechos fundamentales”.⁶¹ Asegurarse de que el desarrollador o proveedor del sistema de IA acepte permitir y colaborar con auditorías algorítmicas realizadas por la organización o por partes independientes y externas.⁶² Si el sistema de IA se desarrolló internamente, facilitar el desarrollo de la capacidad interna para monitorear y auditar la herramienta (con orientación externa si fuera necesario), por ejemplo, estableciendo “mecanismos que faciliten la idoneidad del sistema, como garantizar la trazabilidad y registrar los procesos y resultados del sistema de IA.”⁶³ El poder judicial debería considerar el papel de la sociedad civil y de auditores independientes, especialmente en casos donde las herramientas de IA puedan generar altos riesgos de socavar derechos fundamentales. Estos informes pueden ayudar a tomar decisiones sobre si los sistemas deben desplegarse y recomendar medidas de prevención, mitigación, reparación y monitorización de riesgos. Los análisis algorítmicos deberían estar disponibles para consulta pública. El “Análisis de Impacto Ético: Herramienta de la Recomendación sobre la Ética de la Inteligencia Artificial” de la UNESCO puede utilizarse para examinar posibles impactos en los derechos humanos, como el juicio justo y el debido proceso, el acceso a la justicia y a la reparación eficaz, y la privacidad⁶⁴. Estas evaluaciones deben realizarse en consulta con los grupos de interés relevantes.

2.1.12. Realizar auditorías algorítmicas y evaluaciones de impacto tras el despliegue de sistemas de IA.

⁶⁵ Para ello, el poder judicial debería explorar la creación de directrices para el seguimiento, la auditoría y la evaluación de

los sistemas de IA en el sector judicial.⁶⁶ Las auditorías deben examinar si estos sistemas cumplen con la normativa o con cuestiones más específicas, como el sesgo, la equidad, la transparencia, la explicabilidad, la seguridad o el desempeño.⁶⁷ Las evaluaciones pueden identificar los efectos de los sistemas de IA en los usuarios, en las poblaciones afectadas y en la sociedad. Esto último debería incluir la evaluación del impacto en los derechos humanos y la identificación de los principales efectos del sistema (intencionados y no intencionados) sobre diversos grupos de usuarios y poblaciones. Las evaluaciones integrales de impacto en derechos humanos deberían ser públicas.⁶⁸ Publicar informes periódicos sobre los impactos y el rendimiento (por ejemplo, eficacia y eficiencia) de los sistemas de IA adoptados para cumplir los objetivos de la organización. Además, “establecer procesos para que terceros (por ejemplo, proveedores, consumidores, distribuidores/proveedores) o trabajadores puedan informar sobre posibles vulnerabilidades y riesgos.”⁶⁹ Adicionalmente, la organización que gobierna el poder judicial debe establecer un sistema de gestión de riesgos que permita identificar, detectar, monitorear, clasificar, diagnosticar, controlar y prevenir riesgos, así como mitigar daños. El sistema de gestión debe asignar roles, responsabilidades y procedimientos claros a lo largo del ciclo de vida de la IA.⁷⁰

2.1.13. Dejar de usar sistemas de IA que afecten negativamente a los derechos humanos.

Implementar sistemas desarrollados y ofrecidos en cumplimiento del derecho internacional de derechos humanos, la constitución nacional y las leyes locales, como la protección de datos y la privacidad, la no discriminación, la equidad procesal, las leyes de consumo, competencia y propiedad intelectual. Además, cuando un proveedor externo suministre el sistema de IA, asegurarse de que los términos de uso de la herramienta sean compatibles con las leyes locales y los estándares organizativos de las organizaciones que gobiernan el poder judicial. El poder judicial debería abstenerse de adoptar sistemas de IA que puedan afectar negativamente a los derechos humanos y/o civiles, por ejemplo,

perpetuando resultados discriminatorios y otras injusticias.⁷¹ En la misma línea, como recomienda el Informe del Secretario General de la ONU “Derechos humanos en la administración de justicia”, el poder judicial debería abstenerse “de utilizar o cejen en el uso de aplicaciones de IA cuando este resulte incompatible con el cumplimiento de la legislación internacional en materia de derechos humanos (...) sobre todo en decisiones judiciales y casos penales y en las evaluaciones del riesgo de reincidencia para la adopción de decisiones en materia de fianzas y libertad condicional, salvo y hasta que las autoridades responsables puedan demostrar que con esas aplicaciones se respeta el derecho a un juicio imparcial, en especial un poder judicial independiente, y los derechos a la libertad y la seguridad, a la no discriminación, a la protección frente a torturas y malos tratos y a la intimidad, así como otros derechos humanos afectados, y que las autoridades han afrontado los perjuicios desproporcionados que puede tener el uso de esas tecnologías en grupos concretos.”⁷² Cuando el poder judicial tiene pruebas preliminares de que el uso de una herramienta de IA puede producir efectos negativos significativos o violaciones de derechos humanos, el despliegue y uso del sistema debería suspenderse hasta que se realice una investigación exhaustiva. Cuando las evaluaciones algorítmicas o de impacto detectan que ciertos sistemas de IA violan los derechos humanos, el poder judicial dejará de usarlos y dismantelará dichas herramientas.

2.1.14. Accesibilidad de las herramientas de IA.

Asegurarse de que los miembros del poder judicial que puedan beneficiarse de las herramientas de IA tengan acceso adecuado a los sistemas. Esto último incluye asegurarse de que las personas puedan acceder al hardware necesario para utilizar las herramientas y contar con las habilidades necesarias para hacerlo eficazmente. Para ello, es pertinente diagnosticar, monitorizar y abordar periódicamente las barreras a la adopción de sistemas de IA, especialmente en lo que respecta a las que enfrentan las personas que forman parte del poder judicial.

2.1.15. Medidas de seguridad y ciberseguridad. Los sistemas de IA son propensos a ataques adversarios que pueden alterar sus resultados. Por ejemplo, los sistemas de IA generativa pueden ser vulnerables a “inyecciones de prompts”, extracción de datos, evasión y ataques de envenenamiento.⁷³ Por lo tanto, antes de desplegar una herramienta, el poder judicial debe poner a prueba su resiliencia, protegerla y mantenerla a lo largo de su ciclo de vida. Adoptar medidas técnicas, gerenciales y humanas para prevenir, controlar y mitigar los riesgos e incidentes de ciberseguridad. Cuando el despliegue y uso del sistema de IA impliquen el acceso a servicios en la nube, asegurarse de que el nivel de protección ofrecido por el proveedor cumpla con los estándares legales locales y los estándares de seguridad organizacional.

2.1.16. Mejorar la protección de los datos personales y la privacidad.

Establecer un marco e infraestructura robustos de gobernanza de los datos para proteger los datos personales, promover prácticas responsables de intercambio de datos, tanto personales como no personales, y garantizar que los datos sean inclusivos. La organización también debería establecer “protocolos de datos que regulen el acceso a los datos. Estos protocolos deben especificar quién puede acceder a los datos y bajo qué circunstancias. Solo se debe permitir que el personal debidamente cualificado con la competencia y necesidad de acceder a los datos individuales pueda hacerlo”.⁷⁴ Dada la naturaleza sensible de los datos personales y legales que se manejan en los procesos judiciales, es esencial contar con estrictas medidas de protección de la privacidad. (i) Minimización de datos. Para mitigar el riesgo de filtraciones de datos, deberían utilizarse sistemas de IA que requieran un mínimo de datos personales para funcionar, especialmente en casos que involucren información personal sensible. Por ello, las organizaciones judiciales deberían desplegar sistemas de IA basados en procesos que mejoren la privacidad. (ii) Protocolos de consentimiento. De acuerdo con las leyes locales de protección de datos y privacidad, redactar e implementar protocolos de consentimiento para el tratamiento

de datos personales con herramientas de IA en el ámbito judicial. Esto garantizará que las partes estén informadas y puedan controlar cómo se utilizan sus datos. (iii) Las técnicas de anonimización de datos deben emplearse cuando los sistemas de IA procesan datos personales, especialmente en la creación de bases de datos utilizadas para análisis legal o estudios de precedentes (por ejemplo, anonimizar los datos personales de las partes de decisiones judiciales antes de que se publiquen las sentencias).⁷⁵ Sin embargo, esto debe equilibrarse con la necesidad de contar con conjuntos de datos abiertos, así como con el derecho a la libertad de expresión y al acceso a la información, para no obstaculizar el desarrollo de sistemas de justicia digitales. El equilibrio de estos derechos en la práctica depende de cómo cada jurisdicción regule el derecho a acceder a documentos judiciales y las normas de anonimización derivadas de las leyes de protección de datos y privacidad.

2.1.17. Divulgar proactivamente información clave sobre los sistemas de IA utilizados por el poder judicial. Por ejemplo, las autoridades judiciales pueden publicar un repositorio en línea con información clave sobre los sistemas de IA adoptados para administrar justicia.⁷⁶ El repositorio puede incluir información sobre los datos de entrenamiento, el modelo (por ejemplo, variables, pesos), si es propietario o de código abierto, cómo funcionan los sistemas de IA, quién utiliza las herramientas, cómo se utilizan, las implicaciones para los procesos judiciales de toma de decisiones y si las decisiones adoptadas con el apoyo del sistema son cuestionables.⁷⁷ Esta información debe actualizarse periódicamente, al menos una vez al año, e indicar la fecha de la última actualización. Además, la creación de repositorios públicos de sistemas de IA utilizados por cortes y tribunales podría contribuir a la apertura de los datos. Para ello, evaluar hasta qué punto los instrumentos de transparencia pueden cumplir con los llamados principios FAIR respecto a los datos abiertos: localizables, accesibles, interoperables y reutilizables por cualquiera o para cualquier propósito. Las directrices de la UNESCO para establecer e implementar políticas de datos abiertos pueden conside-

rarse para este propósito.⁷⁸ Paralelamente a este esfuerzo, el poder judicial también podría llevar a cabo campañas de concienciación pública sobre el uso de los sistemas de IA en los procesos judiciales.

2.1.18. Garantizar la impugnabilidad de decisiones apoyadas o sostenidas por sistemas de IA. Implementar las medidas necesarias para garantizar el derecho a oponerse a decisiones tomadas con apoyo o fundamentadas en herramientas de IA. Esto requiere conocimientos significativos sobre cómo se produjeron los resultados, así como el derecho de las partes a auditar y examinar el modelo de IA subyacente para identificar deficiencias que puedan afectar su caso.

2.2. Talento humano

Las siguientes recomendaciones ofrecen orientación sobre las acciones que debe tomar el poder judicial para apoyar a las personas que desplieguen y utilicen sistemas de IA.

