



MEJORES DATOS, MEJOR IA

**Necesidades de datos para una
IA de justicia responsable**

JULIO GABRIEL MERCADO

Esta investigación fue financiada por el Registro de Direcciones de Internet para América Latina y el Caribe (LACNIC).

Descargo de responsabilidad: Los puntos de vista y opiniones expresados en este informe son del autor y no reflejan necesariamente la política o posición de LACNIC.

**Mejores datos, mejor IA:
necesidades de datos para una IA de justicia responsable**

Índice

Introducción	3
Listado de definiciones (en orden alfabético)	5
Parte I. Transformación digital, justicia abierta e inteligencia artificial	7
Transformación digital centrada en las personas	7
Justicia Abierta	8
Inteligencia Artificial en la Justicia	9
Parte II. Datos, sesgos y gobernanza de la IA de justicia	12
Datos e IA de justicia	12
Sesgos basados en datos	14
Parte III. Dimensiones clave en la publicación y uso de datos para la IA de justicia	16
Principios para la publicación de datos de justicia	16
Niveles para la aplicación de los principios	21
Conclusión	24
Bibliografía	26

Introducción

De entre todas las tecnologías legales disruptivas (Susskind, 2013) la inteligencia artificial (IA) es posiblemente la que mayores cuestiones éticas (y ansiedades) plantea. Los cuestionamientos se dirigen a sus aplicaciones actuales, pero también a posibles usos futuros que aún no se han materializado, e incluso a la percepción de las personas respecto de su eventual aplicación (Nyholm, 2024). En este sentido, se habla de la adopción de herramientas de IA en la toma de decisiones públicas como un *arma de doble filo*, que por una parte brinda la oportunidad de abordar desafíos globales y mejorar la efectividad de la justicia, pero que por otra parte precisa de marcos de gobernanza robustos y eficaces, de compleja articulación, particularmente si lo que se busca es lograr resultados equitativos e inclusivos (Engin, 2024).

La implementación de IA en el contexto del sector justicia en América Latina enfrenta desafíos particulares. Por un lado, mientras que las instituciones judiciales expresan cada vez más una voluntad de llevar a cabo una transformación digital abierta y centrada en las personas a las que deben servir, deben hacerlo en un contexto signado por los recursos en general insuficientes. Por otro lado, mientras que la IA genera grandes expectativas como un medio para incrementar los niveles de eficiencia de la administración de justicia, y por ende de acceso a la justicia de las personas, no existe todavía un marco claro para esta adopción que incluya una definición clara del rol de las instituciones judiciales en el proceso, ni de las precondiciones necesarias para que dicha adopción sea realizada de manera responsable.

Un aspecto fundamental a considerar al respecto es el de los datos. Los datos con los cuales se entrenan los sistemas de IA determinan en gran medida los resultados, tanto positivos como negativos, que estos sistemas pueden producir. En este sentido, la gobernanza de los datos no solo moldea en última instancia la gobernanza de la IA, sino que asimismo define en buena medida hasta qué punto sus beneficios se distribuirán de manera equitativa y podrán mitigarse los riesgos asociados con su implementación (Datasphere Initiative, 2024). Entretanto, mientras que su papel elemental en el desarrollo y entrenamiento de sistemas de IA es aceptado como un axioma, la percepción respecto de su valor y sobre la necesidad de establecer mecanismos conscientes y organizados para gobernarlos es todavía incipiente.

En el caso de la justicia, mientras que la Justicia Abierta lleva algunos años orientando a las instituciones judiciales hacia la apertura y disponibilización de los datos que producen, la simple publicación no es suficiente para asegurar una reutilización efectiva en esquemas de IA que aseguren resultados equitativos. Por el contrario, para que esto suceda es necesario establecer infraestructuras robustas, principios de gobernanza y de uso ético de esos datos, que orienten la forma en que las instituciones judiciales los generan, publican y gestionan (Open Data Institute, 2024). Sin un marco adecuado que aborde estas cuestiones, el riesgo de que los esfuerzos de publicación de datos por parte de las instituciones judiciales den como resultado datos fragmentados, inconsistentes o insuficientemente representativos es alto, lo que podría limitar de manera sustancial el desarrollo de una IA responsable que brinde soluciones de justicia inclusivas.

Este trabajo tiene por objetivo brindar a las instituciones judiciales una serie de principios fundamentales, que deben tener en cuenta para gobernar sus datos de manera que estos puedan servir de insumo en el desarrollo de sistemas de IA responsable. En este sentido, plantea cinco principios a ser observados por las

instituciones judiciales a la hora de gestionar y publicar sus datos, al tiempo que establece distintos niveles en los cuales estos principios deben dirimirse. La finalidad de este planteo es garantizar que el uso de IA pueda darse en el marco de procesos de transformación digital de la administración de justicia transparentes, colaborativos y centrados en las personas, que beneficien de manera equitativa a todas las personas, particularmente a aquellas en situación de mayor vulnerabilidad. Al proponer principios para la gestión de datos y abordar los desafíos éticos asociados con la IA, este trabajo busca contribuir a la creación de marcos regulatorios que no solo impulsen la innovación en la administración de justicia, sino que también refuercen los derechos humanos y la inclusión digital¹.

Las fuentes que informan este trabajo son, por una parte, el ingente corpus de conocimiento que se genera cada día sobre las temáticas de IA, justicia y datos (según la opinión de una de las personas entrevistadas, virtualmente inabarcable dada la velocidad de su desarrollo), junto con la normativa sectorial más relevante hasta el momento. Por otra parte, este trabajo contó como fuente primaria de información con una serie de seis entrevistas realizadas a varias de las personas (académicas, investigadoras, especialistas) más destacadas del panorama regional en lo que refiere a la intersección entre estas tres temáticas².

A esta **Introducción** le sigue un **Listado de definiciones** que busca establecer un punto de partida común para la lectura de este trabajo y facilitar la comprensión del lenguaje específico utilizado.

A continuación, la **Parte I** aborda la transformación digital en la justicia, enfocándose en tres dimensiones consideradas clave. La primera es la necesidad de adoptar un enfoque de transformación digital centrada en las personas para la administración de justicia, que mejore su eficiencia, accesibilidad y transparencia, priorizando a la vez el acceso equitativo y adecuado a los servicios y bienes públicos. La segunda dimensión es la visión de Justicia Abierta, que promueve principios como la transparencia, la rendición de cuentas, la participación pública y la colaboración en la justicia, con énfasis en la publicación y disponibilización de datos judiciales reutilizables. Finalmente, se analiza el impacto de los sistemas de IA en la administración de justicia, destacando de qué manera su uso debe lidiar con diversos desafíos éticos para poder ser considerado responsable.

La **Parte II** del trabajo profundiza en la cuestión de los datos, sesgos y gobernanza en el contexto de la IA aplicada a la justicia. Para ello, en primer lugar se examina el papel fundamental de los datos en el desarrollo de sistemas de IA y cómo aspectos tales como su calidad, integridad y transparencia son cruciales para un uso que pueda ser considerado responsable. Se analizan documentos clave como la *Carta Ética Europea sobre el Uso de la Inteligencia Artificial en los Sistemas Judiciales* de 2018, o el *Reglamento Europeo de Inteligencia Artificial* de 2024, que establecen principios éticos sobre el manejo de datos y la necesidad de su alta calidad para garantizar la precisión y equidad de las decisiones judiciales asistidas por IA. A

¹ El germen de este trabajo fue la jornada sobre “Usos responsables de la IA en la Justicia: hacia la co-creación de herramientas y directrices para una transformación digital sostenible”, que tuvo lugar el día 24 de agosto de 2023 en el Centro de Estudios en Tecnología y Sociedad (CETyS) de la Universidad de San Andrés, Argentina.

² Las entrevistas fueron realizadas durante los meses de agosto y septiembre de 2024. Las personas entrevistadas fueron, en orden alfabético: **Carolina Aguerre** (Uruguay); **Sandra Elena** (Argentina); **María Julia Giorgelli** (Argentina); **Laura Márquez-Martínez** (México); **Christian Perrone** (Brasil) y **Juan Pablo Salazar** (Colombia). Este trabajo contó con la tutoría de **Julián Casasbuenas** (Colombia). Al tiempo que el autor les agradece encarecidamente a todas las personas participantes su generosidad y predisposición para participar de esta investigación, las exime, desde ya, de cualquier omisión o error que se haga evidente a lo largo de este trabajo.

continuación, se aborda cómo los sesgos en los datos pueden llevar a decisiones injustas y discriminatorias. Se discute cómo el uso de conjuntos de datos incompletos o sesgados pueden amplificar las desigualdades existentes, con especial atención a su impacto en la administración de justicia, al tiempo que se destaca la importancia de una gestión adecuada y responsable de los datos para evitar la perpetuación de sesgos y garantizar que las herramientas de IA beneficien de manera equitativa a todas las personas involucradas en la provisión de justicia.

La **Parte III** explora las dimensiones clave en la publicación y uso de datos para la implementación de sistemas de IA responsable en la justicia. Basada en entrevistas y análisis de prácticas actuales, esta sección propone un listado de cinco principios cruciales a tener en cuenta a la hora de pensar el vínculo entre datos de justicia e IA responsable. Las dimensiones analizadas son estandarización, accesibilidad, completitud, privacidad y ciberseguridad. Finalmente, en esta sección se enfatiza la necesidad de políticas y directrices claras adaptadas a las especificidades de cada contexto institucional, sustentadas en un nivel primario en políticas de gobernanza de datos y, en un nivel secundario, en espacios institucionales y multi-actor que ayuden a dirimir el balance entre esos principios.

El trabajo finaliza con una sección de **Conclusión** y un listado de la **Bibliografía** utilizada.

Listado de definiciones (en orden alfabético)

Acceso a la justicia: Derecho de toda persona a obtener una resolución eficaz de sus conflictos basada en el derecho y a hacer valer sus derechos ante un tribunal de justicia u otra instancia de resolución de disputas. Este concepto abarca no sólo la posibilidad de acudir a una instancia judicial, sino también la capacidad de comprender, participar y obtener una resolución equitativa y efectiva, independientemente de la situación económica, social o cultural de la persona.

Aprendizaje automático: Rama de la inteligencia artificial que permite a los sistemas aprender y mejorar automáticamente a partir de la experiencia, utilizando algoritmos que identifican patrones en grandes volúmenes de datos y que permiten realizar inferencias (predicciones, recomendaciones, contenidos) a partir de nuevos datos.

