



La Comisión de Coordinación y Seguimiento examinó y aprobó, como cuarto proyecto integrado, la iniciativa denominada “Marco Iberoamericano de Riesgos, Sesgos, Evidencia Digital y Protección de Personas Vulnerables”, destacando su carácter integrador al reunir propuestas vinculadas con inteligencia artificial, discriminación algorítmica, sesgos, tutela de grupos vulnerables, deepfakes y tratamiento probatorio de la evidencia digital.

La iniciativa surge ante la necesidad de ofrecer respuestas judiciales más sólidas frente a los riesgos derivados del uso de la inteligencia artificial, especialmente cuando se encuentran comprometidos derechos fundamentales como la igualdad, la no discriminación, la privacidad, la autenticidad de la prueba y la protección reforzada de personas y grupos vulnerables.

El proyecto propone desarrollar un marco regional iberoamericano orientado al uso ético, transparente y responsable de la inteligencia artificial en la justicia, mediante la elaboración de estándares judiciales, herramientas tecnológicas, lineamientos éticos y mecanismos de cooperación regional.

Este proyecto integrado articula y absorbe iniciativas presentadas por Perú, Colombia, Argentina, Chile, Panamá y Portugal, relacionadas con inteligencia artificial y derechos fundamentales, enfoques de igualdad y accesibilidad, uso respetuoso de los derechos humanos de personas y grupos vulnerables, discriminación algorítmica, radar de ética y sesgos en decisiones judiciales, y observatorios de evidencia digital centrados en deepfakes judiciales.

Propuesta.

Guía Judicial Iberoamericana sobre Inteligencia Artificial y Derechos Humanos con enfoque en Acceso a la Justicia y Protección de Personas en Condición de Vulnerabilidad

1. Introducción

La transformación digital de los sistemas de justicia en Iberoamérica, impulsada por el creciente uso de herramientas de inteligencia artificial (IA) y tecnologías basadas en datos, plantea oportunidades relevantes para mejorar la eficiencia, la coherencia y el acceso a la justicia. Sin embargo, también introduce riesgos significativos que pueden afectar el ejercicio de los derechos fundamentales, particularmente en relación con la igualdad, el acceso a la justicia, la no discriminación, la privacidad, la transparencia y la garantía del debido proceso.

En la literatura especializada se ha identificado que los modelos de IA pueden profundizar sesgos o prejuicios humanos que existen frente a las poblaciones históricamente vulneradas. Estos riesgos se intensifican cuando las decisiones o procesos asistidos por tecnologías afectan a **personas en condición de vulnerabilidad**, tales como:

- Niños, niñas y adolescentes
- Mujeres víctimas de violencia de género
- Personas mayores
- Personas con discapacidad

De esta forma, el proyecto incorpora una mirada específica sobre los derechos de dichas personas en condición de vulnerabilidad, en tanto enfrentan riesgos particulares en los entornos digitales y en el uso de tecnologías basadas en inteligencia artificial.

En estos contextos, el uso inadecuado de la inteligencia artificial puede no solo reproducir desigualdades estructurales, sino también profundizarlas, generar nuevas formas de exclusión o limitar el acceso efectivo a la justicia.

Este proyecto tiene como justificación la heterogeneidad y necesidad de guía de uso de IA en el sector de la justicia. La adopción de IA en los poderes judiciales de la región es desigual en lo regulatorio y lo operativo. La UNESCO Global Judges' Initiative: Survey on the Use of AI Systems by Judicial Operators reporta creciente uso de herramientas de IA y, a la vez, una fuerte demanda de directrices y formación para su empleo responsable por parte de jueces y personal judicial. Y el Código Ético Iberoamericano para el uso de IA en el ámbito judicial exige justamente control

humano significativo, transparencia y registro, evaluación y monitoreo continuo, y suspensión cuando exista riesgo significativo, por lo que contar con evidencia comparada (mapeo) y una lista de verificación es el paso natural para aterrizar esos mandatos.