2.2.1. Revisar el plan de estudios de las Escuelas Judiciales. Considera incorporar temas de IA en el currículo permanente de las Escuelas Judiciales, como los fundamentos de la IA, la aplicación de tecnologías de IA en los sistemas judiciales y el impacto de la IA en la independencia de los jueces, la administración de justicia, especialmente en temas de derechos humanos, ética y gobernanza.⁷⁹

2.2.2. Proporcionar acceso a la formación. Ofrecer oportunidades de formación periódicas a los miembros del poder judicial que utilizarán las herramientas de IA y asegurarse de que desarrollen competencias clave de alfabetización en IA, especialmente para identificar la idoneidad de las herramientas de IA que puedan emplearse para realizar diferentes tipos de tareas, entender cómo utilizarlas y sus posibles impactos y riesgos de acuerdo con las normas internacionales, legislación regional y nacional sobre derechos humanos, y evaluar críticamente sus resultados.⁸⁰ Las organizaciones judiciales deberían considerar si es factible y pertinente establecer programas de formación obligatorios para el

personal judicial, a fin de mantener actualizados sus conocimientos y competencias.

2.2.3. Investigación sobre el uso y los efectos de las herramientas de IA. Las autoridades judiciales deberían estudiar periódicamente cómo sus miembros utilizan los sistemas de IA. La investigación podría documentar estudios de caso para mostrar buenas prácticas y riesgos potenciales a evitar. Esta última ayudaría a los miembros del poder judicial a informar sobre sus experiencias y desafíos, y a compartir las lecciones aprendidas con sus compañeros. Esto debe hacerse en pleno respeto a la independencia del poder judicial (véase arriba, 2.1.1).

2.2.4. Evaluar la necesidad de actualizar las políticas y prácticas de recursos humanos. Dependiendo de las herramientas de IA que se utilicen, el poder judicial puede necesitar revisar la planificación de su plantilla, ya sea mejorando las habilidades del personal existente o contratando nuevo personal. Esto puede requerir actualizaciones en las políticas y procesos de selección para la evaluación del desempeño que aseguren que se evalúen e incentiven los usos adecuados de las herramientas de IA.

2.3. Uso de sistemas de IA generativa

Las siguientes recomendaciones ofrecen orientación sobre qué acciones deben tomar las autoridades judiciales en relación con la adopción, el despliegue y el uso de sistemas de IA generativa.

2.3.1. Conciencia de las funciones y limitaciones. Asegurarse de que los miembros del poder judicial que utilizarán las herramientas de IA generativa sean conscientes de los usos adecuados, limitaciones y riesgos asociados (por ejemplo, resultados sesgados o incorrectos) con su adopción para la redacción de documentos legales y el apoyo a la participación en procedimientos legales, así como para apoyar las actividades de gestión judicial (por ejemplo, traducción).

2.3.2. Asegurar la autenticidad e integridad del contenido. Para lograrlo, deben tomarse varias medidas. (i) Todos los documentos

legales, presentaciones de pruebas u opiniones judiciales generados por IA deben estar claramente etiquetados como asistidos por IA para que jueces, abogados y partes implicadas sean conscientes de la naturaleza del contenido. (ii) Se deben implementar sistemas robustos para seguir el desarrollo y las modificaciones del contenido legal generado por IA. Esto es importante para fines probatorios, ya que permite verificar todos los materiales utilizados en el tribunal y confirmar su autenticidad. Por ejemplo, si se ha utilizado una herramienta de IA para crear o modificar un documento legal, debería haber un registro claro de cuándo y cómo se utilizó dicha herramienta. (iii) Se deben desarrollar protocolos de certificación para las herramientas de IA utilizadas en el poder judicial, verificando que cumplan con las directrices éticas internacionales y locales, así como con los estándares legales de precisión y fiabilidad de cada jurisdicción. Un ejemplo de esto podría ser un proceso de certificación para una herramienta de IA utilizada para analizar pruebas, asegurando que sea lo suficientemente precisa y fiable como para ser utilizada en un tribunal.

2.3.3. Restricciones de uso. Se deben emitir directrices específicas para regular el uso de la IA generativa en el poder judicial, con el fin de prevenir el uso indebido y proteger la integridad del proceso legal, en línea con las normas internacionales. Las aplicaciones específicas de la IA deberían prohibirse o limitarse, teniendo en cuenta su impacto en los derechos humanos, especialmente en la privacidad y la protección de los datos. Por ejemplo, cuando los términos de uso de una herramienta de IA generativa indican que el proveedor utilizará la información provista por el usuario para entrenar sus modelos o que terceros pueden acceder a dicha información, entonces el uso de dicha herramienta debería prohibirse o restringirse para evitar que el poder judicial pierda el control sobre quién puede acceder a información confidencial o datos personales. Además, por ejemplo, debería prohibirse el uso de la IA en ciertas áreas sensibles, como la generación unilateral de decisiones vinculantes o la creación de material probatorio fabricado

sintéticamente. Se debe prestar especial atención al cumplimiento de los derechos de propiedad intelectual, asegurando que cualquier contenido generado por IA respete los derechos de propiedad intelectual del texto original.

3. Orientación específica para los miembros individuales del poder judicial

Las recomendaciones de esta sección son para personas que forman parte del poder judicial, incluidos jueces, magistrados, fiscales, funcionarios judiciales y personal de apoyo judicial que tengan la intención de desplegar y utilizar sistemas de IA. Las recomendaciones se dividen en tres categorías (Figura n° 3).



Figura n° 3. Tipos de recomendaciones para el despliegue y uso de sistemas de IA por parte de los miembros individuales del poder judicial

3.1. Antes de usar sistemas de IA

Las siguientes recomendaciones ofrecen orientación sobre qué acciones deben tomar las personas que forman parte del poder judicial antes de desplegar y utilizar sistemas de IA.

3.1.1. Concienciación en IA, alfabetización en IA y desarrollo de capacidades. Las personas deben ser conscientes de las funcionalidades, fortalezas y precisión de los sistemas de IA, así como de sus limitaciones, sesgos y riesgos en el contexto de actividades legales.⁸¹ Este último también incluye la conciencia de las responsabilidades derivadas del uso negligente de estas herramientas. Las organizaciones judiciales deben garantizar que las personas tengan acceso a programas educativos y formación continua para evaluar críticamente los resultados y el rendimiento de las herramientas.⁸²

3.1.2. Utilizar herramientas de IA probadas mediante análisis algorítmicos y seguir las prescripciones de los informes de debida diligencia en derechos humanos. Como se explicó anteriormente, las organizaciones judiciales deberían implementar marcos de rendición de cuentas antes y después del despliegue de sistemas de IA para identificar implicaciones y riesgos para el acceso a la justicia y los derechos humanos. Sin embargo, si su organización no analiza previamente sus sistemas de IA, debe asegurarse de utilizar herramientas probadas (por parte del proveedor o de terceros expertos) y seleccionar aquellas que hayan sido objeto de análisis sobre el impacto en los derechos humanos.⁸³ Además, se deben seguir las prescripciones de los informes de debida diligencia en materia de derechos humanos sobre sistemas de IA, que deben elaborarse a lo largo de todo su ciclo de vida por las autoridades judiciales.

3.2. Uso de sistemas de IA

Las siguientes recomendaciones ofrecen orientación sobre qué acciones deben tomar las personas que forman parte del poder judicial durante el despliegue y el uso de sistemas de IA.

3.2.1. Evitar depender demasiado de los sistemas de IA para tomar decisiones sustantivas.

Los jueces y magistrados no deben confiar únicamente en los sistemas de IA para tomar decisiones sobre el fondo de un caso ni para decidir sobre cuestiones procesales que puedan afectar a los derechos humanos; en su lugar, deben utilizar los resultados de las herramientas de IA para complementar el análisis jurídico realizado con otros métodos y fuentes de información. La decisión judicial no es un mero resultado, sino una decisión tomada en un proceso.

3.2.2. Cumplir con los términos de uso. Leer y cumplir con los términos de uso de los sistemas de IA publicados por el proveedor externo de la herramienta o proporcionados por tu organización cuando el sistema se desarrolló internamente. Estos términos suelen indicar los usos adecuados y prohibidos de las herramientas, así como los riesgos que deben evitar los responsables de su despliegue y los usuarios. Sin embargo, debe consultar con su organización sobre los términos de uso de los sistemas de IA, especialmente si considera que algunos términos pueden restringir innecesariamente los derechos de los usuarios, negar transparencia o eximir al desarrollador de cualquier responsabilidad.

3.2.3. Transparencia proactiva. Proporcionar información relevante sobre cuándo se utiliza la herramienta de IA y cómo su uso podría afectar a las personas que forman parte de un procedimiento judicial o a los beneficiarios del trabajo legal. Informar qué herramientas se usaron y sus respectivas versiones. Asegurarse de que los materiales producidos o publicados con IA sean reconocidos y distinguidos mediante comillas o citas.

3.2.4. Responsabilidad y rendición de cuentas. Seguir las normas establecidas por su organización. Esta recomendación implica tanto un deber *ex ante* de revelar el uso de un sistema de IA como un deber *ex post* de proporcionar más información cuando sea necesario para determinar la responsabilidad.

3.2.5. Oportunidad para revisar las decisiones y para la impugnación.

Informar a las partes sobre el uso de sistemas de IA y proporcionar oportunidades adecuadas para cuestionar e impugnar decisiones tomadas con o apoyadas por ellos, así como información básica sobre cómo funciona el sistema de IA, cómo fue entrenado, qué entradas se utilizaron para operar el sistema y cómo los resultados producidos por la herramienta de IA informaron la decisión. Garantizar que las personas tengan derecho a un recurso efectivo en circunstancias en las que los sistemas de IA tengan un impacto negativo en sus derechos humanos, incluyendo tanto durante como después de la finalización de los procedimientos penales contra ellas. El poder judicial debe garantizar que cualquier persona que desee impugnar estas decisiones tenga acceso a servicios de asistencia jurídica, recursos informativos o asistencia para navegar por el proceso de apelación.

3.2.6. Informes proactivos para prevenir daños.

Informar a la organización cuando haya sospechas de fallo o posibles o probables impactos negativos, y dejar de usar el sistema de IA si detecta que puede causar daño a los derechos humanos.

3.3. Uso de sistemas de IA generativa

Las siguientes recomendaciones ofrecen orientación sobre qué acciones deben tomar los miembros del poder judicial en relación con la adopción, el despliegue y el uso de sistemas de IA generativa.