Datos abiertos: Datos públicos disponibles para el uso libre e irrestricto, sin barreras técnicas ni económicas para su reutilización y redistribución.

Exclusión digital: Situación de marginalización que afecta a una persona o a un colectivo debida a la falta de acceso a los beneficios de las tecnologías digitales. Dicha exclusión puede ser consecuencia de la carencia de dispositivos adecuados o de las habilidades necesarias para su uso, así como de lagunas de representación de sus particularidades en los datos utilizados como insumo para el desarrollo de dichas tecnologías.

Gobernanza: Proceso mediante el cual se toman decisiones sobre la gestión de recursos, la elaboración de políticas y directrices o la adjudicación de responsabilidades, de manera articulada a través de la interacción entre las distintas partes o actores involucrados.

Inteligencia artificial: Tecnología que simula procesos cognitivos humanos a través de sistemas por máquina, como el aprendizaje automático, para la realización de tareas mediante inferencia, que genera resultados de salida en base a información de entrada.

Inteligencia artificial responsable: Práctica de desarrollo, implementación y uso de sistemas de inteligencia artificial orientada a asegurar el impacto social positivo, el respeto de los derechos de las personas y la minimización de los sesgos y errores, basada en la observancia de los estándares legales vigentes y principios éticos compartidos.

Justicia abierta: Práctica o visión orientada a la implementación de políticas de transparencia y rendición de cuentas, datos abiertos de justicia y participación y colaboración por parte de las instituciones que conforman el sistema de administración de justicia, con la finalidad de propiciar una justicia centrada en las personas.

Justicia centrada en las personas: Modelo que pone las necesidades y la perspectiva de las personas en el centro de la administración de justicia, principalmente mediante el diseño de procesos y servicios enfocados a incrementar su accesibilidad.

Policentrismo: Sistema en el cual operan de manera simultánea distintos puntos de poder, toma de decisiones o influencia.

Sesgo: En el marco de un sistema de inteligencia artificial, es la propensión a producir resultados que reflejan prejuicios o desigualdades, como resultado de sesgos preexistentes ya sea en los datos utilizados para su entrenamiento o en los algoritmos que lo configuran.

Transformación digital: Proceso de integración de tecnologías digitales en diversos aspectos funcionales de una institución o sector, que tienen como resultado un cambio fundamental en su forma de operar y proporcionar valor a las personas a las que sirve.

Parte I. Transformación digital, justicia abierta e inteligencia artificial

Transformación digital centrada en las personas

La transformación digital está redefiniendo las instituciones. Este proceso, que se da en un contexto de creciente omnipresencia de las tecnologías digitales, da lugar a disrupciones (entendidas como un quiebre o rompimiento con la manera tradicional de ejecutar algo), a las que las instituciones responden implementando cambios en los modos mediante los cuales sus intervenciones generan valor social (Vial, 2019).

Por otra parte, es un error considerar la transformación digital como un fenómeno sólo ligado a la aplicación de tecnología a procesos preexistentes para acelerarlos o hacerlos más eficientes, sin revisar su finalidad o su sentido. En contraste, la transformación digital debe enfocarse en el uso de tecnologías con vistas a ampliar los niveles de eficiencia, accesibilidad y transparencia de las instituciones. Esto requiere poner en línea varios factores clave, como la introducción de cambios culturales y organizativos que lleven a mejoras significativas en cómo se brindan los servicios a las personas, cómo se optimizan las operaciones o qué servicios completamente nuevos pueden implementarse (Brown et al., 2014; PNUD, 2024).

Para ser efectiva e inclusiva, la transformación digital debe apuntar a incrementar la confianza de las personas en las instituciones. Un modo de lograr esto es orientar dicha transformación hacia un enfoque de transformación digital centrada en las personas, que tome para sí las diversas realidades sociales y ponga el énfasis en facilitar un acceso equitativo y adecuado a los servicios y bienes públicos (Santiso, 2022). En este contexto, resulta clave asegurar ese acceso en particular a las personas y comunidades en situación desaventajada, dado que son quienes sufren en mayor medida de exclusión digital y de una continua ampliación de las brechas digitales existentes.

La transformación digital puede tener efectos directos y duraderos sobre la vida de las personas, por lo que revertir la subrepresentación en los datos que sirven como insumo para informar y diseñar dichas políticas es una manera de abordar las brechas digitales. Este es un aspecto que debe estar en el centro de los procesos de transformación digital. Por otra parte, mientras que la transformación digital puede, por una parte, facilitar el ejercicio de derechos como la libertad de expresión o el acceso a algunos bienes públicos, esta puede asimismo tener efectos negativos que sobre los vínculos y tensiones preexistentes entre otros derechos, como por ejemplo el derecho a la privacidad. El enfoque de transformación digital centrada en las personas sirve para mitigar esos posibles efectos negativos, dado que parte de una comprensión cabal de cómo los derechos de las personas pueden verse afectados como consecuencia de ese mismo proceso. En el marco de dicha comprensión, la perspectiva de las personas debe integrarse en todas las etapas del ciclo de vida de la transformación digital, lo que incluye tanto la toma de decisiones para su diseño, como su posterior evaluación y monitoreo (OCDE, 2024a).

En el caso particular de la administración de justicia, nos encontramos ante un servicio que viene experimentando distintos procesos de transformación digital pero que, en sus aspectos más fundamentales, ha experimentado relativamente muy pocos cambios a lo largo de los últimos doscientos años (Muller et al., 2013). En este sentido, la implementación de políticas de transformación digital en la justicia no debería

centrarse únicamente en implementar herramientas digitales, como el expediente electrónico o las audiencias virtuales. Requiere, por el contrario, un cambio cultural y organizativo profundo hacia el interior de las instituciones. Ignorar la necesidad de dicho cambio podría dar lugar a un *tecno-solucionismo* peligroso, que lleve a intentar resolver problemas complejos únicamente mediante soluciones tecnológicas, sin considerar su verdadera naturaleza multidimensional y, por ende, la necesidad de abordarlos de manera integral y centrada en las personas. Este tecno-solucionismo podría, de hecho, dar como resultado el fracaso de los proyectos de transformación digital emprendidos, la pérdida de autonomía de las personas a las que se buscaba beneficiar o, incluso, la perpetuación y el agravamiento de los niveles de exclusión digital preexistentes (Addo et al., 2021).

El proceso de transformación digital para la administración de justicia debe orientarse hacia la construcción de una justicia centrada en las personas (De Langen et al., 2022; OCDE, 2023). Ello supone una inclusión significativa de las personas y comunidades usuarias, tanto actuales como potenciales, en las distintas fases de su implementación, en particular aquellas en situación de vulnerabilidad, que son quienes al mismo tiempo sufren de manera más acentuada la falta de acceso a la justicia (Pathfinders for Peaceful, Just and Inclusive Societies, 2019). De manera concomitante, la transformación digital en la administración de justicia debe comprender tanto el nivel micro, esto es, de cada institución o juzgado, como el nivel sistémico, es decir, el de toda la cadena de provisión de justicia, a modo de incluir a todas las instituciones encargadas de proveer de manera articulada seguridad y justicia (poderes judiciales, juzgados, fiscalías, ministerios públicos, fuerzas policiales y de seguridad, ministerios de justicia o instituciones penitenciarias). Estas instituciones se caracterizan tanto por su papel en el cumplimiento del Estado de derecho como por la necesidad de un flujo continuo y colaborativo de datos e información entre ellas (Seepma et al., 2021).

Justicia Abierta

La creciente adopción de la visión de Justicia Abierta por parte de las instituciones del sector justicia resulta un factor decisivo en el contexto de estas necesidades y puede posibilitar la implementación exitosa de una transformación digital en el sector justicia enfocada en la creación colectiva de valor público para la justicia (Jiménez-Gómez, 2017; Elena et al., 2018; Hess et al., 2019; Sharp Vargas, 2020). En la práctica, la Justicia Abierta consiste en que las instituciones que forman parte de la cadena de administración de justicia adopten políticas centradas en las personas que pueden agruparse en tres grandes conjuntos.

Por un lado, las políticas de transparencia y rendición de cuentas proporcionan a las personas un acceso claro y comprensible al funcionamiento del sistema de justicia. Estas políticas, que pueden ser activas (es decir, la divulgación proactiva de información por parte de las instituciones) o pasivas (acceso a la información previo requerimiento), también ayudan a que las instituciones de justicia sean más responsables en su desempeño, haciendo que su funcionamiento sea más predecible y trazable.

En segundo lugar, la Justicia Abierta impulsa la publicación por parte de las instituciones de justicia de datos sobre su funcionamiento, en particular sobre sentencias y resoluciones, en formatos abiertos y reutilizables para su libre uso. Estos datos pueden ser libremente reutilizados en distintos esquemas de

creación de valor, que comprenden tanto esquemas informativos, como de rediseño de procesos basado en evidencia, o de innovación tecnológica basada en datos.

En tercer lugar, la Justicia Abierta promueve la creación de instancias de participación ciudadana y colaboración. A través de estos espacios, las personas usuarias del sistema pueden contribuir a la creación y revisión de políticas públicas de justicia desde una perspectiva centrada en las personas. Al mismo tiempo, las instituciones judiciales, junto con otros actores locales y supralocales, colaboran entre sí en la gestión de temas transversales.

En lo que respecta a la gestión de los datos, la Justicia Abierta impulsa desde hace más de una década la publicación de datos del sistema de justicia, en formatos preferentemente abiertos y reutilizables (Elena et al., 2014; Marković et al., 2020; Mercado, 2021). Estos datos consisten fundamentalmente en resoluciones y sentencias judiciales (documentos en formato de texto), en datos jurisdiccionales (que reflejan la labor jurisdiccional, como por ejemplo indicadores de actividad judicial) y datos estructurales (que reflejan aspectos financieros, edificios o de gestión de recursos humanos de las instituciones).

La existencia de un ecosistema de datos de justicia robusto es fundamental para que las instituciones y actores involucrados en la administración de justicia puedan resolver los problemas de justicia de las personas mediante la generación de procedimientos más efectivos y centrados en sus necesidades legales. Entretanto, esta disponibilidad de datos de justicia se apoya a su vez en dos factores interdependientes. Por una parte, requiere de una comunidad de actores que utilice y demande datos para informar y mejorar los servicios de justicia basándose en evidencia y desde una perspectiva centrada en las personas. Por otra parte, precisa de un suministro oportuno de datos relevantes y de calidad sobre los resultados de los procesos judiciales, que no sólo respondan a esta demanda específica, sino que también se encuentren estandarizados y disponibles para los diversos contextos de creación de valor (World Justice Project, 2023).