Además, existe un riesgo de sesgos y discriminación hacia poblaciones históricamente vulneradas. Los documentos de UN Women/UNESCO sobre impactos de género y las investigaciones sectoriales (No Woman Left Behind) muestran brechas de participación y liderazgo, así como sesgos en datos y sistemas que afectan de manera desproporcionada a las mujeres.

En general, la literatura especializada advierte sobre los riesgos de sesgos y discriminación hacia poblaciones históricamente vulneradas, entre ellas los niños, niñas y adolescentes y las personas con discapacidad.

A su vez, los principios de la CEPEJ (2018), también acogidos en el Código Ético (derechos, no discriminación, calidad/seguridad de datos, transparencia/imparcialidad y “bajo control del usuario”) ofrecen una guía concreta para evaluar y mitigar sesgos también en el caso de personas con discapacidad (accesibilidad, comprensibilidad y trazabilidad). No obstante, hay un déficit de evidencia y coordinación. La OCDE Governing with AI (2024) subraya la escasez de evaluaciones empíricas y de métricas comparables en justicia.

Por su parte la Observación General N.º 6 del Comité sobre los Derechos de las Personas con Discapacidad establece que los Estados partes tienen obligaciones de información en relación con el artículo 5 de la Convención, ya que deben recopilar y analizar datos e información de investigación adecuados para identificar desigualdades, prácticas discriminatorias y patrones de desventaja, y analizar la eficacia de las medidas que promueven la igualdad.¹

Asimismo, establece que las medidas específicas que adopten los Estados partes deben ser compatibles con todos sus principios y disposiciones. En particular, no deben dar lugar a la perpetuación del aislamiento, la segregación, los estereotipos, la estigmatización ni ninguna otra forma de discriminación contra las personas con discapacidad.²

En relación con los NNA, el uso de IA en el sistema judicial requiere especial cautela. Conforme al Código Ético Iberoamericano, toda herramienta de IA debe estar sujeta

a control humano posterior, operando como asistencia y no como sustituto de la labor jurisdiccional. En este sentido, los casos que involucran derechos de NNyA demandan una diligencia reforzada, en consonancia con los estándares internacionales de protección.

La Observación General Nro. 12 del Comité de los Derechos del Niño dispone que “los Estados Partes garantizarán al niño que esté en condiciones de formarse un juicio propio, el derecho de expresar su opinión libremente en todos los asuntos que afectan al niño, teniéndose debidamente en cuenta las opiniones del niño en función de la edad y madurez”. En esa misma línea, la Corte Interamericana de Derechos Humanos ha indicado que el derecho debe ser interpretado a la luz del art.12 de la Convención sobre los Derechos del Niño, que establece el derecho a ser oído, con el objeto de que la intervención de la niña o niño se ajuste a las condiciones de este y no redunde en perjuicio de su interés genuino. 3

Por su parte las 100 Reglas de Brasilia sobre acceso a Justicia en condición de vulnerabilidad establecen que “Todo niño, niña y adolescente debe ser objeto de una especial tutela por parte de los órganos del sistema de justicia en consideración a su desarrollo evolutivo.4 Es decir, la participación efectiva de los niños en el proceso implica el acceso a la justicia, el respeto a su dignidad al considerarlos verdaderos protagonistas, dejando de ser meros espectadores a la luz del nuevo paradigma de la Convención de los Derechos del Niño.5

El Comité de los Derechos del Niño, en su Observación General nro. 25 (2021), ha enfatizado que los derechos reconocidos en la Convención sobre los Derechos del Niño se aplican plenamente al entorno digital. Asimismo, ha instado a los Estados a garantizar su cumplimiento mediante medidas legislativas, administrativas, judiciales y educativas. 6

En consecuencia, las obligaciones derivadas de los instrumentos internacionales de derechos humanos deben guiar la aplicación de la IA en los poderes judiciales, concretando en cada caso el respeto y la vigencia del interés superior y la participación del niño en el proceso como principios rectores de toda decisión judicial.