3.3.1. Proteger los datos personales y confidenciales.

No incluir datos personales ni información confidencial en las instrucciones (*prompts*) cuando se utilicen herramientas externas de IA generativa.⁸⁴ Tener en cuenta que cualquier información que introduzca en un chatbot de IA comercial (por ejemplo, como parte del *prompt*) debe considerarse disponible para todos, entre otros, porque los términos de servicio de la mayoría de las empresas de IA que ofrecen acceso gratuito a estos sistemas indican que las entradas se usarán para entrenar modelos futuros.⁸⁵

3.3.2. Usos principales de los LLM. Los LLMs pueden utilizarse para diferentes tareas, incluyendo, pero no limitándose a, redactar documentos legales básicos, redactar discursos y presentaciones, resumir, traducir, hacer correcciones gramaticales, modificar el tono de un texto (por ejemplo, de informal a formal), mejorar su legibilidad, explorar temas específicos y llevar a cabo tareas administrativas (por ejemplo, redactar correos electrónicos).⁸⁶ Sin embargo, todas las tareas anteriores implican verificar la producción del sistema de IA y contrastarla con fuentes fiables. Las herramientas de IA pueden ayudar a localizar material que la persona conoce y puede evaluar por su cuenta. Por el contrario, estas herramientas resultan ineficaces para investigar información que no puede ser evaluada o verificada de forma independiente.⁸⁷ Se recomienda consultar los usos sugeridos para herramientas de IA generativa al final de esta sección.

3.3.3. Falta de fiabilidad como motores de búsqueda y para el análisis legal. Los LLMs comerciales de propósito general no son fuentes fiables de información ni medios adecuados para realizar análisis legales ni tareas matemáticas. “Incluso con los mejores *prompts*, la salida puede ser inexacta, incompleta, engañosa o sesgada.”⁸⁸

3.3.4. No sustituir los testimonios de peritos por LLMs. El testimonio de los peritos abarca los procesos y las metodologías mediante los cuales se elaboran los informes periciales y las descripciones, razones, argumentos, análisis, evaluaciones y conclusiones expuestos en el texto. No se deben utilizar LLMs para producir los tipos de productos que los peritos requerirían, no solo por la falta de fiabilidad de los informes generados por la herramienta, sino también por la opacidad con la que el sistema infiere a partir de sus datos de entrenamiento.⁸⁹

3.3.5. Conciencia de las limitaciones y riesgos de los LLM. Tener en cuenta que los resultados generados por los LLMs actuales pueden incluir información incorrecta, imprecisa o ficticia sobre cuestiones fácticas, legales (leyes y jurisprudencia) y técnicas. Los LLM pueden

producir respuestas que carecen de referencia al mundo real y que son textos sin sentido.⁹⁰ Además, debe ser consciente del posible sesgo en el texto generado por los LLM: “Los desequilibrios en los datos de entrenamiento también pueden conducir a sesgos en el modelo.”⁹¹ Además, ciertos LLM pueden haber sido entrenados con datos hasta cierto punto en el tiempo (por ejemplo, con datos disponibles hasta el último momento en que fueron entrenados). Por lo tanto, la información sobre casos legales recientes, leyes y hechos puede no haberse incluido en los datos de entrenamiento; por ello, el resultado producido por el *chatbot* que opera a partir de un LLM puede estar desactualizado o inexacto.

3.3.6. Verificar las salidas antes de usarlas. La estructura convincente del texto producido por un LLM no debe conducir a una confianza excesiva en la factualidad y la veracidad del resultado.⁹² “Por diversas razones, los LLM no ofrecen garantías sobre la veracidad, la calidad ni el formato deseado (por ejemplo, un formato de código específico) de sus salidas.”⁹³ Por lo tanto, evita una confianza excesiva en las herramientas de IA, asegurando siempre que los resultados de los sistemas de IA generativa estén verificados y contrastados con fuentes de información fiables antes de su uso en materiales y documentos legales. El pensamiento crítico humano debe ser central en todas las interacciones con sistemas de IA generativa.

3.3.7. Uso transparente. Revelar el uso de sistemas de IA generativa para redactar textos —sentencias, opiniones y otros documentos que puedan tener consecuencias legales— o cuando se emplee explícitamente en audiencias judiciales. Para ello, debe distinguir el texto producido por el *chatbot* de IA utilizado en una decisión mediante comillas y un sistema de citas.⁹⁴ Cuando el texto producido por el *chatbot* se incluye en una decisión judicial, toda la interacción que conduce al resultado final —tanto las indicaciones del usuario como las respuestas del *chatbot*— debe documentarse. Dependiendo de la longitud y complejidad de la interacción, esta documentación puede incluirse en el cuerpo principal del texto, en una nota al pie o en un anexo.

3.3.8. Integridad. No hacer pasar el texto sintético como propio. Además, evitar posibles infracciones de derechos de autor y de propiedad intelectual asociadas al uso de contenido generado por sistemas de IA generativa.

3.3.9. Responsabilidad. Asumir la responsabilidad de utilizar los resultados generados por los sistemas de IA generativa para redactar sentencias y decisiones judiciales, e informar sobre su uso para participar en audiencias judiciales. Asegurarse de la precisión y fiabilidad de la salida generada por el *chatbot* antes de usarla.⁹⁵

FINALMENTE, AQUÍ SE PRESENTAN DIEZ EJEMPLOS DE USOS DE LOS SISTEMAS DE IA GENERATIVA:

1. Corregir o mejorar los borradores iniciales de informes administrativos, actas de reuniones, correos electrónicos, apuntes, decisiones judiciales y resoluciones. Ejemplos de uso de casos:
 - Mejora la gramática, la ortografía y la estructura de los textos.
 - Editar el tono del texto, por ejemplo, hacerlo más accesible para diferentes públicos objetivo o adoptar un lenguaje sencillo.
 - Preparando borradores de traducción de pasajes.
2. Reorganizar las referencias según un sistema de citas de fuentes dado.
3. Comparar textos y fuentes.
4. Presentaciones de diseño y creación de visualizaciones.
5. Generación y ajuste de código informático.
6. Realizar exploraciones iniciales de temas y seguir las siguientes etapas de búsqueda a partir de fuentes fiables.
7. Lluvia de ideas.
8. Estructuración de argumentos.
9. Explorar explicaciones alternativas o críticas sobre cuestiones de derecho.
10. Generar resúmenes: verificar siempre la exactitud de la salida.



*Las herramientas de IA no sustituyen
el razonamiento jurídico cualificado,
el juicio humano o el asesoramiento
jurídico personalizado.*

Conclusión

Como se subraya al principio de este documento, las herramientas de IA pueden apoyar los objetivos fundamentales del sector judicial. Aun así, la adopción de instrumentos defectuosos y el uso negligente de sistemas de IA por parte del poder judicial también pueden socavar los derechos humanos, como el juicio justo y el debido proceso, el acceso a la justicia y a la eficacia de los recursos, la privacidad y la protección de datos, la igualdad ante la ley y la no discriminación, así como valores judiciales como la imparcialidad, la independencia y la rendición de cuentas.

Las *Directrices* ofrecen principios y recomendaciones para maximizar los beneficios de las herramientas de IA en el sistema judicial y prevenir posibles daños, en especial violaciones de derechos humanos. Aunque el poder judicial puede implementar directamente los principios y recomendaciones para organizaciones e individuos, también pueden utilizarse como **modelo para desarrollar directrices nacionales y subnacionales específicas para cada contexto**.

Sin embargo, la implementación efectiva de herramientas de IA en el sector judicial requiere mucho más que la mera adopción de normas y directrices. También es esencial garantizar que el talento humano esté preparado para utilizar las herramientas de IA de forma ética y responsable, y que las herramientas sean accesibles, fiables y cumplan con los derechos humanos.

Las herramientas de IA no sustituyen el razonamiento jurídico cualificado, el juicio humano o el asesoramiento jurídico personalizado. Por ejemplo, los *chatbots* de IA que operan a partir de LLMs generan texto uniendo secuencias de formas lingüísticas detectadas en sus datos de entrenamiento “según información probabilística sobre cómo se combinan, pero sin ninguna referencia al significado”.⁹⁶ Estas herramientas de IA no pueden sustituir el razonamiento jurídico cualificado porque no tienen justificación ni comprensión contextual de un problema legal. El uso de LLMs por parte de personas no abogadas con fines legales requiere precaución, ya que la salida de los sistemas puede parecer autoritativa y coherente, al tiempo que presenta hechos que pueden ser inexactos y respuestas que no son coherentes con hechos legales.⁹⁷

En resumen, el poder judicial debería aspirar a utilizar herramientas de IA para mejorar, en lugar de reemplazar, el juicio humano.

Anexo

Revisores externos (en orden alfabético):

Karine Caunes (Directora Global del Programa, Derecho de la UE, Centro de IA y Política Digital, Estados Unidos), Giulia Gentile (Profesora Adjunta, Essex Law School, Reino Unido de Gran Bretaña e Irlanda del Norte), Felipe Guzmán (Consultor en instrumentos regulatorios para la transformación digital del Poder Judicial, Consejo Superior del Poder Judicial de Colombia), Ameen Jauhar (Jefe de equipo, Centro de Derecho Aplicado y Tecnología), Vidhi Centre for Legal Policy), Amanda Leal (Especialista en Gobernanza y Políticas de IA en HealthAI), Martin Okumu Masiga (Secretario General, Africa Jurists and Judges Forum, Johannesburg, Sudáfrica), Grace Mutung'u (Investigadora y Abogada del Tribunal Superior de Kenia), Juez Jean Aloise Ndiaye (Tribunal Supremo de Senegal), Alfonso Peralta Gutiérrez (Juez de Primera Instancia e Investigación Penal, Granada, España), Matthieu Quinou (Doctor en Derecho Privado y Abogado en el Colegio de Abogados de París, Francia), Anca Radu (Investigadora, Instituto Universitario Europeo, Departamento de Derecho, Florencia, Italia), Dr. Krishna Ravi Srinivas (Profesor Adjunto en la Universidad de Derecho NALSAR, India), Dra. Catherine Regis (Profesora, Facultad de Derecho, Universidad de Montreal, Canadá), Dra. Dory Reiling (Jueza Senior jubilada, Experta en Tribunales Internacionales y TIC, Ámsterdam, Reino de los Países Bajos), Agneris Sampieri (abogada mexicana), Dra. Miriam Stankovich (Especialista Principal en Política Digital, Centro de Aceleración Digital, Estados Unidos), Kalliopi Terzidou (Investigadora Doctoral, Facultad de Derecho, Universidad de Luxemburgo), Colin van Noordt (Investigadora, Reino de los Países Bajos), Carolina Villadiego (Asesora Legal para América Latina, Comisión Internacional de Juristas, Ginebra, Suiza), Anthony Wong (Presidente de IFIP, Federación Internacional de Procesamiento de Información, Austria).

Más allá de las consultas a expertos, se organizaron consultas públicas sobre el borrador de directrices entre agosto y septiembre de 2024. Las directrices finales se han enriquecido con más de cien contribuciones de individuos y organizaciones de 41 países, junto con comentarios colectivos de eventos organizados por la UNESCO con agentes judiciales en diferentes partes del mundo.