Inteligencia Artificial en la Justicia

El advenimiento de la IA en el marco de la transformación digital, en particular en el contexto del sistema de justicia, ha creado en muy poco tiempo un escenario completamente distinto al de años anteriores. Para definir qué es la IA puede tomarse la definición reciente incluida en el *Reglamento de Inteligencia Artificial de la Unión Europea*, que habla de un sistema de IA como uno “...basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales...” [Reglamento (UE) 2024/1689, Art. 3].

Por otra parte, de entre la plétora de definiciones existentes sobre qué es la IA, la de Rocha y Carvalho (2022) la entiende como un conjunto de técnicas cuyo objetivo es reproducir, mediante una máquina, las capacidades cognitivas de los seres humanos, basadas en buena medida en aplicaciones de aprendizaje automático (*machine learning*). Estas técnicas en particular requieren de una ingente cantidad de datos para ser entrenadas con la finalidad de poder inferir, a partir de las entradas que recibe, resultados tales como predicciones, contenidos, recomendaciones o decisiones.

Para esta definición, Rocha y Carvalho utilizan como punto de partida la premisa de que el despliegue de la IA, en particular en el contexto de la administración de justicia, puede mejorar de manera global los niveles actuales de acceso a la justicia, en un contexto de aumento constante de sus costes, tanto en términos del tiempo necesario para resolver litigios como de la carga de trabajo del personal judicial. Por otra parte, los autores destacan que una buena parte de los litigios son, en la práctica, simples, similares (o incluso idénticos), repetitivos y con resultados predecibles. En consecuencia, el uso de IA para automatizar procesos manuales puede, en estos casos, agilizar las decisiones y reducir de manera significativa los tiempos procesales y, así, la cantidad de casos pendientes (*backlog*). Según los autores, el uso de IA permitiría que el personal judicial (en particular las juezas y los jueces) dediquen más tiempo a su función jurisdiccional principal de decidir los casos judiciales, al mejorar la asignación de recursos como consecuencia de la agilización de tareas mecánicas y repetitivas.

En un trabajo pionero en el cual se exploró el nivel de preparación de las instituciones judiciales en cuatro países de América Latina para la adopción de estas tecnologías, Aguerre y Bustos Frati (2021) describen tres grandes niveles para el uso de IA en la justicia. En un nivel más básico, se encuentra la asistencia mediante el informe, el apoyo y el asesoramiento a las personas involucradas en el sistema de justicia (definido como *tecnologías de apoyo*). En un nivel intermedio, las autoras definen la sustitución de las personas humanas en la realización de algunas funciones y actividades (*tecnologías de sustitución*). Por último, en un tercer nivel, se encuentra un cambio más fundamental en la forma de trabajar del sistema de justicia, o bien el establecimiento de formas diferentes de justicia, en particular a través de cambios en los procesos (*tecnologías disruptivas*).

Por otra parte, definir hoy los usos de la IA en la justicia resulta especialmente complejo como producto de la rápida evolución de las plataformas de IA generativa y su capacidad de trabajar con información, transformarla, gestionarla y generar nuevo texto, lo que ha vuelto su uso más ubicuo y transversal a todo el ámbito judicial. Más allá de la naturaleza intrínsecamente dinámica e impredecible de la IA, existen algunos usos en el campo de la administración de justicia que se encuentran más consolidados que otros, lo que permite clasificar su aplicación en tres categorías principales (Reiling, 2022). En primer lugar, se encuentran los sistemas de organización de la información, como el reconocimiento automatizado de patrones y términos clave, que pueden analizar grandes volúmenes de documentos legales y jurisprudencia y ofrecer argumentaciones complejas y fundamentadas (Socol de la Osa et al., 2024). En segundo lugar, se encuentran los sistemas de apoyo para la toma de decisiones, que, basados en el aprendizaje automatizado, pueden condensar argumentos legales complejos, evaluar la relevancia y solidez de las pruebas presentadas, e incluso sugerir posibles resultados (Wang et al., 2022). Por último, los sistemas de predicción ya permiten identificar veredictos en textos legales mediante la búsqueda de palabras clave, categorizar decisiones según sus resultados y establecer predicciones basadas en jurisprudencia relevante (Medvedeva et al., 2023), mientras que algunos enfocados en la evaluación de riesgos se basan en modelos estadísticos para intentar prever comportamientos futuros en el marco de un proceso (Webster, 2024).

En cuanto al uso de herramientas de IA en la justicia, existen ya datos empíricos a nivel global (UNESCO, 2024) que señalan un nivel de familiaridad superior a 9 de cada 10 personas operadoras judiciales. En términos de uso efectivo en la labor judicial, la evidencia señala que un 45% de estas personas utilizan de manera cotidiana herramientas de IA. Los principales usos incluyen la investigación (59%), la redacción

de documentos (45%), la redacción de correos electrónicos (38%) y la comprensión de nuevos conceptos legales (25%). Entretanto, según esta misma evidencia, existe un relativo equilibrio entre las preocupaciones éticas en torno al uso de la IA (en particular la IA generativa) en la profesión legal y una actitud abierta hacia esta tecnología y sus posibles beneficios.

Este alto nivel de familiaridad conlleva el desafío de asegurar que el desarrollo y la implementación de la IA se realicen de manera responsable, esto es, teniendo en cuenta su impacto social y en línea con los valores de las partes involucradas, los estándares legales vigentes y los principios éticos compartidos (Stryker, 2024). En este sentido, si bien no existe una definición universalmente aceptada de IA responsable, existe un consenso en torno a la idea de que la implementación de sistemas de IA en el contexto de la transformación digital de la justicia puede tener efectos tanto positivos como negativos, por lo que es preciso utilizarla de manera controlada, teniendo en cuenta su impacto tanto a nivel institucional como social. Todo esto implica reconocer en el proceso de desarrollo de estos sistemas la perspectiva de las personas a las que la IA debe servir, como manera de buscar un impacto positivo a partir de su utilización (Adams, 2024).

Establecer los mecanismos necesarios para asegurar un uso responsable de la IA no está exento de diversas complejidades y obstáculos, como lo evidencia el enorme corpus de evaluaciones, índices y documentos de políticas elaborados en el marco de la explosión del uso de IA en la escena global. Un ejemplo es el *Global Index on Responsible AI* (Adams et al., 2024), que explora la existencia de prácticas de IA responsable en 138 países. Entre sus conclusiones generales, la primera edición de este estudio establece una distinción taxativa entre la existencia de una gobernanza de la IA y una IA efectivamente responsable. En este sentido, del estudio se desprende que, si bien para un uso responsable de la IA es necesario contar con un marco de gobernanza de la IA, la mera existencia de dicho marco no garantiza que el uso de la IA sea considerado como responsable, dado que ello requiere adoptar provisiones específicas para la protección de los derechos de las personas en su implementación.

Esta conclusión se encuentra asimismo en línea con el valor que le asignan Belli y Gaspar (2024) a valores como la transparencia y la rendición de cuentas para el establecimiento de políticas de gobernanza de la IA. Estos valores son considerados por los autores como la clave para orientar los desarrollos futuros hacia el respeto de los derechos de las personas, el cumplimiento de los requisitos legales existentes y la promoción de la sostenibilidad y, por ende, para el desarrollo responsable de la IA.

Por otra parte, los resultados del *Global Index on Responsible AI* levantan alertas sobre lo que se considera una ausencia generalizada de preparación en los países evaluados para desarrollar sistemas de IA inclusivos y equitativos. En este contexto, aspectos como la cooperación entre las partes interesadas y las distintas comunidades de práctica, así como la incorporación del punto de vista de la sociedad civil, son considerados como factores clave para impulsar la IA responsable. Por otra parte, otro análisis como el aportado por el *Human-Centered Artificial Intelligence (HAI) Index Report* de 2024 (Maslej et al., 2024) hace hincapié en la creciente dependencia de la disponibilidad de datos para los modelos de IA, al tiempo que alerta sobre la escasez de datos de calidad disponibles para el desarrollo futuro de la IA.

Todos estos aspectos llevan a revisar algunos de los aspectos fundamentales de las políticas de Justicia Abierta, tales como la búsqueda de mayor disponibilidad y calidad de los datos judiciales, o la colaboración

entre diversas partes interesadas, los cuales toman un nuevo cariz a la luz de la necesidad de desarrollar sistemas de IA en la justicia que sean responsables y centrados en las necesidades de las personas. Así, mientras que los datos abiertos de justicia han apoyado hasta el momento principalmente esquemas de creación de valor enfocados en mejorar la comprensión sobre el funcionamiento de las instituciones y generar conocimiento basado en evidencia sobre el sistema legal (Albrecht et al., 2022), usos que, por otra parte han dado lugar a la creación de estándares específicos para que estos permitan reflejar el vínculo con la justicia de poblaciones en situación de vulnerabilidad (National Center for State Courts, 2020; Hivos, 2023), la relevancia prevista para estos mismos datos como insumo para el desarrollo de sistemas de IA responsable se prevé en aumento, motivo por el cual resulta pertinente replantear los principios que deberían guían su publicación en base a las nuevas necesidades detectadas.

Parte II. Datos, sesgos y gobernanza de la IA de justicia

Datos e IA de justicia

El rol fundamental de los datos para el desarrollo la IA es reconocido en prácticamente todos los instrumentos institucionales de políticas y documentos estratégicos que abordan o buscan regular la implementación de este tipo de herramientas (entre muchos otros, Cumbre Ministerial de Altas Autoridades de América Latina y el Caribe, 2023; UK Judiciary, 2023; OCDE, 2024b; Canada Federal Court, 2024). En particular, estos dan cuenta de la necesidad de considerar la calidad, integridad y transparencia de estos como aspectos clave que posibilitan un uso responsable, que sirva para proteger de los derechos de las personas involucradas en procesos judiciales.

En el ámbito particular de la administración de justicia, la Comisión Europea para la Eficiencia de la Justicia (CEPEJ), órgano dependiente del Consejo de Europa creado para mejorar la eficiencia y el funcionamiento de los sistemas de justicia en sus Estados miembro, ya había presentado en el año 2018 una serie de principios éticos³ para un buen uso de la IA en la administración de justicia a través de la *Carta Ética Europea sobre el Uso de la Inteligencia Artificial en los Sistemas Judiciales y su Entorno* (CEPEJ, 2018).