Frente a este escenario, los sistemas judiciales tienen el deber de asegurar que la incorporación de tecnologías digitales sea **compatible con los estándares de derechos humanos** y, en particular, con las **Reglas de Brasilia sobre Acceso a la**

Justicia de las Personas en Condición de Vulnerabilidad, que establecen la obligación de remover barreras y garantizar un acceso real, efectivo y no discriminatorio.

2. Finalidad de la Guía

La presente Guía tiene como propósito brindar a juezas, jueces, fiscales y demás operadores judiciales un **instrumento práctico, claro y accesible** que permita:

- Comprender los impactos de la inteligencia artificial en la función judicial
- Identificar riesgos asociados al uso de sistemas automatizados
- Adoptar decisiones informadas y respetuosas de los derechos fundamentales
- Garantizar el acceso a la justicia en condiciones de igualdad
- Prevenir la discriminación algorítmica y la revictimización
- Asegurar el tratamiento adecuado de la evidencia digital

A diferencia de otros instrumentos de carácter técnico, esta Guía **no busca explicar el funcionamiento de la tecnología**, sino orientar la actuación judicial frente a sus efectos.

3. Enfoque del documento

La Guía adopta un enfoque:

a) Centrado en derechos humanos

Todas las recomendaciones se fundamentan en estándares constitucionales e internacionales de derechos humanos.

b) Basado en las Reglas de Brasilia

Se incorpora de manera transversal el principio de acceso efectivo a la justicia, mediante:

- accesibilidad
- lenguaje claro
- ajustes razonables
- trato digno
- protección diferenciada

c) Enfocado en personas en situación de vulnerabilidad

Se prioriza el análisis del impacto de la tecnología en:

- infancia y adolescencia
- violencia de género
- envejecimiento
- discapacidad

d) Orientado a la toma de decisiones judiciales

El contenido está diseñado para responder a una pregunta clave:

¿Qué debe hacer una jueza o juez cuando una tecnología incide en el caso?

4. Estructura de la Guía

La Guía se organiza en cuatro módulos integrados, que combinan estándares jurídicos, evidencia comparada y herramientas prácticas.

Módulo 1: Estándares jurídicos y acceso a la justicia

Desarrolla los principios que deben orientar el uso de inteligencia artificial en la justicia, incluyendo:

- Igualdad y no discriminación
- Acceso efectivo a la justicia
- Derecho a la explicación
- Supervisión humana obligatoria
- Accesibilidad y ajustes razonables

Incorpora aplicaciones específicas para:

- Niños, niñas y adolescentes (interés superior del niño)
- Mujeres víctimas de violencia de género (no revictimización)
- Personas mayores (trato digno y accesibilidad)
- Personas con discapacidad (inclusión plena y autonomía)

Módulo 2: Mapeo regional y buenas prácticas

Presenta un informe que contenga las iniciativas de inteligencia artificial en sistemas judiciales de Iberoamérica, incluyendo:

- Casos de uso reales
- Buenas prácticas identificadas
- Salvaguardas implementadas
- Riesgos observados

Incluye fichas simplificadas para facilitar su comprensión y replicabilidad.

Consideramos importante contar con evidencia del uso de la IA en la región para conocer las tendencias regulatorias y de desarrollo de la tecnología, para así analizar la viabilidad de aplicar un especial enfoque diferencial que abarque el ciclo de vida de la IA.

Módulo 3: Riesgos, evidencia digital y herramientas para la decisión judicial

Este es el núcleo práctico de la Guía. Integra en un mismo espacio:

a) Identificación de riesgos

Explicados en lenguaje claro, incluyendo:

- Sesgos y discriminación algorítmica
- Revictimización
- Exclusión digital
- Falta de transparencia
- Uso indebido de datos sensibles

b) Criterios para la actuación judicial

Orientaciones sobre:

- Cuándo exigir intervención humana
- Cómo evaluar decisiones automatizadas
- Cuándo solicitar peritajes
- Cómo actuar frente a riesgos de discriminación

c) Lista de verificación

Herramienta práctica con preguntas clave aplicables a cualquier caso.