ACNUDH, Organización Mundial de la Salud, Access Now (Estados Unidos), Access to Justice Centre for Excellence, ACT (Canadá): The App Association (Estados Unidos), Advocacy and Awareness Centre Africa (AAC Africa) Initiative (Kenia), African IGF & African Digital Rights Network (Etiopía, colaborador desde Egipto), Anticorruption Research Centre (Kenia), ARTICLE 19 (Brasil, y Office for Mexico and Central America), Auticon (Francia), Cambridge Existential Risk

Initiative (Reino Unido de Gran Bretaña e Irlanda del Norte), Centro de IA y Política Digital (Estados Unidos), Centro para la Equidad en Salud, Derecho y Política (C-HELP, India), Centro para Internet y Sociedad (CEI, India), Centro para el Espacio, el Ciberespacio y el Derecho de los Datos (Bond University, Australia), Centro para el Estudio del Derecho y la Gobernanza (Universidad Jawaharlal Nehru, India), Centro de Capacitación y Perfeccionamiento 'Ricardo C. Núñez' del Poder Judicial de la Provincia de Córdoba (Argentina), Centro Federal de Conciliación y Registro Laboral (México), Centro para la Investigación del Ciberdelito y la Ciberseguridad (Centro para el CIC, Colombia), Centro de Derecho de la Propiedad Intelectual y Tecnologías de la Información (CIPIT, Kenia), Consejo de la Judicatura Federal (México), Consejo para la Investigación en Relaciones Económicas Internacionales (ICRIER, India), Tribunal de Apelación de El Cairo (Egipto), Cuevas Mella (República Dominicana), Sociedad Daksh (India), Rights Nepal, El Veinte (Colombia), Laboratorio Empatía (Perú), Escuela Judicial Electoral (Escuela Judicial Electoral, México), Instituto Europeo de Administración Pública (Luxemburgo) en colaboración con el Laboratorio de Derecho y Tecnología de Maastricht de la Universidad de Maastricht (Reino de los Países Bajos), Faculdade de Tecnología de Teresina (Brasil), Facultad de Derecho, Universidad Tecnológica de Sídney (Australia), Universidad Federal Báltica Immanuel Kant (Federación Rusa), Instituto Helénico de Derecho Internacional y Extranjero (Grecia), Derecho de Ikigai (India), Instituto Indio de Comercio Exterior de Calcuta (India), Inno Canyon Consulting (Nigeria), Asociación Internacional de Fiscales (Austria), Kictanet (Kenia), Laboratorio de Innovación Tecnológica de la Facultad de Derecho de la Universidad Nacional de Lomas de Zamora UNLZ-LAB (Argentina), Abogados Sin Fronteras (Canadá, El Salvador), Mano Erudita

(Estados Unidos), Comisión Nacional de Lituania para la UNESCO (Lituania), Tribunal de Magistrados de Maseru (Lesoto), Ministerio Público (Panamá), Ministerio de Asuntos Exteriores de España, Ministerio de Justicia, Asuntos Parlamentarios (Nepal), Academia Nacional de Estudios e Investigaciones Jurídicas (NALSAR, India), Academia Nacional Judicial (Nepal), Asociación de Abogados de Nepal, Sociedad de Abogados de Nueva Zelanda, Osh Technology & Artificial Intelligence Industries (Egipto), Unión Panafricana de Abogados (Tanzania), Poder Judicial Asunción (Paraguay), Priva Sapien (India), Rama Judicial (Colombia), Rencontre Africaine pour la Défense des Droits de l'Homme (RADDHO, Senegal), S.B Law firm (Islamabad, Pakistán), Saffi technologies/red de acción de internet de Kenia (Kenia), Facultad de Derecho, Trinity College Dublin (Irlanda), Tribunal Regional de Sofía (Bulgaria), Tribunal Supremo de Paraguay, Tribunal Supremo de Justicia del Estado de Aguascalientes (México), Universidad Taiz (Yemen), TeachSomebody (Reino de los Países Bajos), Equipo de Derecho Civil de la Asociación de Jueces Polacos IUSTITIA (Polonia), Tribunal Supremo de Nepal, SGH Escuela de Economía de Varsovia (Polonia), UBA – IALAB (Argentina), Universidad Americana (Paraguay), Universidad Galileo (Guatemala), Universidad de Bolonia (Italia), Universidad de Minas Gerais (UFMG, Brasil), Universidad de New Brunswick, Facultad de Derecho (Canadá), Universidad de Estrasburgo (Francia), Western Sydney University, IFIP, ACS (Australia).

Contribuciones individuales de jueces (Argentina, Costa Rica, Egipto, Guatemala, Paraguay, Tanzania), personal judicial (Nepal), abogados y profesionales del derecho (Camerún, Colombia, Guatemala, Kenia, México), organizaciones de la sociedad civil (Nepal), académicos (Bélgica, Costa Rica, Filipinas, Suiza) y periodistas (El Salvador).

Notas finales

- 1 GPAI, *Transparencia Algorítmica: Informe de Estado de la Arte* (Global Partnership on Artificial Intelligence (GPAI), noviembre de 2024) [🔗](#); Nestor Maslej y otros, *Informe Anual del Índice de IA 2025* (Comité Directivo del Índice de IA, Instituto para la IA Centrada en el Ser Humano, Universidad de Stanford, abril de 2025), p. 456 [🔗](#); Sarah Muñoz-Cadena y Juan David Gutiérrez, 'Repositorios de Algoritmos Públicos En El Mundo' (Sistemas de Algoritmos Públicos, Universidad de los Andes, marzo de 2025) [🔗](#).
- 2 Sarah Muñoz-Cadena y otros, 'Sistemas de IA En El Sector Público de América Latina y El Caribe' (Universidad de los Andes, marzo de 2025) [🔗](#).
- 3 Yannis Charalabidis, Rony Medagiglia y Colin van Noordt, *Manual de investigación sobre gestión pública e inteligencia artificial* (Edward Elgar Publishing, 2024), doi:10.4337/9781802207347; M. Manzoni y otros, *AI Watch Camino hacia la adopción de la inteligencia artificial por el sector público: Manual para responsables políticos, administraciones públicas y partes interesadas relevantes* (Oficina de Publicaciones de la Unión Europea, 2022), doi:10.2760/288757 (en línea), 10.2760/693531 (impreso); Regine Paul, Emma Carmel y Jennifer Cobbe, *Manual sobre Políticas Públicas e Inteligencia Artificial* (Edward Elgar Publishing, 2024), doi:10.4337/9781803922171.
- 4 Para una visión general de las diferentes tareas realizadas con sistemas de IA en el sector de la justicia, véase F. Bell y otros, *IA Decision-Making and the Courts: A Guide for Judges, Tribunal Members and Court Administrators* (Australasian Institute of Judicial Administration, 2022); UNESCO, *Kit Global de Herramientas sobre IA y Estado de Derecho para el Poder Judicial* (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), 2023) [🔗](#).

Para informes y artículos sobre herramientas específicas de IA utilizadas por operadores judiciales, véase Nikolaos Aletras y otros, 'Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective', ed. por Lexing Xie, *PeerJ Computer Science*, 2 (2016), pp. 1–19, doi:10.7717/peerj-cs.93; Urvashi Aneja y Dona Mathew, *Inteligencia artificial en el sistema judicial de la India: ¿un caso de irresponsabilidad organizada?* (Digital Futures Lab, marzo de 2023) [🔗](#); Elsa Estevez, Sebastián Linares y Pablo Fillottrani, *PROMETEA: Transformando La Administración de Justicia Con Herramientas de Inteligencia Artificial* (Banco Interamericano de Desarrollo, 2020) [🔗](#); Collen Zvandasara Kufakwababa, 'Herramientas de inteligencia artificial en automatización del trabajo legal: el uso y percepción de herramientas para el descubrimiento de documentos y los procesos de clasificación de privilegios en despachos jurídicos del sur de

África' (Universidad de Stellenbosch, 2021) [🔗](#); Goretty Carolina Martinez, 'La Inteligencia Artificial y Su Aplicación al Campo Del Derecho', *Alegatos*, n° 82 (2012), pp. 827–46; Masha Medvedeva, Michel Vols y Martijn Wieling, 'Usando el aprendizaje automático para predecir decisiones del Tribunal Europeo de Derechos Humanos', *Artificial Intelligence and Law*, 28.2 (2020), pp. 237–66, doi:10.1007/s10506-019-09255-y; Ministro do Superior Tribunal de Justiça, *Inteligencia Artificial: Tecnología Aplicada a la Resolución de Conflictos en el Poder Judicial Brasileño* (Fundação Getulio Vargas, 2020) [🔗](#); Marion Oswald y otros, 'Modelos de Vigilancia de Evaluación de Riesgos Algorítmica: Lecciones del Modelo HART de Durham y proporcionalidad "experimental"', *Information & Communications Technology Law*, 27.2 (2018), pp. 223–50, doi:10.1080/13600834.2018.1458455; Víctor Saavedra y Juan Carlos Upegui, *PretorIa y La Automatización Del Procesamiento de Causas de Derechos Humanos* (Derechos Digitales América Latina y Dejusticia, marzo de 2021) [🔗](#).

Para un repositorio de herramientas y proyectos de IA desarrollados por CNJ de Brasil, véase CNJ, *Repositório Nacional de Projetos de Software e Versionamento de Arquivos do – Git.jus* (Conselho Nacional de Justiça (CNJ)) [🔗](#) [consultado el 27 de mayo de 2024].

- 5 Por ejemplo, la Comisión Europea para la Eficiencia de la Justicia del Consejo de Europa (CEPEJ) ha documentado más de cien herramientas de IA que contribuyen a la gestión de procesos judiciales y a las tareas de toma de decisiones de jueces y magistrados. Véase CEPEJ, *Centro de Recursos sobre Ciberjusticia e IA* (Comisión Europea para la Eficiencia de la Justicia (CEPEJ), 6 de mayo de 2024) [🔗](#).
- 6 Kahneman y sus colaboradores se refieren al "ruido" como un error en el juicio humano –distinto del sesgo– que consiste en una variabilidad no deseada en la toma de decisiones que debería haber sido idéntica, causando inconsistencias aleatorias en las decisiones. En este sentido, "los juicios son ruidosos, ya que las personas que se espera que estén de acuerdo acaban en puntos muy diferentes". Además, su libro analizó cómo los algoritmos de aprendizaje automático podrían reducir el ruido en los escenarios de toma de decisiones judiciales. Daniel Kahneman, Olivier Sibony y Cass R. Sunstein, *Ruido. Un*

fallo en el juicio humano, 1ª edn (Little, Brown and Company, 2021), p. 4.