Este documento, que sirvió oportunamente para sentar las bases del debate sobre el futuro uso de IA en la justicia (en un momento en el cual esta era una temática todavía incipiente) deja entrever la importancia que desde ese momento ya se le asignaba a los datos, en particular a través del principio de Calidad y Seguridad (principio número 3), mediante el cual se establece la necesidad de que los datos basados en decisiones judiciales que se ingresen en las herramientas aprendizaje automático provengan de fuentes certificadas y se mantengan inalterados hasta su utilización. Por otra parte, este principio establece la necesidad de trazabilidad del proceso, como manera de garantizar que no se produzcan modificaciones que puedan alterar el contenido o el significado de la decisión judicial procesada.

³ Los principios éticos planteados por la CEPEJ en la Carta son: respeto por los derechos fundamentales; no discriminación; calidad y seguridad; transparencia, imparcialidad y equidad; y uso bajo control del usuario.

En esta Carta Ética, la CEPEJ explora asimismo las políticas existentes en materia de publicación de datos abiertos por parte de los distintos sistemas de administración de justicia, especialmente en lo que compete a la publicación, de manera abierta y reutilizable, de fallos y sentencias. Dos cuestiones se establecen como premisas en este estudio, cuya continuidad será analizada en lo que resta de este trabajo por tratarse de cuestiones en torno a las cuales se dirime el futuro de la implementación de la IA en la justicia. Por una parte, se establece la disponibilidad de datos como un prerequisite esencial para el desarrollo de la IA, así como su disponibilidad en una modalidad abierta como condición fundamental para su reutilización en el contexto de desarrollo de herramientas de IA. En segundo lugar, se establece la cuestión de la privacidad y la protección de datos como uno de los aspectos a tener en cuenta en el marco del procesamiento de esta información por parte de sistemas de IA basados en aprendizaje automático.

La entrada en vigor, en el año 2024, del *Reglamento Europeo de Inteligencia Artificial* significa un paso significativo de cara a generar en el corto plazo un marco normativo y jurídico único y específico para regular los sistemas de IA que operen en la Unión Europea (UE), con independencia de si están establecidos en la UE o en un tercer país. Por otra parte, el Reglamento establece una clasificación en base a los distintos niveles de riesgo, en base a la cual se crea la categoría de sistemas de IA de *alto riesgo*, cuyos proveedores deben por lo tanto someterse a obligaciones específicas de transparencia, documentación, supervisión humana, gestión de calidad, entre otras prácticas.

Los sistemas de alto riesgo, definidos según este Reglamento, son aquellos que forman parte de actividades tales como la biometría, la administración de infraestructuras críticas, la educación (en particular, los sistemas de admisión en centros educativos, o la evaluación de calificaciones, aprendizajes o conductas), el empleo y la gestión de las personas trabajadoras (sistemas de selección de personal y toma de decisiones contractuales), el acceso a servicios y prestaciones esenciales, tanto públicos como privados (acceso a asistencia pública, calificación crediticia), o la migración, el asilo y la gestión del control fronterizo (evaluación de solicitudes de asilo, visados o permisos de residencia, o la entrada en un país). Entretanto, los sistemas de garantía de cumplimiento del derecho y la administración de justicia también figuran como sistemas que son considerados de alto riesgo en el Reglamento, motivo por el cual se encuentran sujetos a obligaciones particulares que buscan regular y supervisar su uso [Reglamento (UE) 2024/1689, Anexo III].

En los considerandos del Reglamento se explicita la necesidad de regular la utilización de la IA en la justicia. En particular, se señalan los efectos potencialmente relevantes que estas herramientas pueden tener para la democracia, el cumplimiento del Estado de Derecho, las libertades individuales y el derecho a la tutela judicial efectiva y a un juez imparcial. Por otra parte, se subraya la importancia de la regulación para poder hacer frente a posibles sesgos, errores y opacidades de estos sistemas, en particular en los casos en los que son utilizados por una autoridad judicial para investigar e interpretar los hechos y el Derecho en el marco de la aplicación de la ley, al tiempo que se explicita que, mientras que las herramientas de IA pueden apoyar la toma de decisiones, esta sigue siendo en última instancia una responsabilidad humana [Reglamento (UE) 2024/1689, considerandos puntos 61 y ss.].

El Reglamento señala asimismo la importancia de los datos de alta calidad y su acceso como forma de proporcionar una estructura y garantizar el funcionamiento de muchos sistemas de IA, con vistas a garantizar que los sistemas funcionen del modo previsto y en condiciones de seguridad, evitando convertirse en una fuente de discriminación. Para ello, se insta a establecer prácticas adecuadas de gestión y gobernanza

de los datos utilizados en el marco de los sistemas de IA con el fin de lograr que los conjuntos de datos para el entrenamiento, la validación y la prueba sean de alta calidad. En este sentido, se llama la atención sobre los sesgos, que se consideran inherentes a los conjuntos de datos utilizados y que, como tales, pueden afectar a los resultados de los sistemas de IA, perpetuando y amplificando así la discriminación existente contra determinados colectivos vulnerables. Para ello, se establece como requisito que los conjuntos de datos, en la mayor medida posible, sean completos y libres de errores, lo cual no debe afectar al uso de técnicas de protección de la intimidad de las personas, si bien se deben tener en cuenta las características, rasgos o elementos particulares del entorno geográfico, contextual, conductual o funcional específico en el que esté previsto que se utilice el sistema de IA [Reglamento (UE) 2024/1689, considerandos punto 67].

Este punto del Reglamento pone el foco en una de las cuestiones centrales en lo que respecta a las necesidades de datos surgidas en el marco del desarrollo de sistemas de IA, en particular en el ámbito de la justicia. Esta consiste en que, al tiempo que se prevé un uso intensivo de estos sistemas en la labor judicial, su implementación depende en buena medida de contar con acceso a grandes cantidades de datos para el entrenamiento de estos sistemas, algo que a menudo no se cumple (Campbell, 2020). En este sentido, si bien parece existir una creciente conciencia sobre la importancia de aumentar la disponibilidad de datos para mejorar el acceso a la justicia, revelar problemas sistémicos, impulsar reformas judiciales y empoderar a quienes no tienen representación legal (Rostain, 2024), falta todavía un largo camino para entender y valorar el potencial de los datos judiciales y la importancia de su disponibilidad. Particularmente en América Latina, la visión de Justicia Abierta ha intentado mejorar esta situación en los últimos años, con algunos casos de éxito pero sin lograr saldar del todo la deuda existente en la materia.

Entretanto, para que las herramientas de IA funcionen de manera efectiva, se requieren tres factores críticos, conocidos como *las tres v*. Estos son, un gran volumen de datos (una cantidad significativa de información), una gran variedad de datos (distintas fuentes que amplíen las capacidades de la IA y reduzcan posibles sesgos) y un alto nivel de veracidad (datos provenientes de fuentes confiables), precondiciones que pueden ser satisfechas a través de su disponibilización mediante estrategias de apertura de datos a través de portales y repositorios institucionales (Comisión Europea, 2024).

Por otra parte, un documento reciente de la CEPEJ, publicado en el contexto de la irrupción global de la IA generativa (CEPEJ, 2024), da cuenta de una coherencia con aquel documento publicado en el año 2018 mencionado más arriba, al señalar que cualquier herramienta de IA se encuentra esencialmente determinada por los datos con los que ha sido entrenada, por lo que los vacíos u omisiones (y los sesgos) presentes en las bases de datos utilizadas para su entrenamiento tienen una clara injerencia sobre la validez de los resultados que genera. Por lo tanto, es necesario explorar el porqué de la importancia de los datos en materia de la generación de sesgos que limitan la validez de los resultados de la implementación de la IA.

Sesgos basados en datos

Entendidos de manera general como aquellos casos en los cuales las herramientas de IA toman decisiones que, de manera sistemática, dejan en desventaja a ciertas poblaciones o grupos de personas (Friis et al., 2023), los sesgos son fallas sistemáticas que ocurren en un proceso de toma de decisiones y llevan a resultados injustos. En el contexto del desarrollo de la IA, el sesgo puede ser resultado de diversas causas

(entre ellas, una recolección defectuosa de los datos), a raíz de las cuales un modelo de aprendizaje automático puede aprender y replicar patrones de sesgo presentes en los datos utilizados para entrenarlo, lo que puede derivar en resultados injustos o discriminatorios (Ferrara, 2023).

La cuestión de los sesgos basados en datos en el marco de la IA se debatió en la sesión de junio y julio de 2024 del Consejo de Derechos Humanos de Naciones Unidas, instancia en la cual se presentó un informe de relatoría que hace hincapié en el uso de IA en la administración de justicia (Consejo de Derechos Humanos, 2024). En este informe se destaca, por una parte, la digitalización de datos a gran escala y el uso que los sistemas de IA hacen de esos datos para tomar decisiones y llevar a cabo acciones en diversos sectores. Como contrapartida, se explica que los conjuntos de datos con los que se entrenan los algoritmos a menudo son incompletos o pueden sub-representar (o sobre-representar) a ciertos grupos de personas, lo cual puede tener como consecuencia el surgimiento de un sesgo algorítmico, al igual que aquellos casos en los que los conjuntos de entrenamiento incluyen datos ya sesgados. De esta manera, en aquellos casos en los cuales los datos de entrenamiento son insuficientes, los algoritmos pueden hacer predicciones que resultan sistemáticamente discriminatorias para los grupos que están ausentes o sub-representados en los datos. Por otra parte, el sesgo algorítmico puede ocurrir con algoritmos basados en datos no representativos, en los cuales la sobre-representación de un grupo de personas también puede producir resultados sesgados⁴.

Así, si los datos en los que se basan los modelos están sesgados contra ciertos grupos, los modelos informáticos tienden a reproducir y amplificar esos sesgos, por lo que el uso de esos datos sesgados o defectuosos para tomar decisiones en la vida real puede seguir perjudicando a los grupos en situación de vulnerabilidad, replicando y profundizando las disparidades existentes. Así, según el informe, en lo que respecta a los datos incompletos o no representativos, si los datos utilizados para entrenar la IA no son lo suficientemente diversos o inclusivos, el algoritmo puede no representar ni predecir con precisión los resultados para los grupos subrepresentados. Por otra parte, si los datos de entrenamiento reflejan sesgos históricos o sociales, la IA probablemente los perpetúe e incluso amplíe.