Módulo 4: Decálogo para el uso de la IA en la justicia

La Guía concluye con un conjunto de principios orientadores, entre ellos:

1. Acceso efectivo a la justicia
2. No discriminación algorítmica
3. Interés superior del niño
4. Protección frente a la violencia de género
5. Accesibilidad universal
6. Transparencia y explicabilidad
7. Supervisión humana obligatoria
8. Protección de datos sensibles
9. Uso responsable de evidencia digital
10. Responsabilidad institucional

5. Carácter del instrumento

Esta Guía es:

- **Práctica:** orientada a la toma de decisiones judiciales
- **Accesible:** redactada en lenguaje claro
- **Aplicable:** utilizable en casos concretos
- **Integradora:** combina estándares, evidencia y herramientas
- **Adaptable:** pensada para distintos contextos judiciales

6. Impacto esperado

La implementación de esta Guía permitirá:

- Fortalecer la confianza en la justicia en entornos digitales
- Reducir riesgos de discriminación y exclusión
- Mejorar la calidad de las decisiones judiciales
- Garantizar el acceso a la justicia para personas vulnerables
- Promover un uso ético, responsable y transparente de la IA

7. Planificación del proyecto

Fase 1. Diseño metodológico

Meses 1 a 3

Actividades

- Definición de la estructura final de la Guía
- Diseño metodológico del mapeo regional
- Desarrollo de la matriz de riesgos
- Definición del enfoque transversal basado en las Reglas de Brasilia
- Elaboración de herramientas de trabajo:
 - fichas de buenas prácticas
 - fichas de riesgos
 - estructura para casos tipo

Productos

- Índice definitivo de la Guía
- Documento metodológico aprobado
- Plantillas y herramientas de levantamiento de información

Fase 2. Mapeo regional y recopilación de buenas prácticas

Meses 4 a 8

Actividades

- Identificación de iniciativas de IA en sistemas de justicia
- Sistematización de información sobre:
 - objetivos
 - tecnologías utilizadas
 - grupos impactados
 - salvaguardas implementadas
- Identificación de riesgos asociados a cada iniciativa
- Análisis comparado de marcos regulatorios
- Clasificación por tipo de uso y población en situación de vulnerabilidad

Productos

- Repositorio regional estructurado (al menos 10–15 iniciativas)
- Fichas de buenas prácticas
- Informe comparado de tendencias regulatorias y tecnológicas

Fase 3. Desarrollo del núcleo práctico de la Guía

Meses 9 a 12

Actividades

1. Identificación de riesgos
2. Traducción de riesgos en criterios judiciales
3. Desarrollo de herramientas prácticas
4. Elaboración de casos tipo

Productos

- Matriz de riesgos validada
- Lista de verificación judicial
- Casos prácticos desarrollados
- Lineamientos sobre tratamiento de evidencia digital

Fase 4. Validación práctica con operadores judiciales

Meses 13 a 15

Actividades

- Aplicación de simulaciones de casos
- Validación de:
 - claridad del lenguaje
 - utilidad del checklist
 - pertinencia de los criterios
- Recopilación de retroalimentación
- Ajuste del contenido de la Guía

Productos

- Informe de validación
- Ajustes metodológicos
- Versión mejorada del Módulo 3

Fase 5. Integración y redacción final de la Guía

Meses 16 a 17

Actividades

- Redacción final de los módulos:
 - Módulo 1: estándares jurídicos y acceso a la justicia
 - Módulo 2: mapeo regional y buenas prácticas
 - Módulo 3: riesgos y herramientas prácticas

- Módulo 4: decálogo regional
- Integración del repositorio como anexo
- Revisión jurídica y editorial

Productos

- Versión final de la Guía Judicial
- Decálogo integrado
- Anexos con información comparada

Fase 6. Validación institucional y publicación

Mes 18

Actividades

- Validación final por los países participantes
- Elaboración de versión accesible (lenguaje claro y formatos inclusivos)

Productos

- Guía publicada oficialmente

Cronograma resumido

Fase 1: Meses 1 a 3

Fase 2: Meses 4 a 8

Fase 3: Meses 9 a 12

Fase 4: Meses 13 a 15

Fase 5: Meses 16 a 17

Fase 6: Mes 18