- 7 Nathan-Ross Adams, 'Parker contra Forsyth NO: Lecciones para el uso de la IA en la investigación jurídica', *Michalsons*, 11 de julio de 2023 [🔗](#); Robert J. Ambrogi, 'Cuatro meses después de lanzar su herramienta GenAI "nacional", el bufete de abogados Gunderson Dettmer informa sobre los resultados hasta ahora, nuevas características y una sorpresa en el coste', *Lawsites*, 20 de diciembre de 2023 [🔗](#); Francesca Bennetts y otros, 'IA responsable: navegando los riesgos y abrazando las posibilidades', *Allen Overy*, 12 de noviembre de 2023 [🔗](#);

Juan David Gutiérrez, 'Jueces y magistrados en Perú y México tienen fiebre de ChatGPT', *Tech Policy Press*, 19 de abril de 2023 [🔗](#); Juan David Gutiérrez, 'ChatGPT en los tribunales colombianos: por qué necesitamos tener una conversación sobre la alfabetización digital del poder judicial', *VerfBlog*, 23 de febrero de 2023 [🔗](#); Juan David Gutiérrez, 'Tecnologías de IA en el Poder Judicial: Evaluación Crítica de Grandes Modelos de Lenguaje en la Toma de Decisiones Judiciales', en *Handbook on Public Policy and AI*, ed. por Regine Paul, Emma Carmel y Jennifer Cobbe (Edward Elgar Publishing, 2014) [🔗](#); LexisNexis, 'LexisNexis lanza Lexis+ AI, una solución de IA generativa con citas legales vinculadas sin alucinaciones', *LexisNexis*, 25 de octubre de 2023 [🔗](#);

Benjamin Weiser, 'Esto es lo que ocurre cuando tu abogado usa ChatGPT', *The New York Times*, 27 de mayo de 2023, sección New York [🔗](#); Benjamin Weiser y Jonah E. Bromwich, 'Michael Cohen utilizó casos falsos citados por A.I. para solicitar el fin de la supervisión judicial', *The New York Times*, 29 de diciembre de 2023 [🔗](#);

Mark Witten, 'Aplicación de la IA generativa al derecho: oportunidades y riesgos | Queen's Law', *Queen's University*, 23 de octubre de 2023 [🔗](#);

Kieran Woboditsch-Velasco, 'La revolución silenciosa: el auge de la IA generativa en los despachos de abogados canadienses', *MyOpenCourt*, 25 de noviembre de 2023 [🔗](#);

Nuria Morcillo, 'Sentencias falsas, leyes extranjeras y filtrado de datos: los riesgos de usar ChatGPT se cuelan en los despachos de abogados', *El País*, 18 de febrero de 2025, sección Economía [🔗](#);

Sara Merken, 'Abogados en una demanda de Walmart admiten citaciones de caso "alucinadas" por IA', *Reuters*, 10 de febrero de 2025, sección Legal Industry [🔗](#);

Sam Harden, 'Uno pensaría que los abogados ya lo sabrían', *Team Do Something*, 14 de febrero de 2025 [🔗](#);

Kalliopi Terzidou, 'Sistemas de IA generativa en la práctica jurídica que ofrecen servicios legales de calidad mientras defienden la ética jurídica', *International Journal of Law in Context*, 2025, pp. 1–22, doi:10.1017/S1744552325000047.

- 8 Juan David Gutiérrez, 'Capítulo 24: Evaluación crítica de grandes modelos lingüísticos en la toma de decisiones judiciales', en *Handbook on Public Policy and Artificial Intelligence* (Edward Elgar Publishing, 2024), pp. 323–38, doi:10.4337/9781803922171.00033.

- 9 Para las directrices y políticas emitidas por los órganos que rigen el poder judicial y por los tribunales, véase:

UK Courts and Tribunals Judiciary, *Artificial Intelligence (IA) - Guidance for Judicial Office Holders* (Reino Unido de Gran Bretaña e Irlanda del Norte, 12 de diciembre de 2023) [🔗](#);

Conselho Nacional de Justiça, *Dispõe Sobre a Ética, a Transparência e a Governança Na Produção e No Uso de Inteligência Artificial No Poder Judiciário e Dá Outras Providências* (Conselho Nacional de Justiça (CNJ), 21 de agosto de 2020) [🔗](#);

Tribunales de Nueva Zelanda, *Directrices para el uso de la inteligencia artificial generativa en cortes y tribunales* (Nueva Zelanda, 7 de diciembre de 2023) [🔗](#);

Tribunal Federal, *Principios y Directrices Interinas sobre el Uso de la Inteligencia Artificial por parte del Tribunal* (Tribunal Federal, 20 de diciembre de 2023) [🔗](#);

Tribunal Federal, *Aviso a las Partes y a la Profesión: El Uso de la Inteligencia Artificial en Procedimientos Judiciales* (Tribunal Federal, 7 de mayo de 2024) [🔗](#);

Queensland Courts, *Directrices para el uso responsable de la inteligencia artificial generativa (IA) para el uso responsable por personas no abogadas*, 14 de mayo de 2024 [🔗](#);

Tribunal Supremo de Victoria, *Directrices para litigantes: Uso responsable de la inteligencia*

artificial en litigios (Tribunal Supremo de Victoria, mayo de 2024) [🔗](#);

Consejo Judicial Canadiense, *Directrices para el Uso de la Inteligencia Artificial en los Tribunales Canadienses* (Consejo Judicial Canadiense, septiembre de 2024) [🔗](#);

Tribunal Supremo del Estado de Delaware, *Política Interina sobre el Uso de GenAI por Funcionarios Judiciales y Personal de los Tribunales* (Tribunal Supremo del Estado de Delaware, 21 de octubre de 2024) [🔗](#);

Tribunal Supremo de la República de Singapur, *Guía sobre el uso de herramientas de inteligencia artificial generativa por parte de los usuarios de tribunales* (Tribunal Supremo de la República de Singapur, 23 de septiembre de 2024) [🔗](#);

Consejo Superior de la Judicatura, *Lineamientos Para El Uso y Aprovechamiento Respetuoso, Responsable, Seguro y Ético de La Inteligencia Artificial En La Rama Judicial* (Consejo Superior de la Judicatura, de diciembre de 2024) [🔗](#);

Conselho Nacional de Justiça (CNJ), *Resolução N° 615, de 11 de Março de 2025. Estabelece Diretrizes para o Desenvolvimento, Utilização e Governança de Soluções Desenvolvidas Com Recursos de Inteligência Artificial no Poder Judiciário*. (Conselho Nacional de Justiça (CNJ), 11 de marzo de 2025) [🔗](#).

Cabe destacar que los Tribunales de Nueva Zelanda publicaron tres directrices separadas para jueces, funcionarios judiciales, miembros de tribunales y personal de apoyo judicial; abogados; y, respectivamente, no abogados.

De manera similar, el Tribunal Federal de Canadá emitió "Principios y Directrices sobre el Uso de la Inteligencia Artificial por parte del Tribunal" y un Aviso a las partes, litigantes autorepresentados e intervinientes sobre el uso de la IA en procedimientos judiciales.

Asimismo, el Tribunal Supremo de Singapur, en septiembre de 2024, publicó su "Guía sobre el uso de herramientas de inteligencia artificial generativa por parte de los usuarios de tribunales" que se aplica a los usuarios de tribunales ("cualquier persona involucrada en un caso judicial, incluidos fiscales, abogados, personas que se representan a sí mismas o testigos") en "todos los asuntos del Tribunal Supremo, los tribunales estatales y los tribunales de justicia familiar".

En octubre de 2024, el Tribunal Superior de Justicia de Río Negro en Argentina aprobó su “Protocolo de Buenas Prácticas para el Uso de la Inteligencia Artificial Generativa (IAGen)” para magistrados, funcionarios y empleados.

Para las directrices emitidas por otros organismos públicos, véase: CEPEJ, *Carta Ética Europea sobre el Uso de la Inteligencia Artificial en los Sistemas Judiciales y su Entorno* (Comisión Europea para la Eficiencia de la Justicia (CEPEJ), 2019); CEPEJ-GT-CYBERJUST, *Uso de la Inteligencia Artificial Generativa (IA) por Profesionales Judiciales en un Contexto Laboral* (Comisión Europea para la Eficiencia de la Justicia (CEPEJ). Grupo de trabajo CEPEJ sobre Ciberjusticia e Inteligencia Artificial (CEPEJ-GT-CYBERJUST), 12 de febrero de 2024) [🔗](#).

- 10 Punto 8(a) del Anexo III del Reglamento de IA de la Unión Europea.
- 11 Véanse las secciones 2 a 5 del Capítulo III del Reglamento de IA de la Unión Europea
- 12 Bell y otros, *Toma de Decisiones por IA y los Tribunales: Guía para Jueces, Miembros de Tribunales y Administradores de Tribunales*; Juan David Gutiérrez, ‘Retos Éticos de la Inteligencia Artificial en el Proceso Judicial.’, en *derecho procesal. #NuevasTendencias. XLI Congreso Colombiano de Derecho Procesal*, ed. por ICDP (Instituto Colombiano de Derecho Procesal (ICDP) y Universidad Libre., 2020) [🔗](#); MSI-NET, *Algoritmos y Derechos Humanos - Estudio sobre las dimensiones de derechos humanos de técnicas automatizadas de procesamiento de datos y posibles implicaciones regulatorias* (Consejo de Europa, 2018) [🔗](#); OHCHR, *Taxonomía de los Riesgos de Derechos Humanos Relacionados con la IA Generativa - Suplemento al Documento Fundamental de B-Tech sobre el Desarrollo y Despliegue Responsables de la IA Generativa* (Oficina de Derechos Humanos de las Naciones Unidas para el Alto Comisionado (OHCHR), 2023) [🔗](#); Kalliopi Terzidou, ‘El uso de la inteligencia artificial en el poder judicial y su cumplimiento del derecho a un juicio justo’, *Journal of Judicial Administration*, 31.3 (2022), pp. 154–68; Giulia Gentile, ‘La IA en la sala de juicio y la independencia judicial: una perspectiva de la UE’, *EUideas*, 22 de agosto de 2022 [🔗](#).
- 13 Terzidou, ‘El uso de la inteligencia artificial en el poder judicial y su cumplimiento del derecho a un juicio justo’; Kalliopi Terzidou, ‘Anonimización

automatizada de decisiones judiciales: facilitando la publicación de sentencias judiciales mediante sistemas algorítmicos’, en *Actas de la Decimonovena Conferencia Internacional sobre Inteligencia Artificial y Derecho*, ICAIL ‘23 (Association for Computing Machinery, 2023), pp. 297–305, doi:10.1145/3594536.3595151.