En particular en el caso de la administración de justicia, el informe hace referencia a las herramientas de evaluación de la posibilidad de reincidencia, que son utilizadas para la toma de decisiones en diferentes etapas del sistema de justicia penal. Se trata de herramientas de IA que emplean datos históricos para evaluar la probabilidad de que las personas procesadas actúen de ciertas maneras, especialmente de la probabilidad de que cometan un nuevo delito en el futuro. Estas herramientas generan puntuaciones de riesgo utilizando información procedente de registros criminales y encuestas a las personas procesadas. En este sentido, los datos del sistema de justicia penal utilizados para entrenar estos algoritmos reflejan desigualdades sistémicas, a menudo basadas en un historial de conductas policiales racistas⁵.

Teniendo en cuenta lo mencionado, es posible identificar algunas necesidades clave en términos de disponibilidad y publicación de datos de justicia. Para empezar, es esencial comprender que el proceso de

⁴ En el documento se hace referencia a un caso de sesgo algorítmico derivado del uso de bases de datos de imágenes para el reconocimiento facial en los Estados Unidos, en las cuales una sobre-representación de personas de ascendencia africana, en última instancia reflejo de patrones de racismo sistémico, hizo que estas personas fueran más propensas a ser identificadas de forma errónea.

⁵ El caso más documentado y estudiado en este sentido es el de la herramienta *Correctional Offender Management Profiling for Alternative Sanctions* (COMPAS) en los Estados Unidos, *vid.* Fuchs, 2018, o Lagioia et al., 2023.

generación, clasificación y uso de los datos de justicia es fundamental. Este proceso implica la necesidad de que la producción de datos y el diseño de los algoritmos se realicen con responsabilidad e inclusividad. Esto es especialmente crucial en contextos como el de América Latina, donde una parte significativa de la población es invisibilizada tanto por el Estado como por el sistema de justicia, lo que perpetúa los sesgos presentes en nuestras sociedades (ILDA, 2020).

Parte III. Dimensiones clave en la publicación y uso de datos para la IA de justicia

Principios para la publicación de datos de justicia

En base a las entrevistas realizadas y al relevamiento de experiencias en materia de publicación y uso de datos para la implementación de sistemas de IA responsable⁶, se propone una serie de consideraciones que las instituciones de justicia deben tener en cuenta para la publicación de sus datos. Estas consideraciones buscan informar un marco a través del cual pueda orientarse el trabajo futuro en materia de gobernanza de los datos y consecuente publicación. En tanto tales, se considera que estos deben ser tenidos en cuenta en base a la particularidad de cada contexto, a la vez que buscan brindar un enfoque general, basado en los desafíos y necesidades detectados en la práctica, recogidos durante la investigación para este trabajo.

Estas dimensiones en juego son: estandarización, accesibilidad, completitud, privacidad y ciberseguridad. Se trata de cinco grandes dimensiones clave, que se encuentran en buena medida interrelacionadas, respecto de las cuales las instituciones de justicia encargadas de publicar datos deben establecer o adoptar una postura basada en sus necesidades y posibilidades institucionales, su marco legal específico y, en definitiva, los aspectos políticos que conforman los procesos de toma de decisiones específicos. Buscan brindar, así, un marco general a ser tenido en cuenta a fin de que los datos publicados puedan eventualmente formar parte de procesos de implementación de IA responsable.

Antes de pasar a la definición de cada uno de estos principios, se recoge como buena práctica el uso de espacios controlados (*sandboxes*) para el desarrollo, experimentación y puesta a punto de los sistemas de IA, previa utilización generalizada. Estos espacios controlados permiten el uso de nuevos sistemas eliminando y administrando el riesgo que conlleva su uso en la práctica real. Estos espacios controlados son propuestos en el *Reglamento Europeo de IA* [Reglamento (UE) 2024/1689, considerandos, punto 139]

⁶ En particular se destaca la experiencia de *IA LucIA*, una herramienta de IA desarrollada por ÌO Justice, organización mexicana de desarrollo de herramientas de tecnología legal con foco en la apertura y el acceso a la justicia. La herramienta fue desarrollada a partir de un ejercicio de innovación abierta con el poder judicial federal en México en el año 2021, a raíz de la necesidad de contar con una herramienta que permitiese encontrar precedentes de manera más eficiente y evaluar sistemáticamente la coherencia en las decisiones judiciales. Para ello, utiliza IA generativa para proporcionar respuestas basadas en la estadística de decisiones judiciales previas, con el objetivo de ofrecer una respuesta precisa y alineada con la normativa jerárquica. Notablemente, *IA LucIA* es entrenada con decisiones judiciales que pasan por un proceso previo de curaduría para mitigar los sesgos en los resultados. Vid. <https://ialucia.com/>

como una manera de impulsar la innovación, estableciendo al mismo tiempo un entorno de experimentación y prueba controlado para ser utilizado en la fase de desarrollo, previa a su aplicación.

Los espacios controlados ofrecen un enorme potencial para mejorar la gobernanza de datos en la IA. Dado que permiten a los responsables de políticas evaluar la efectividad de las regulaciones propuestas, identificar posibles lagunas y refinar iterativamente las políticas para asegurar que se alineen con los estándares éticos, los requisitos legales y los marcos de gobernanza de datos (Datasphere Initiative, 2024). En el caso de Brasil, por ejemplo, la Autoridad Nacional de Protección de Datos (ANPD), autoridad responsable de velar por la protección de los datos personales, cuenta con el primer caso documentado de utilización de un espacio controlado para el desarrollo de IA (CAF, 2024).

Para la publicación de datos en el marco de prácticas de IA de justicia responsable, una primera dimensión a tener en cuenta es la **estandarización** de los datos. En la actualidad, las instituciones judiciales que publican sus datos tienden a hacerlo de manera aislada, sin un marco común que asegure la uniformidad en la publicación. Esta situación resulta problemática, dado que, al no existir un estándar unificado para todos los tribunales, se pierde la oportunidad de garantizar que los datos se presenten de manera coherente y que cumplan con criterios de calidad uniformes (Hivos, 2023).

La falta de aplicación de estándares comunes para la publicación de los datos judiciales no solo crea disparidades significativas entre las diferentes instituciones, sino que también dificulta la comparación y el análisis transversal de la información. En última instancia, esto no solo limita el potencial de innovación y generación de valor público que trae aparejada la reutilización de los datos publicados por la administración de justicia, sino que también pone en riesgo la calidad y la fiabilidad de los resultados derivados del uso de esos datos. En particular en materia de IA, la ausencia de estandarización dificulta el uso efectivo de los datos, dado que dificulta su integración y reutilización para el entrenamiento de sistemas que puedan aplicarse de manera efectiva en la administración de justicia. Por otra parte, un aspecto a tener en cuenta para el establecimiento de estos estándares compartidos para la publicación de datos es que se trata en buena medida de un problema de acción colectiva, en el cual se requiere una cooperación que a menudo es compleja de alcanzar por cuestiones de intereses institucionales en conflicto que desalientan dicha cooperación.

Existen en América Latina algunos casos en los cuales se avanzó en torno a la estandarización de los datos, entre los cuales pueden citarse el caso de Argentina, donde el Ministerio de Justicia estableció oportunamente, de manera colaborativa con los subsistemas de justicia provinciales, un protocolo técnico de datos y de procesos para guiar la generación de datos en base a criterios y estándares comunes que los hicieran interoperables y comparables (Programa Justicia Abierta, 2022), o el caso de Brasil, donde el Conselho Nacional de Justiça trabajó en la creación de estándares comunes para la publicación de casos y de documentos, en particular con vistas a su reutilización en sistemas de IA (Conselho Nacional de Justiça, Resolução N° 332 de 21/08/2020)⁷. La utilización de formularios únicos de curaduría para sentencias, que incluyan extractos de hechos, decisiones y pruebas, se valora en este contexto como un posible camino para comenzar a subsanar esta ausencia de estandarización, que requerirá de desarrollo posterior en los términos particulares de cada sistema de justicia.

⁷ Cabe destacar que en la actualidad se encuentra en proceso de elaboración una nueva normativa actualizada para establecer nuevos estándares. Dicha normativa se encuentra aún en etapa de debate hacia la finalización de este trabajo.

En segundo lugar, se encuentra el principio de la **accesibilidad** de los datos. Por principio, los datos judiciales deben ser fácilmente accesibles para posterior su utilización en sistemas de IA. En este sentido, se identifica a la *convergencia*, entendida como la disponibilidad de datos provenientes de varias fuentes de manera integrada y unificada, como un aspecto fundamental a tener en cuenta para facilitar dicha accesibilidad. Por otra parte, si bien existen argumentos a favor de la adopción de distintas metodologías para la disponibilización de los datos, como el fideicomiso de datos (CAF, 2024) o los recursos comunes de datos (*data commons*)⁸, las entrevistas recogen la modalidad abierta como un modo de brindar accesibilidad cuyas ventajas no deberían ser soslayadas.

Así, la disponibilidad de los datos en formato abierto ofrece varias ventajas, tales como brindar accesibilidad a los datos de manera directa, mediante repositorios o portales públicos y abiertos, lo que permite su descarga y reutilización casi inmediata y elimina barreras de acceso. En segundo lugar, las fuentes en el caso de los datos abiertos son verificables y ayudan a triangular distintos hallazgos, dado que representan un recurso compartido que refleja el mismo conjunto de hechos públicos. En tercer lugar, los datos abiertos son explicables y rastreables. Esto significa que quienes los utilizan pueden explicar y rastrear su origen. De hecho, se espera que los sistemas de IA construidos en base a estos datos mejoren a medida que los datos mismos también mejoren, lo cual es crucial en un contexto en el que cada vez se depende más de datos de entrenamiento sintetizados (Jacobsen, 2023). Por último, otra ventaja identificada es que existen instituciones gubernamentales responsables de su calidad, lo que contribuye a que las solicitudes para corregir datos reciban una respuesta adecuada, garantizando así la responsabilidad en su publicación (Carfi et al., 2024).

En América Latina, la accesibilidad de los datos judiciales es una cuestión incipiente, que se encuentra sujeta a cuestiones tales como la ausencia de consciencia sobre el valor de los datos, la carencia de políticas claras y sostenidas de gobernanza de datos, y resistencias institucionales propias de la justicia como institución en general poco proclive a la apertura. En todo caso, este es un aspecto en el cual existen avances (Red Internacional de Justicia Abierta, 2021) que se espera que sigan teniendo lugar en lo próximo, en particular a medida que el uso de IA y las necesidades de datos sigan siendo documentadas y debatidas.

Un tercer aspecto a tener en cuenta es el de la **completitud** de los datos. En los sistemas de justicia, al menos en América Latina, son escasos los datos que existen sobre las personas detrás de las sentencias, lo que representa un problema en términos de visibilidad de la diversidad en los datos judiciales. La subrepresentación, plasmada en los datos, de personas en situación de vulnerabilidad en términos de su vínculo con la justicia, es un problema significativo. Esto afecta a grupos específicos como mujeres, minorías étnicas, personas con discapacidades y personas de bajos recursos, quienes quedan invisibilizados, dado que sus características particulares no se ven reflejadas en los conjuntos de datos disponibles. Como resultado de esto, los sistemas de IA tienen una capacidad limitada para ofrecer respuestas equitativas, lo que puede empeorar las desigualdades en términos de acceso a la justicia.

⁸ Una experiencia destacable de utilización de la metodología recursos comunes de datos, que permite disponibilizar datos de justicia, pero de una manera más restringida, es la del *Civil Justice Data Commons* de la Facultad de Derecho de la Georgetown University. Vid. <https://www.law.georgetown.edu/tech-institute/initiatives/georgetown-justice-lab/civil-justice-data-commons/cjdc-blog/a-civil-justice-data-commons/>

En consecuencia, es esencial realizar esfuerzos para mejorar la recolección y representación de datos sobre estos grupos en los sistemas judiciales, particularmente a través de medidas proactivas, como las que ofrece el principio de *equidad de datos* (World Economic Forum, 2024). Este principio se define como la asunción de una responsabilidad por parte de las instituciones de aplicar prácticas en el manejo de datos que respeten y promuevan los derechos humanos, la oportunidad y la dignidad. La equidad de datos exige una acción para garantizar que los sistemas basados en datos promuevan resultados justos y equitativos para todos los individuos, grupos y comunidades, en particular para aquellos que han sido históricamente marginados.

Para lograr esto se busca que los datos no sólo sean recopilados y gestionados de manera inclusiva, sino que también sean utilizados para generar y difundir resultados que reflejen de manera precisa y representativa la diversidad de la población, contribuyendo así a una administración de justicia más equitativa y efectiva.

Por otra parte, la creación de un estándar para la publicación de datos que incluyera variables clave en los conjuntos de datos a producir y publicar por parte de las instituciones judiciales, tales como la pertenencia a grupos en situación de vulnerabilidad, tales como minorías lingüísticas, étnicas o de género (Hivos, 2023), o bien la generación de capacidades por parte de las instituciones para recolectar datos sobre feminicidios (D'Ignazio et al., 2024) son esfuerzos focalizados en lograr esta completitud que deben ser sostenidos.

Un cuarto aspecto clave a tener en cuenta para la publicación de datos de justicia es el de la **privacidad**. Este aspecto se encuentra directamente ligado a la discusión respecto del acceso a los datos, debido a que la amplia accesibilidad propuesta debe estar conjugada con la protección del derecho a la privacidad y a la confidencialidad de información de índole personal.

El balance entre la cuestión de la privacidad y la necesidad de reutilizar datos de justicia en esquemas de creación de valor, como en este caso mediante la IA aplicada a la administración de justicia, es un tema de particular complejidad que requiere de un enfoque equilibrado que garantice los derechos de las personas sin dejar de lado la necesidad de contar con la mayor completitud posible en materia de los datos. En este sentido, ya existen pronósticos respecto de esta cuestión que hablan de una incipiente *voracidad* de datos de parte de los desarrollos de sistemas de IA que se espera que, en lo venidero, aumente la presión sobre todo el ecosistema de datos existente, al incrementar la cantidad y los tipos de datos recolectados. Esto pondrá, sin lugar a dudas, en tensión los marcos regulatorios para la protección de los derechos de las personas, en particular el derecho a la privacidad (King et al., 2024).

En junio de 2024, la Oficina del Comisionado de Privacidad de los Datos Personales de Hong Kong publicó una guía para instituciones y organizaciones que implementan y utilizan sistemas de IA con datos personales (PCPD, 2024). Esta guía propone para la protección de datos personales un marco basado en distintas fases, que comprende un enfoque de evaluación de riesgos para el uso de los datos mediante el cual las organizaciones puedan de manera sistemática identificar, analizar y evaluar los riesgos a lo largo de todo el ciclo de vida de los sistemas de IA. Entretanto, entre las prácticas más concretas se destaca la minimización de datos, esto es, limitar el uso de la información sólo a la estrictamente necesaria para la personalización y el funcionamiento de los sistemas de IA. Por otra parte, la guía subraya la importancia de la intervención humana para llevar adelante esta supervisión continua y la evaluación de riesgos, en

particular en aplicaciones consideradas como de alto riesgo, en conjunto con una comunicación transparente sobre el uso y el impacto de los sistemas de IA.

En todo caso, la privacidad es un aspecto que se encuentra lejos de estar resuelto y que forma parte de un debate continuo y multifacético, en el cual existen distintas partes interesadas y marcos regulatorios que, a la vez, sostienen perspectivas variadas y a menudo opuestas. Por ello, es preciso hacer hincapié en la importancia de continuar con los debates y considerar cuidadosamente las diversas perspectivas para lograr un equilibrio entre valores que parecen entrar en competencia, como son la transparencia y la accesibilidad de las decisiones judiciales para su uso en sistemas de IA y la protección del derecho a la privacidad de las personas (Van Opijnen, 2023).

La quinta y última dimensión que entra en discusión es la de la **ciberseguridad**. Esta dimensión posee un impacto claro en el desarrollo e implementación de los sistemas de IA, tanto a raíz de la necesidad de proteger los datos y los sistemas utilizados, así como a la necesidad de abordar problemas éticos y de privacidad de los datos.

Si bien el tema de la ciberseguridad se extiende a un contexto más amplio en el campo de la IA, esto es, en torno a la seguridad de los sistemas en sí mismos, en lo que respecta al uso de datos se requiere de una mirada holística que abarque no solo la protección de los datos en su primera dimensión (esto es, captura, almacenamiento y gestión), sino también los controles de riesgo en torno a la seguridad de la información utilizadas. Algunos de esos posibles riesgos pueden ser un ataque de *envenenamiento de datos*, que puede alterar el conjunto de entrenamiento de un modelo de IA y llevarlo a decisiones incorrectas o sesgadas (Cinà et al., 2023), o bien un acceso indebido a datos sensibles a través de la utilización de técnicas avanzadas de análisis y predicción que exploten las debilidades de los algoritmos utilizados (Bhati, 2024).

Desde un punto de vista regulatorio, el *Reglamento General de Protección de Datos* europeo establece la necesidad de adoptar políticas internas y aplicar medidas que cumplan en particular los principios de protección de los datos desde la fase de diseño y por defecto. Algunas de las medidas propuestas van en la línea de reducir a la mínima expresión el uso de datos personales, o establecer medidas eficaces de anonimización, o bien transparentar las funciones y el tratamiento de esos datos [Reglamento (UE) 2016/679, considerandos punto 78].

Gráfico 1. Principios para la publicación de datos de justicia



Niveles para la aplicación de los principios

Una vez establecidas las cinco dimensiones en torno a las cuales gira la cuestión de la disponibilización de datos de justicia para el uso en sistemas de IA de justicia responsable, es necesario establecer los distintos niveles o ámbitos en los cuales las instituciones de justicia deben dirimir estas dimensiones y establecer su postura en tanto instituciones responsables por la generación y publicación de esos datos.

Como sucede en general con la gobernanza de los datos digitales, las dimensiones arriba mencionadas deben dirimirse, por una parte, en un contexto institucional, pero asimismo en un contexto policéntrico, que involucra las distintas partes interesadas en el desarrollo de los sistemas. Se trata de un contexto definido por su naturaleza compleja y a menudo fragmentaria (Aguerre, 2024), que comprende entre otros actores involucrados a las instituciones de justicia que publican los datos, pero asimismo a empresas u organizaciones que harán uso de esos datos para el desarrollo de sistemas de IA.

Por ende, se establecen dos escalas o niveles de práctica y debate para dirimir estas cinco dimensiones. En un **nivel primario**, se espera que las instituciones de justicia, puedan dirimir las posturas institucionales en torno a la disponibilización de sus datos a través del desarrollo de políticas de gobernanza específicas para sus datos, que podrían incluir la adaptación de políticas, estándares, códigos y directrices específicos,

mientras que en algunos casos podrían asimismo requerir el establecimiento de nuevos marcos regulatorios. Estas iniciativas deberían tener como objetivo guiar el desarrollo ético y responsable de los sistemas de IA en el sector público y los enfoques adoptar podrían comprender tanto medidas *suaves*, como el establecimiento de directrices o estándares éticos, hasta nuevas regulaciones o legislación (OCDE, 2024c).

En este sentido, existen propuestas que se enfocan en la creación de modelos de gestión o gobernanza de los datos que prevén el establecimiento de unidades responsables de la dirección y calidad de los datos publicados, a las que se les puede encomendar el establecimiento de guías e instrumentos para garantizar la trazabilidad de los datos, su registro y conservación, así como otros aspectos ligados a su calidad (CLAD, 2023).

Pero por otra parte, el enfoque tradicional para la gobernanza de datos, que se centra únicamente en los requisitos legales de cada jurisdicción particular (es decir, en el ámbito en el que lleva a cabo su labor jurisdiccional cada institución), no es compatible con las características propias del uso que se hace de los datos para el desarrollo de sistemas de IA. Dicho uso, que probablemente trascienda los límites impuestos por una jurisdicción en particular, requiere alcanzar, en un **nivel secundario**, acuerdos entre diversos actores (incluso transnacionales), entre los que se encuentran órganos legislativos, las mismas instituciones que generan los datos, empresas u organizaciones que desarrollan y utilizan los sistemas de IA y personas usuarias del sistema de justicia. Así, la naturaleza policéntrica del desarrollo y uso de los sistemas de IA plantea desafíos en cuanto a su gobernanza, que deben ser abordados (Xue, 2024).