- 14 Para un marco de diferentes tipos y clasificaciones de herramientas de IA, véase OCDE, ‘OECD Framework for the Classification of AI Systems’ (OECD Publishing, 2022) [🔗](#).
- 15 La resolución 78/265 de la Asamblea General de la ONU está disponible aquí: [🔗](#)
- 16 Los Principios de Bangalore establecen estándares para la conducta ética de los jueces. “En julio de 2006, el Consejo Económico y Social de las Naciones Unidas (ECOSOC) adoptó una resolución reconociendo los Principios de Bangalore como representativos de un desarrollo adicional y como complementarios a los Principios Básicos de las Naciones Unidas sobre la Independencia del Poder Judicial de 1985. La ECOSOC invitó a los Estados a animar a sus judicaturas a tener en cuenta los Principios al revisar o desarrollar normas respecto a la conducta judicial.” Los Principios de Bangalore están disponibles aquí: [🔗](#)
- 17 Sobre la utilidad y aceptabilidad del uso de herramientas de IA generativa por parte de árbitros y los posibles riesgos de adoptar estas tecnologías, véase João Ilhão Moreira y Jiawei Zhang, ‘ChatGPT como cuarto árbitro? La ética y los riesgos del uso de grandes modelos de lenguaje en el arbitraje’, *Arbitration International*, 2024, pp. 1–14, doi:10.1093/arbint/aiae031; David M. den Bakker, ‘Arbitraje e IA: El papel de los LLMs en la redacción de premios’ (tesis de máster inédita, Universidad Radboud de Nimega, 2024), doi:10.5281/zenodo.14601385.
- 18 Para directrices y recomendaciones emitidas por universidades, colegios de abogados, colegios de abogados y asociaciones profesionales, véase:

Bell y otros, *IA Decision-Making and the Courts: A Guide for Judges, Tribunal Members and Court Administrators*;

James E. Baker, Laurie Hobart y Mathew Mittels teadt, *IA para jueces: un marco* (Center for Security and Emerging Technology, diciembre de 2021) [🔗](#);

Juan G. Corvalán y Mariana Sánchez Caparrós, *Guía de Directrices Para El Uso de ChatGPT e IA Generativa de Texto En La Justicia* (UBA IALAB, 24 de noviembre de 2023) [🔗](#) ;

Fédération des Barreaux d'Europe, *European Lawyers in the Era of ChatGPT: Directrices sobre cómo los abogados deben aprovechar las oportunidades que ofrecen los grandes modelos de lenguaje y la IA generativa* (Comisión de Nuevas Tecnologías de la Asociación Europea de Abogados, 25 de julio de 2023) [🔗](#) ;

El Colegio de Abogados del Estado de California, *Recomendaciones del Comité de Responsabilidad y Conducta Profesional sobre la Regulación del Uso de IA Generativa por Licenciarios*, 16 de noviembre de 2023;

El Colegio de Abogados, *Consideraciones al utilizar ChatGPT y software de inteligencia artificial generativa basado en grandes modelos de lenguaje*, 30 de enero de 2024 [🔗](#) ;

Queensland Law Society, *Declaración de Orientación n° 37 Inteligencia artificial en la práctica jurídica*, 6 de junio de 2024 [🔗](#) ;

Asociación de Abogados de Nueva Gales del Sur, *Cuestiones derivadas del uso de modelos de lenguaje de IA (incluido ChatGPT) en la práctica jurídica*, 12 de julio de 2023 [🔗](#) ;

CIARB, *Directrices sobre el Uso de la IA en el Arbitraje* (The Chartered Institute of Arbitrators (CIARB), marzo de 2025) [🔗](#) .

La Asociación Americana para el Avance de la Ciencia (AAAS) desarrolló materiales para jueces (artículos y podcasts) sobre IA en los tribunales, incluyendo recomendaciones para el uso de la IA en litigios e investigación jurídica, entre otros. Los materiales están disponibles aquí: [🔗](#) .

19 UNESCO, 'Formación de Operadores Judiciales', 15 de octubre de 2024 [🔗](#) .

20 UNESCO, 'El MOOC Global de la UNESCO sobre IA y el Estado de Derecho Involucró a miles de operadores judiciales', 20 de abril de 2023 [🔗](#) .

21 UNESCO, *Iniciativa Global de Jueces de la UNESCO: Encuesta sobre el Uso de Sistemas de IA por Operadores Judiciales* (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), 2024) [🔗](#) .

22 UNESCO, 'Recomendación sobre la ética de la inteligencia artificial' (2022) [🔗](#) . Para una

evaluación de las herramientas de IA generativa a través de la recomendación de la UNESCO, véase: UNESCO, *Modelos de Fundación como ChatGPT a través del prisma de la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial* (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), 2023) [🔗](#) .

23 UNESCO, *Análisis del Impacto Ético: Una herramienta de la Recomendación sobre la Ética de la Inteligencia Artificial* (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), 2023) [🔗](#) .

24 UNESCO, *Kit de herramientas global sobre IA y el estado de derecho para el poder judicial*.

25 En el proceso de consulta pública de los borradores de las directrices, entre agosto y septiembre de 2024, la UNESCO recibió comentarios individuales de 99 personas y organizaciones con sede en Argentina, Australia, Bélgica, Brasil, Bulgaria, Camerún, Canadá, Colombia, Costa Rica, Dinamarca, República Dominicana, Egipto, El Salvador, Francia, Guatemala, India, Irlanda, Italia, Kenia, Lesoto, Liberia, Lituania, México, Nepal, Reino de los Países Bajos, Nueva Zelanda, Nigeria, Pakistán, Panamá, Paraguay, Perú, Filipinas, Polonia, Senegal, España, Suiza, Tanzania, Reino de los Países Bajos, Reino Unido de Gran Bretaña e Irlanda del Norte, Estados Unidos de América y Yemen. Además, la UNESCO recibió contribuciones de miembros del poder judicial en eventos organizados en India, Kenia, Madagascar y Nepal.

26 Para una explicación de lo que implica el "respeto, protección y promoción de los derechos humanos, las libertades fundamentales y la dignidad humana" en el contexto de las tecnologías de IA, consulte el Preámbulo y las secciones III.1 y III.2 de la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial. Para orientaciones sobre análisis de impacto ético, consulte el "Análisis de Impacto Ético: Una herramienta de la recomendación sobre la ética de la inteligencia artificial" de la UNESCO. Para una visión general de cómo los derechos humanos que pueden verse afectados por la IA generativa, véase ACNUDH, *Taxonomía de los Riesgos de Derechos Humanos Relacionados con la IA Generativa - Suplemento al Documento Fundamental de B-Tech sobre el Desarrollo y Despliegue Responsable de la IA Generativa*.

- 27 Para una explicación del principio de “equidad y no discriminación”, consulte la sección III.2 de la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial.
- 28 Ibid. Véase también, ONU, *Derechos Humanos en la Administración de Justicia: Informe del Secretario General* (Naciones Unidas (ONU), 7 de agosto de 2024) [🔗](#) ; sobre el sesgo de género, véase UNESCO, *Desafiando los prejuicios sistemáticos: Una investigación sobre el sesgo de género en los grandes modelos de lenguaje*, 2024, [🔗](#).
- 29 Los Principios de Bangalore proponen las siguientes aplicaciones del principio de “igualdad” que son pertinentes para estas Directrices: “5.1. Un juez debe ser consciente y comprender de la diversidad social y de las diferencias que surjan de diversas fuentes, incluyendo, pero no limitándose a, raza, color, sexo, religión, origen nacional, casta, discapacidad, edad, estado civil, orientación sexual, estatus social y económico y otras causas similares (“motivos irrelevantes”). 5.2. Un juez no deberá, en el desempeño de sus funciones judiciales, mediante palabras o conductas, manifestar parcialidad o prejuicio hacia ninguna persona o grupo por motivos irrelevantes. 5.3. Un juez ejercerá sus funciones judiciales con la debida consideración para todas las personas, como las partes, testigos, abogados, personal judicial y colegas judiciales, sin diferenciación por motivos irrelevantes que sean irrelevantes para el correcto desempeño de dichas funciones.” *Los Principios de Conducta Judicial de Bangalore* (Naciones Unidas, 2018), pp. 14–15 [🔗](#).
- 30 UNSDG, ‘No dejar a nadie atrás’, *Grupo de las Naciones Unidas para el Desarrollo Sostenible (UNSDG)* [🔗](#) [consultado el 26 de marzo de 2025].
- 31 Ibid.
- 32 Para una explicación del principio de “privacidad y protección de datos”, consulte la sección III.2 de la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial. Véase también, ONU, *Derechos Humanos en la Administración de Justicia: Informe del Secretario General*.
- 33 Para una explicación del principio de “proporcionalidad y no daño”, consulte la sección III.2 de la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial.
- 34 Un principio análogo denominado “principio de justificación” fue propuesto por Dan Jerker B. Svantesson, ‘Ciberdelitos y la adopción de sistemas de inteligencia artificial para la toma de decisiones judiciales en sistemas de justicia penal’, *Commonwealth Cybercrime Journal*, 1.1 (2023), pp. 152–79.
- 35 Para una explicación del principio de “seguridad y protección”, consulte la sección III.2 de la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial.
- 36 Para un análisis de riesgos publicado para autoridades que buscan incorporar modelos de IA generativa en sus flujos de trabajo, así como posibles contramedidas, véase BSI, *Generative AI Models. Oportunidades y riesgos para la industria y las autoridades* (Oficina Federal de Seguridad de la Información - Bundesamt für Sicherheit in der Informationstechnik (BSI), 4 de abril de 2024) [🔗](#). Para más información sobre las leyes de acceso a la información, consulte la página web de la UNESCO: [🔗](#).
- 37 Véase IA HLEG, *Directrices éticas para una IA de confianza* (Comisión Europea, 8 de abril de 2019), p. 17 [🔗](#).
- 38 IA HLEG, *Directrices éticas para IA confiable*.
- 39 Para una explicación del principio de “transparencia y explicabilidad”, consulte la sección III.2 de la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial. De manera similar, como argumentan las directrices publicadas por el Consejo Judicial Canadiense: “La explicabilidad es la necesidad de que estas herramientas de IA proporcionen explicaciones claras y comprensibles para su resultado, facilitando a los usuarios (y a quienes se ven afectados) interpretar, confiar, impugnar o aceptar la producción de IA en flujos de trabajo críticos. La explicabilidad es similar al requisito de que los jueces proporcionen explicaciones razonadas para sus decisiones jurídicas”. Véase, Consejo Judicial Canadiense, *Directrices para el uso de la inteligencia artificial en los tribunales canadienses*, p. 9.
- 40 Consulta AI HLEG, *Directrices éticas para IA de confianza*. Para una visión general del alcance y proceso de auditorías algorítmicas, véase Ada Lovelace Institute y DataKind UK, *Examining the Black Box: Tools for Evaluating Algorithmic Systems*, 29 de abril de 2020 [🔗](#). ONU, *Derechos*

Humanos en la Administración de Justicia: Informe del Secretario General.