En este sentido, la gobernanza de la IA posee múltiples capas, entre las cuales se encuentran diversos actores que forman parte de una cadena en la cual coexisten y entran en tensión las perspectivas de las empresas u organizaciones desarrolladoras y los intereses y necesidades de la industria de la IA, pasando por regulaciones nacionales, o patrones regulatorios internacionales (como es el caso del citado *Reglamento Europeo de IA*). Esto hace que la implementación de principios robustos y generales relativos a la disponibilización de los datos y a su utilización sea especialmente compleja debido al gran número de partes involucradas. En todo caso, los mecanismos de gobernanza para la IA que busquen dirimir las dimensiones sobre la disponibilidad de datos de justicia pueden ser, en un nivel primario, producto de procesos internos de las instituciones, pero deberán encuadrarse dentro de un ecosistema de gobernanza más amplio, en el cual interactúan las diversas partes interesadas tanto en el ámbito de los ecosistemas nacionales de justicia como al ámbito global.

La necesidad de una gobernanza global para la IA es resultado, por una parte, del hecho de que las materias primas que utiliza, incluyendo los datos para su entrenamiento, se obtienen a nivel mundial. Al mismo tiempo, su despliegue atraviesa fronteras y genera múltiples aplicaciones en todo el mundo, lo cual tiene como resultado una concentración de poder a escala global, que tiene a su vez implicaciones geopolíticas y geoeconómicas que deben ser abordadas (High-level Advisory Body on AI, 2024). Dado que estas cuestiones deben resolverse con la colaboración de múltiples actores autónomos que deben gestionar y proporcionar acceso a recursos compartidos, es esencial establecer un modelo de gobernanza distribuido a nivel global y sectorial que facilite la coordinación, el acceso y la colaboración entre todos los actores involucrados (Verhulst, 2024).

Así, a fin de lograr esta gobernanza más efectiva de la IA de justicia, es necesario por lo tanto llevar la discusión a instancias institucionales de discusión internacional, o bien espacios multi-actor o *policy networks*, que involucren a las distintas partes que forman parte de la cadena para la generación de sistemas de IA específicamente basadas en datos del sistema de justicia⁹.

Una de las instituciones más relevantes que se han ocupado de este tema es UNESCO, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, agencia especializada de las Naciones Unidas que aporta a estos temas mediante la cooperación internacional en educación, ciencia, cultura y comunicación. UNESCO viene jugando un rol relevante en lo que respecta a la implementación de la IA en la labor judicial mediante una perspectiva ética y responsable. Ejemplos de esto son el establecimiento del *Global Network of Experts on AI and the Rule of Law*, un grupo de expertos que provee asistencia técnica y capacitación a operadores judiciales sobre la temática, una investigación que mide el nivel de conocimiento y uso de IA por parte de operadores judiciales (UNESCO, 2024), o bien el desarrollo actual de directrices para el uso de sistemas de IA en los tribunales¹⁰.

Existen algunos ejemplos de espacios multi-actor en los cuales tiene lugar el debate en torno a los datos y a su reutilización en sistemas de IA, que por lo tanto son relevantes para plantear y establecer recomendaciones de políticas en torno a las dimensiones planteadas. Uno de estos espacios institucionales es la *Dynamic Coalition on Data and Artificial Intelligence*, un espacio inserto en el Internet Governance Forum en el cual se discuten aspectos relacionados con el vínculo entre los datos y el desarrollo de sistemas de IA desde diversos puntos de vista de las partes involucradas e interesadas en el proceso¹¹.

La *Global Partnership on Artificial Intelligence* es otra iniciativa internacional que guía el desarrollo responsable de la IA mediante una colaboración entre los múltiples actores involucrados en el proceso, lo que incluye gobiernos, organizaciones internacionales y el sector privado¹². Finalmente, la *Alianza para el Gobierno Abierto* (OGP) representa un espacio en el cual el vínculo entre apertura de datos, Justicia Abierta y el uso de esos datos para el desarrollo de sistemas de IA responsable puede discutirse y en especial materializarse a través de la elaboración de compromisos de acción por parte de las instituciones a cargo de accionar estos principios¹³.

⁹ La naturaleza internacional de los desafíos planteados por la IA se puso de relevancia en noviembre de 2023 entre los países signatarios de la *Bletchley Declaration*, mediante la cual se hizo un llamado a la acción a partir de que “...Many risks arising from AI are inherently international in nature, and so are best addressed through international cooperation. We resolve to work together in an inclusive manner to ensure human-centric, trustworthy and responsible AI that is safe, and supports the good of all through existing international fora and other relevant initiatives, to promote cooperation to address the broad range of risks posed by AI...” (Bletchley Declaration, párrafo 5), *vid.* <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>

¹⁰ Estas directrices se encuentran en proceso de desarrollo al momento de concluir esta investigación. *Vid.* <https://unesdoc.unesco.org/ark:/48223/pf0000390781>

¹¹ *Vid.* <https://intgovforum.org/en/content/dynamic-coalition-on-data-and-artificial-intelligence-governance-dc-daig>

¹² *Vid.* <https://www.oecd.org/en/about/programmes/global-partnership-on-artificial-intelligence.html>

¹³ *Vid.* <https://www.opengovpartnership.org/policy-area/justice/>

Conclusión

Para ser centrados en las personas, los procesos de transformación digital en la justicia no deben limitarse a la adopción de nuevas tecnologías. Deben apoyarse en políticas de Justicia Abierta que orienten el cambio cultural y organizativo necesario para mitigar la exclusión digital y así garantizar que los beneficios de la digitalización alcancen de manera equitativa a todas las personas. En este sentido, la integración de la IA en la administración de justicia, aunque prometedora, debe ser abordada con cautela. La posibilidad de mejorar la eficiencia y la calidad del servicio judicial a través de estas tecnologías requiere un marco ético robusto que asegure la protección de los derechos fundamentales y el respeto a los principios de transparencia y rendición de cuentas. En este sentido, la colaboración entre actores gubernamentales, judiciales y de la sociedad civil será esencial para desarrollar sistemas de IA que sean verdaderamente inclusivos y responsables.

El rol de los datos en el desarrollo sistemas de IA responsable para la justicia debe abordarse, puesto que existe una interrelación clave entre esos datos, cómo se lleva a cabo su gobernanza y la necesidad de mitigar sesgos y errores que pueden afectar la equidad y la justicia en el marco de la aplicación de herramientas de IA al quehacer judicial.

Particularmente en contextos como América Latina, donde la invisibilidad de ciertos grupos y colectivos en su vínculo con el sistema judicial es un problema persistente, es importante promover una disponibilidad y apertura de datos que contemple no solo la transparencia y la calidad de los mismos, sino también un compromiso real con la inclusión y la equidad. Estos aspectos se presentan como pilares fundamentales para garantizar un posterior uso responsable de la IA, conforme a los lineamientos de organizaciones como la CEPEJ y el Reglamento Europeo de Inteligencia Artificial. Este último establece un marco normativo que busca regular los sistemas de IA considerados de alto riesgo, entre los cuales se incluye a aquellos utilizados en el ámbito judicial, resaltando la necesidad de mitigar sesgos y errores en los datos que pueden afectar la equidad y la justicia.

La implementación de sistemas de IA en el ámbito de la justicia requiere un enfoque riguroso y bien estructurado para la publicación y gestión de datos. A partir de las consideraciones discutidas, se identifican cinco dimensiones fundamentales a ser consideradas al respecto, cada una de las cuales presenta desafíos y oportunidades que deben ser abordados de manera conjunta, tanto a nivel primario por las instituciones que publican los datos como a nivel secundario, dada la naturaleza policéntrica de la cadena de valor de la IA, por parte de los diversos actores involucrados en el proceso.

La *estandarización* es crucial para asegurar la calidad y comparabilidad de los datos, facilitando así su uso en el entrenamiento de sistemas de IA. Sin un marco común, la integridad y utilidad de los datos se ven comprometidas, limitando la capacidad de innovación en el sector.

La *accesibilidad* es esencial para que los datos puedan ser utilizados efectivamente en el desarrollo de soluciones basadas en IA. La promoción de formatos abiertos y repositorios públicos contribuye a eliminar barreras, permitiendo que los datos sean reutilizados de manera inmediata y efectiva.

La *completitud* de los datos es fundamental para asegurar que todos los grupos sociales estén representados, evitando así la perpetuación de desigualdades. Es necesario implementar prácticas que garanticen la inclusión y representación adecuada de las poblaciones vulnerables.

La *privacidad* debe ser un principio rector en el manejo de datos, equilibrando la necesidad de acceso con la protección de derechos individuales. Esto implica desarrollar marcos regulatorios robustos y enfoques que prioricen la minimización de datos y la intervención humana.

Finalmente, la *ciberseguridad* se presenta como un aspecto vital para la protección de los datos utilizados en sistemas de IA. Se requiere de políticas que integren medidas de seguridad desde la fase de diseño, asegurando la integridad y confidencialidad de la información.

En conclusión, establecer un marco de gobernanza de datos robusto y colaborativo es esencial para avanzar en la utilización responsable de la IA en el ámbito judicial. La interacción entre múltiples actores, tanto a nivel local como internacional, será clave para abordar estos desafíos y fomentar un desarrollo ético y responsable de la IA en la justicia. Estas son las cuestiones que deben ser dirimidas de manera ineludible para asegurar que los beneficios de la IA se traduzcan en mejoras tangibles para el sistema de justicia y, en última instancia, para las personas en su conjunto.