41 ONU, *Derechos Humanos en la Administración de Justicia: Informe del Secretario General.*

42 Para una explicación del principio de “transparencia y explicabilidad”, consulte la sección III.2 de la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial. Para una visión general de los diversos instrumentos de transparencia algorítmica adoptados por el sector público en todo el mundo, véase GPAI, *Algorithmic Transparency: A State-of-the-Art Report*. Para una visión general del principio de “justicia abierta” en el contexto de los sistemas de IA, véase Brian M Barry, ‘IA para asistir a la toma de decisiones judiciales: implicaciones para el futuro de la justicia abierta’, *Australian Law Journal*, 98.9 (2024), pp. 656–69.

43 Para una explicación del principio “Conciencia y alfabetización”, consulte la sección III.2 de la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial. Véase también, ONU, *Derechos Humanos en la Administración de Justicia: Informe del Secretario General.*

44 Para una explicación del principio de “Responsabilidad y rendición de cuentas”, consulte la sección III.2 de la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial.

45 Para una discusión exhaustiva sobre qué significa “rendición de cuentas” en el contexto de la gobernanza de la IA, véase Claudio Novelli, Mariarosaria Taddeo y Luciano Floridi, ‘Accountability in Artificial Intelligence: What It Is and How It Works’, *AI & SOCIETY*, 2023, doi:10.1007/s00146-023-01635-y.

46 Para una explicación del principio de “supervisión humana”, consulte la sección III.2 de la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial. Véase también IA HLEG, *Directrices de ética para IA de confianza.*

47 Consulta AI HLEG, *Directrices éticas para IA de confianza.*

48 Para una explicación del principio de “Gobernanza y colaboración multiactor y adaptativa”, consulte la sección III.2 de la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial. Véase también IA HLEG, *Directrices éticas para IA de confianza*; UNESCO e i4Policy,

Desarrollo Multiactor de la IA: 10 Bases para el Diseño Inclusivo de Políticas (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), Innovation for Policy Foundation (i4Policy), 2022) [🔗](#) ; Anri Van der Spuy, *¿y si todos gobernamos Internet? Avanzar la participación multiactor en la gobernanza de Internet* (UNESCO, 2017) [🔗](#) .

49 Los Principios de Bangalore proponen las siguientes aplicaciones del principio de “independencia judicial” que son pertinentes para estas Directrices: “1.1. Un juez ejercerá la función judicial de forma independiente en función de la evaluación de los hechos por parte del juez y de acuerdo con una comprensión consciente de la ley, libre de influencias, incentivos, presiones, amenazas o interferencias externas, directas o indirectas, de cualquier parte o por cualquier motivo. 1.2. Un juez será independiente en relación con la sociedad en general y respecto a las partes particulares en un conflicto que el juez debe resolver. 1.3. Un juez no solo estará libre de conexiones inapropiadas y de influencia por parte de los poderes ejecutivo y legislativo del gobierno, sino que también deberá parecer a un observador razonable libre de ello.” *Los Principios de Conducta Judicial de Bangalore*, pp. 8–9.

50 Los Principios de Bangalore proponen las siguientes aplicaciones del principio de “imparcialidad” que son pertinentes para estas Directrices: “2.1. Un juez ejercerá sus funciones judiciales sin favor, 2.2. Un juez debe garantizar que su conducta, tanto dentro como fuera del tribunal, mantenga y refuerce la confianza del público, la profesión legal y los litigantes en la imparcialidad del juez y del poder judicial.” *Los Principios de Conducta Judicial de Bangalore*, p. 10.

51 El poder judicial debería considerar los sacrificios de adoptar herramientas hechas a medida frente a herramientas ya disponibles. Este último puede ser menos costoso y más rápido de implementar, pero menos transparente y con una personalización limitada. La primera puede ser más cara (aunque los costes pueden reducirse cuando se construye sobre código abierto) y requerir más tiempo de implementación (incluyendo desarrollo y pruebas), pero es más transparente y alineada con las necesidades y objetivos de la organización.

52 IA HLEG, *Directrices éticas para IA confiable.*

- 53 Por ejemplo, errores inducidos por humanos durante la selección o etiquetado de datos en herramientas de aprendizaje automático de aprendizaje supervisado; falta estadística de entrada, representación o datos en bruto que pueda reforzar el sesgo en herramientas de aprendizaje automático no supervisado; Errores impulsados por el entorno en herramientas de aprendizaje automático de refuerzo, etc.
- 54 “Las tarjetas modelo son documentos breves que acompañan modelos entrenados de aprendizaje automático que proporcionan evaluaciones comparativas en diversas condiciones, como en diferentes grupos culturales, demográficos o fenotípicos (por ejemplo, raza, ubicación geográfica, sexo, [...] Tipo de piel [...] y grupos interseccionales (por ejemplo, edad y raza, o sexo y tipo de piel de Fitzpatrick) que son relevantes para los dominios de aplicación previstos. Las tarjetas modelo también revelan el contexto en el que se pretende que se usen los modelos, detalles de los procedimientos de evaluación del rendimiento y otra información relevante.” Margaret Mitchell y otros, ‘Tarjetas modelo para informes modelo’ (presentado en la Conferencia sobre Justicia, Responsabilidad y Transparencia, Atlanta, GA, EE. UU., 2019), pp. 220–29, doi:10.1145/3287560.3287596. Véase también: GPAL, *Transparencia Algorítmica: Un informe de estado de la arte*; Matías Valderrama, María Paz Hermosilla y Romina Garrido, *Estado de la evidencia: Transparencia algorítmica* (Open Government Partnership y GobLab (Universidad Adolfo Ibáñez), mayo de 2023) ; Alianza para el Gobierno Abierto, ‘CAPÍTULO 8: Transparencia algorítmica’, en *The Skeptic’s Guide to Open Government*, 2022.
- 55 UNESCO, *Kit Global de Herramientas sobre IA y el Estado de Derecho para el Poder Judicial*, p. 42.
- 56 Mira, Brian Calvert, ‘La IA ya consume tanta energía como un país pequeño. Es solo el principio.’, *Vox*, 28 de marzo de 2024 ; Alexandra Sasha Luccioni, Sylvain Viguier y Anne-Laure Ligozat, ‘Estimando la huella de carbono de BLOOM, un modelo de lenguaje con parámetros 176B’, *Journal of Machine Learning Research*, 24.253 (2023), pp. 1–15; Sasha Luccioni, Yacine Jernite y Emma Strubell, ‘Procesamiento Consumido de Energía: ¿Watts impulsando el coste del despliegue de IA?’, en *Actas de la Conferencia ACM 2024 sobre Equidad, Responsabilidad y Transparencia*, FAccT ’24 (Association for Computing Machinery, 2024), pp. 85–99, doi:10.1145/3630106.3658542.
- 57 ONU, *Derechos Humanos en la Administración de Justicia: Informe del Secretario General*. 7 de agosto de 2024.
- 58 Para una visión general de los diferentes marcos de rendición de cuentas que pueden implementarse antes del despliegue de sistemas de IA, véase Dillon Reisman y otros, *Informe de Evaluaciones de Impacto Algorítmicas: Un marco práctico para la rendición de cuentas de agencias públicas* (AI Now Institute, 9 de abril de 2018) ; Instituto Ada Lovelace y DataKind UK, *Examinando la caja negra: herramientas para evaluar sistemas algorítmicos*; Alessandro Mantelero, ‘El análisis del impacto de los derechos fundamentales (FRIA) en el Reglamento de IA: raíces, obligaciones legales y elementos clave para una plantilla modelo’, *Computer Law & Security Review*, 54 (2024), p. 106020, doi:10.1016/j.clsr.2024.106020; UNESCO, *Análisis del Impacto Ético: Una Herramienta de la Recomendación sobre la Ética de la Inteligencia Artificial*; Gobierno de los Países Bajos, *Análisis del Impacto de Derechos Fundamentales y Algoritmos (FRAIA)* (Reino de los Países Bajos, 31 de julio de 2021) .
- 59 Una “evaluación algorítmica de impacto” es un proceso que evalúa los impactos futuros —a nivel individual, comunitario y social— que pueden estar asociados con el despliegue y uso de un sistema de IA e identifica cómo abordar posibles consecuencias negativas. La evaluación puede centrarse en implicaciones sociales y éticas, efectos económicos, impacto en los derechos humanos y consecuencias medioambientales. Reisman y otros, *Informe de Evaluaciones de Impacto Algorítmicas: Un marco práctico para la rendición de cuentas de agencias públicas*.
- 60 Una “evaluación algorítmica de riesgos” es un proceso que mapea riesgos específicos que pueden generarse con el despliegue y uso de un sistema de IA, analiza vulnerabilidades y posibles daños, e identifica las medidas necesarias para prevenir y abordar los riesgos. Las evaluaciones se centran en posibles errores o fallos del sistema, vulnerabilidades de seguridad y protección,

y brechas regulatorias. Instituto Ada Lovelace y DataKind Reino Unido, *examinando la caja negra: herramientas para evaluar sistemas algorítmicos*.

- 61 Un “análisis del impacto en derechos fundamentales” es un proceso que identifica, comprende y evalúa los posibles impactos adversos de los sistemas de IA en relación con los derechos fundamentales y prescribe cómo abordar y prevenir los impactos negativos. Mantelero, ‘El análisis del Impacto de los Derechos Fundamentales (FRIA) en el Reglamento de IA: Raíces, Obligaciones Legales y Elementos Clave para una Plantilla Modelo’.
- 62 ONU, *Derechos Humanos en la Administración de Justicia: Informe del Secretario General*, p. 19.
- 63 IA HLEG, *Directrices éticas para IA de confianza*, p. 31.
- 64 Además, la posible implicación de las herramientas de IA para el ecosistema de la información podría requerir una evaluación de riesgos más específica basada en las *Directrices para la Gobernanza de las Plataformas Digitales de la UNESCO: Salvaguardia de la Libertad de Expresión y el Acceso a la Información mediante un Enfoque Multiactor* (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), 2023).
- 65 Instituto Ada Lovelace y DataKind Reino Unido, *examinando la caja negra: herramientas para evaluar sistemas algorítmicos*.
- 66 Alessandro Mantelero, ‘Análisis del impacto en los derechos humanos y la IA’, en *Más allá de los datos: Derechos humanos, análisis de impacto ético y social en IA*, ed. por Alessandro Mantelero (T.M.C. Asser Press, 2022), pp. 45–91, doi:10.1007/978-94-6265-531-7_2; Alessandro Mantelero y Maria Samantha Esposito, ‘Una metodología basada en la evidencia para el análisis del impacto en derechos humanos (HRIA) en el desarrollo de sistemas intensivos en datos de IA’, *Computer Law & Security Review*, 41 (2021), p. 105561, doi:10.1016/j.clsr.2021.105561; Evaluación de Impacto en Derechos Fundamentales y Algoritmos del Gobierno de los Países Bajos (FRAIA); Gobierno de Canadá, *Herramienta de Análisis de Impacto Algorítmico* (Canadá, 12 de abril de 2023) .
- 67 Instituto Ada Lovelace y DataKind Reino Unido, *Examinando la caja negra: herramientas para evaluar sistemas algorítmicos*.
- 68 Una “evaluación del impacto en derechos humanos” es un proceso que identifica el impacto de los sistemas de IA en los titulares de derechos humanos y sociales, y ofrece recomendaciones sobre cómo prevenirlos o mitigarlos. Normalmente se realiza en respuesta a situaciones críticas. Mantelero, ‘El Análisis del Impacto de los Derechos Fundamentales (FRIA) en el Reglamento de IA: Raíces, Obligaciones Legales y Elementos Clave para una Plantilla Modelo’. Véase también, ONU, *Derechos Humanos en la Administración de Justicia: Informe del Secretario General*.
- 69 IA HLEG, *Directrices éticas para IA de confianza*, p. 31.
- 70 Alessandro Mantelero, Análisis del impacto en los derechos humanos y la IA’, en *Más allá de los datos: Derechos humanos, análisis de impacto ético y social en IA*, ed. por Alessandro Mantelero (T.M.C. Asser Press, 2022), pp. 45–91, doi:10.1007/978-94-6265-531-7_2; Alessandro Mantelero y Maria Samantha Esposito, ‘Una metodología basada en la evidencia para el análisis del impacto en derechos humanos (HRIA) en el desarrollo de sistemas intensivos en datos de IA’, *Computer Law & Security Review*, 41 (2021), p. 105561, doi:10.1016/j.clsr.2021.105561; Evaluación de Impacto en Derechos Fundamentales y Algoritmos del Gobierno de los Países Bajos (FRAIA); Gobierno de Canadá, *Herramienta de Análisis de Impacto Algorítmico* (Canadá, 12 de abril de 2023) .
- 71 Por ejemplo, el 21 de junio de 2022, el Tribunal de Justicia de la Unión Europea (TJUE) emitió una sentencia que prohíbe instrumentos específicos de “detección predictiva automática de amenazas” (herramientas que “filtran automáticamente grandes cantidades de datos para predecir posibles amenazas a la seguridad pública”). El TJUE concluyó que las herramientas de IA basadas en aprendizaje automático “capaces de modificarse sin intervención humana ni revisión del proceso de evaluación” eran incompatibles con los derechos a la privacidad y la protección de datos personales. El TJUE argumentó que “la opacidad que caracteriza la forma en que funciona la tecnología de inteligencia artificial ... pueden privar también a los sujetos de los datos de su derecho

a un recurso judicial efectivo". La sentencia del TJUE está disponible aquí: [🔗](#) Para resumen y análisis del caso, véase Christian Thönnies y Niovi Vavoula, 'Detección Predictiva Automatizada de Amenazas tras la Ligue des Droits Humains: Implicaciones para ETIAS y CSAM (Parte I)', *VerfBlog*, 8 de mayo de 2023 [🔗](#).

72 ONU, *Derechos Humanos en la Administración de Justicia: Informe del Secretario General*, p. 21.

73 Para un relato detallado de las vulnerabilidades de las herramientas de IA generativa, véase BSI, *Modelos de IA Generativa. Oportunidades y riesgos para la industria y las autoridades*.

74 IA HLEG, *Directrices éticas para IA de confianza*, p. 17.

75 Véase Terzidou, 'Anonimización automatizada de decisiones judiciales: facilitando la publicación de decisiones judiciales mediante sistemas algorítmicos'.

76 Por ejemplo, la Comisión Europea para la Eficiencia de la Justicia del Consejo de Europa (CEPEJ) y el poder judicial brasileño publicaron repositorios que documentan los sistemas de IA adoptados en sus jurisdicciones.

77 Para una revisión de diversos instrumentos de transparencia algorítmica, con énfasis en repositorios o registros de algoritmos públicos, que buscan divulgar información sobre los sistemas de IA adoptados en el sector público, véase GPAI, *Algorithmic Transparency: A State-of-the-Art Report*.

78 UNESCO, *Datos Abiertos para la IA: ¿Y ahora qué?* (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), 2023) [🔗](#).

79 Véase, por ejemplo, el plan de estudios del MOOC Global de la UNESCO sobre IA y Estado de Derecho. UNESCO, 'El MOOC Global de la UNESCO sobre IA y el Estado de Derecho implicó a miles de operadores judiciales'.

80 Instrumentos de evaluación de impacto, como la "Análisis de Impacto Ético: Herramienta de la Recomendación sobre la Ética de la Inteligencia Artificial" de la UNESCO y el marco de análisis del impacto en derechos humanos propuesto por Mantelero, pueden utilizarse para evaluar los impactos en los derechos humanos de las herramientas empleadas por los operadores

judiciales. Véase Mantelero, 'Análisis del impacto en derechos humanos e IA'.

81 En este sentido, los Principios de Bangalore proponen las siguientes aplicaciones del principio de "competencia y diligencia" que son pertinentes para estas Directrices: "6.3. Un juez deberá tomar medidas razonables para mantener y mejorar los conocimientos, habilidades y cualidades personales necesarios para el adecuado desempeño de sus funciones judiciales, aprovechando para ese fin la formación y otras facilidades que deberían estar disponibles bajo control judicial para los jueces." *Los Principios de Conducta Judicial de Bangalore*, p. 16.

82 Juan David Gutiérrez, '¿Están los Jueces En Capacidad de Usar Modelos de Lenguaje a Gran Escala (LLMs)?', *Revista EXCEJENCIA*, 7 (2023), pp. 10–15; David Bawden, 'Orígenes y conceptos de la alfabetización digital', en *Digital Literacies: Concepts, Políticas y Prácticas*, editado por Colin. Lankshear y Michele. Knobel (Peter Lang, 2008), pp. 17–32 [🔗](#).

83 Mantelero, 'Análisis del Impacto en Derechos Humanos e IA'; Mantelero y Esposito, 'Una metodología basada en la evidencia para la análisis del impacto en derechos humanos (HRIA) en el desarrollo de sistemas intensivos en datos de IA'.

84 Para recomendaciones generales sobre cómo los operadores judiciales pueden evaluar cuestiones de privacidad y protección de datos, véase UNESCO, *Directrices para Actores Judiciales sobre Privacidad y Protección de Datos* (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), 2022) [🔗](#).

85 Cortes y Tribunales del Reino Unido Judicial, *Inteligencia Artificial (IA) - Orientación para titulares de cargos judiciales*, p. 3.

86 Véase Tribunales de Nueva Zelanda, *Directrices para el uso de la inteligencia artificial generativa en cortes y tribunales*; Cortes y Tribunales del Reino Unido Judicial, *Inteligencia Artificial (IA) - Orientación para titulares de cargos judiciales*, p. 6.

87 Cortes y Tribunales del Reino Unido Judicial, *Inteligencia Artificial (IA) - Orientación para titulares de cargos judiciales*, p. 3.

88 Tribunales de Nueva Zelanda, *Directrices para el uso de la inteligencia artificial generativa en cortes y tribunales*, p. 2.

- 89 Para una orientación similar sobre la generación de pruebas dirigidas a los usuarios de tribunales, véase Tribunal Supremo de la República de Singapur, *Guía sobre el uso de herramientas de inteligencia artificial generativa por usuarios de tribunales*.
- 90 Las *Directrices* no aluden al término “alucinación” porque este uso contribuye a antropomorfizar sistemas de IA, especialmente los chatbots.
- 91 BSI, *Modelos de IA Generativa. Oportunidades y riesgos para la industria y las autoridades*, p. 9; Véase también, UNESCO, *Desafiando los prejuicios sistemáticos: Una investigación sobre el sesgo de género en grandes modelos de lenguaje*, 2024, [🔗](#).
- 92 “Los LLMs generalmente generan textos lingüísticamente correctos y convincentes, y son capaces de hacer afirmaciones sobre una amplia variedad de temas. Esto puede crear la impresión de un rendimiento similar al humano, lo que lleva a una confianza excesiva en las afirmaciones y el rendimiento de los modelos (lo que se denomina sesgo de automatización). Para los usuarios, esto puede resultar en sacar conclusiones incorrectas de los textos generados o aceptar afirmaciones sin cuestionarlas” BSI, *Modelos de IA Generativa. Oportunidades y riesgos para la industria y las autoridades*, p. 10.
- 93 BSI, *Modelos de IA Generativa. Oportunidades y riesgos para la industria y las autoridades*, p. 9.
- 94 Por ejemplo, para directrices sobre cómo citar la salida de chatbots de IA con el sistema de citación APA, véase Timothy McAdoo, ‘Cómo citar ChatGPT’, *APA Style*, 2024.
- 95 Para una orientación similar, véase Tribunal Supremo del Estado de Delaware, *Política Interina sobre el Uso de GenAI por Funcionarios Judiciales y Personal de los Tribunales*.
- 96 Emily M. Bender y otros, ‘Sobre los peligros de los loros estocásticos: ¿Pueden los modelos de lenguaje ser demasiado grandes?, en *Actas de la Conferencia ACM 2021 sobre Equidad, Responsabilidad y Transparencia*, FAccT ‘21 (Association for Computing Machinery, 2021), pp. 610–23 (p. 616), doi:10.1145/3442188.3445922.
- 97 Matthew Dahl y otros, ‘Grandes ficciones legales: perfilando alucinaciones legales en grandes modelos de lenguaje’, 2024 [🔗](#); Varun Magesh y otros, ‘¿Sin alucinaciones? Evaluando la fiabilidad de las principales herramientas de investigación legal en IA’, 2024.

La IA ha entrado en la sala del tribunal. Jueces y profesionales del derecho han empezado a utilizar sistemas de IA para ahorrar tiempo y centrarse en tareas importantes. Este cambio tecnológico puede conducir a mejoras significativas en los derechos humanos y en la impartición de la justicia.

Sin embargo, ¿cómo podemos asegurarnos de que los beneficios de la IA superen sus riesgos (sesgos, discriminación, errores y alucinaciones)?

Los tribunales tienen el poder de controlar y moldear cómo se desarrollan y utilizan los sistemas de IA para salvaguardar la independencia judicial, el Estado de derecho y los derechos humanos.

Los actores del sector de justicia han instado a la UNESCO a desarrollar estas Directrices para proporcionar recomendaciones y principios clave sobre los pasos necesarios a considerar antes, durante y después del despliegue de sistemas de IA.

Estas Directrices ayudarán a los líderes judiciales a diseñar marcos políticos relevantes para el contexto local. Ayudarán a jueces, personal judicial y profesionales del derecho, independientemente de sus habilidades digitales, a desarrollar un uso profesional de la IA.

El programa de IA y Estado de Derecho de la UNESCO dota a jueces y profesionales del derecho de todo el mundo con las habilidades necesarias para defender la justicia, el Estado de derecho y los derechos humanos en la era de la IA.

DIRECTRICES PARA EL USO DE SISTEMAS DE IA EN CORTES Y TRIBUNALES



unesco