Bibliografía

- Adams, R., Adeleke, F., Florido, A., de Magalhães Santos, L. G., Grossman, N., Junck, L., & Stone, K. (2024). *Global Index on Responsible AI 2024* (1st ed.). Global Center on AI Governance. <https://global-index.ai/>
- Adams, R. (2024, 27 de mayo). Responsible AI practices for business leaders (Episodio 13) [Episodio de podcast de video]. *Unpacked*. Lab 45. <https://lab45thinktank.com/podcast/responsible-ai-practices-for-business-leaders-with-dr-rachel-adams-ceo-gcg/>
- Addo, P. M., Baumann, D., McMurren, J., Verhulst, S. G., Young, A., & Zahuranec, A. J. (2021). *Usages émergents des technologies au service du développement : Un nouveau paradigme des intelligences*. Policy paper. AFD. <https://www.afd.fr/fr/ressources/technologies-developpement>
- Aguerre, C. (2024). Internet Interoperability and Polycentric Attributes in Global Digital Data Ordering. En C. Aguerre, M. Campbell-Verduyn & J.A. Scholte (eds.). *Global Digital Data Governance. Polycentric Perspectives*. Routledge, 34-51. https://www.researchgate.net/publication/377140969_Internet_Interoperability_and_Polycentric_Attributes_in_Global_Digital_Data_Ordering
- Aguerre, C. & Bustos Frati, L. (2021). *IA, justicia y políticas de transformación digital en el ámbito público latinoamericano*. Proyecto Preparación del sector judicial para la Inteligencia Artificial en América Latina. CETyS. <https://repositorio.udes.edu.ar/jspui/handle/10908/18634>
- Albrecht, K. & Filip, K. (2022). Public Records Aren't Public: Systemic Barriers to Measuring Court Functioning & Equity. *CJC Publications*, 41. https://scholarworks.gsu.edu/cj_facpub/41
- Belli, L. & Gaspar, W. (2024). AI Transparency, AI Accountability, and AI Sovereignty: An Overview. En L. Belli & W. Gaspar (eds.). *The Quest for AI Sovereignty, Transparency and Accountability*. Official Outcome of the UN IGF Data and Artificial Intelligence Governance Coalition. Fundação Getulio Vargas, 21-28. <https://diretorio.fgv.br/en/publication/quest-ai-sovereignty-transparency-and-accountability>
- Bhati, D. (2024, 6 de marzo). AI worm that can steal private data: What is it, how it works, and how to stay safe. *India Today*. <https://www.indiatoday.in/technology/news/story/ai-worm-that-can-steal-private-data-what-is-it-how-it-works-and-how-to-stay-safe-2511369-2024-03-06>
- Brown, A., Fishenden, J. & Thompson, M. (2014). Organizational Structures and Digital Transformation. En A. Brown, J. Fishenden & M. Thompson. *Digitizing Government. Business in the Digital Economy*. Palgrave Macmillan, 165-183. https://doi.org/10.1057/9781137443649_10
- CAF (2024). *Diseño de políticas públicas de inteligencia artificial. Desarrollo de habilitadores para su implementación en América Latina y el Caribe*. Comunidad Andina de Fomento. <https://scioteca.caf.com/handle/123456789/2241>
- Campbell, R.W. (2020). Artificial Intelligence in the Courtroom: The Delivery of Justice in the Age of Machine Learning. *Colorado Technology Law Journal*, 18(2), 323-349. https://ctlj.colorado.edu/wp-content/uploads/2020/08/2-Campbell_06.25.20.pdf

Canada Federal Court (2024). *Interim Principles and Guidelines on the Court's Use of Artificial Intelligence*. <https://www.fct-cf.gc.ca/en/pages/law-and-practice/artificial-intelligence#cont>

Carfi, N., Foti, J. & Fumega, S. (2024, 30 de julio). Beyond the Hype-Cycle for AI and Open Data. *OGP Horizons*. <https://medium.com/ogp-horizons/beyond-the-hype-cycle-for-ai-and-open-data-9f9b01f7b069>

CEPEJ (2024). *Use of Generative Artificial Intelligence (AI) by judicial professionals in a work related context*. Nota informativa preparada por el CEPEJ Working group on Cyberjustice and Artificial Intelligence. <https://www.coe.int/en/web/cepej/-/information-note-on-the-use-of-generative-artificial-intelligence-ai-by-judicial-professionals-in-a-work-related-context>

CEPEJ (2018). *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment*. Documento adoptado en la 31 sesión plenaria de la CEPEJ. <https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment>

Cinà, A.E., Grosse, K., Demontis, A., Vascon, S., Zellinger, W., Moser, B.A., Oprea, A., Biggio, B., Pelillo, M. & Roli, F. (2023). Wild Patterns Reloaded: A Survey of Machine Learning Security against Training Data Poisoning. *ACM Computing Surveys*, 55(13). <https://doi.org/10.1145/3585385>

CLAD (2023). *Carta Iberoamericana de Inteligencia Artificial en la Administración Pública*. <https://clad.org/declaraciones-consensos/carta-iberoamericana-de-inteligencia-artificial-en-la-administracion-publica/>

Conselho Nacional de Justiça, Resolução Nº 332 de 21/08/2020, *Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências*. <https://atos.cnj.jus.br/atos/detalhar/3429>

Comisión Europea (2024). *The Role of Artificial Intelligence in Processing and Generating New Data. An exploration of legal and policy challenges in open data ecosystems*. data.europa.eu. <https://data.europa.eu/sites/default/files/report/The%20Role%20of%20Artificial%20Intelligence%20in%20Processing%20and%20Generating%20New%20Data%20.pdf>

Consejo de Derechos Humanos (2024). *Contemporary forms of racism, racial discrimination, xenophobia and related intolerance*. A/HRC/56/68. Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance, Ashwini K.P. <https://undocs.org/A/HRC/56/68>

Cumbre Ministerial de Altas Autoridades de América Latina y el Caribe (2023). *Para promover una inteligencia artificial ética en América Latina y el Caribe (Declaración de Santiago)*. Firmada en Santiago de Chile, 23-24 de octubre de 2023. https://minciencia.gob.cl/uploads/filer_public/40/2a/402a35a0-1222-4dab-b090-5c81bbf34237/declaracion_de_santiago.pdf

Datasphere Initiative (2024, 25 de junio). *6 reasons why Data matters for AI*. The Datasphere. <https://www.thedatasphere.org/news/6-reasons-why-data-matters-for-ai/>

De Langen, M. & Muller, S. (2022, 17 de marzo). *Opinion: We must talk about better justice systems*. Devex. <https://www.devex.com/news/opinion-we-must-talk-about-better-justice-systems-102840>

- D'Ignazio, C., Cruxên, I., Martínez Cuba, A., Suárez Val, H., Dogan, A., & Ansari, N. (2024). Geographies of missing data: Spatializing counterdata production against femicide. *Environment and Planning D: Society and Space*, 0(0). <https://doi.org/10.1177/02637758241275961>
- Elena, S., Aquilino, N. & Pichón Rivière, A. (2014). *Emerging Impacts in Open Data in the Judiciary Branches in Argentina, Chile and Uruguay*. Case Study. CIPPEC. https://webfoundation.org/docs/2017/09/Arg_ChI-Case-study-CIPPEC.pdf
- Elena, S. & Mercado, J.G. (2018). Justicia Abierta: una aproximación teórica. En S. Elena (coord.), *Justicia Abierta: aportes para una agenda en construcción*. SAJJ, 17-42. <http://www.bibliotecadigital.gob.ar/items/show/1818>
- Engin, Z. (2024, 2 de septiembre). *Frontier AI: double-edged sword for public sector*. AI for Public Good, Joseph Rowntree Foundation. <https://www.jrf.org.uk/ai-for-public-good/frontier-ai-double-edged-sword-for-public-sector>
- Ferrara, E. (2023). Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies. *Sci* 6(1). <https://doi.org/10.3390/sci6010003>
- Friis, S. & Riley, J. (2023, 29 de septiembre). Eliminating Algorithmic Bias Is Just the Beginning of Equitable AI. *Harvard Business Review*. <https://hbr.org/2023/09/eliminating-algorithmic-bias-is-just-the-beginning-of-equitable-ai>
- Fuchs, D.J. (2018). The Dangers of Human-Like Bias in Machine-Learning Algorithms. *Missouri S&T's Peer to Peer* 2(1). <https://scholarsmine.mst.edu/peer2peer/vol2/iss1/1>
- Hess, B. & Koprivica Harvey, A. (2019). Open Justice in Modern Societies: What Role for Courts? En B. Hess & A. Koprivica Harvey (eds.). *Open Justice. The Role of Courts in a Democratic Society*. Studies of the Max Planck Institute Luxembourg for International, European and Regulatory Procedural Law, Vol. 13. Nomos Verlag, 9-46. <https://doi.org/10.5771/9783845297620>
- High-level Advisory Body on Artificial Intelligence (2024). *Governing AI for Humanity*. Final Report. September 2024. United Nations AI Advisory Body. https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf
- Hivos (2023). *Estándares de Datos Abiertos de Justicia y su guía de implementación*. Proyecto “Más datos, más inclusión: avanzando hacia una Justicia Abierta en Guatemala”. <https://america-latina.hivos.org/document/estandares-de-datos-abiertos-de-justicia-y-su-guia-de-implementacion/>
- ILDA (2020). *Data for development: the road ahead*. Iniciativa Latinoamericana para los Datos Abiertos. <https://idatosabiertos.org/en/publicaciones/english-data-for-development-a-road-ahead/>
- Jacobsen, B.N. (2023). Machine learning and the politics of synthetic data. *Big Data & Society*, 10(1). <https://doi.org/10.1177/20539517221145372>
- Jiménez-Gómez, C. (2017). Hacia el Estado abierto: justicia abierta en América Latina y el Caribe. En A. Naser, A. Ramírez-Alujas & D. Rosales (eds.). *Desde el gobierno abierto al Estado abierto en América Latina y el Caribe*. CEPAL, 231-251. <https://www.cepal.org/es/publicaciones/44769-gobierno-abierto-al-estado-abierto-america-latina-caribe>

King, J. & Meinhardt, C. (2024). *Rethinking Privacy in the AI Era. Policy Provocations for a Data-Centric World*. White Paper. HAI. <https://hai.stanford.edu/white-paper-rethinking-privacy-ai-era-policy-provocations-data-centric-world>

Lagioia, F., Rovatti, R. & Sartor, G. (2023). Algorithmic fairness through group parities? The case of COMPAS-SAPMOC. *AI & Society* 38, 459–478. <https://doi.org/10.1007/s00146-022-01441-y>

Marković, M. & Gostojić, S. (2020). Open Judicial Data: A Comparative Analysis. *Social Science Computer Review*, 38(3), 295-314. <https://doi.org/10.1177/0894439318770744>

Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J.C., Shoham, Y., Wald, R. & Clark, J. (2024). *The AI Index 2024 Annual Report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. <https://hai.stanford.edu/research/ai-index-report>

Medvedeva, M., Wieling, M. & Vols, M. (2023). Rethinking the field of automatic prediction of court decisions, *Artificial Intelligence and Law*, 31, 195–212. <https://doi.org/10.1007/s10506-021-09306-3>

Mercado, J.G. (2021). Justicia Abierta y datos abiertos de justicia: herramientas conceptuales y estudio de caso sobre datos abiertos subnacionales en Argentina. En L.F. Aguilar Villanueva, E.A. Ruvalcaba & J.I. Criado (coords.). *Gobierno Abierto en Iberoamérica. Estudios de caso en gobiernos subnacionales*. Volumen 1. Tirant Lo Blanch, 129-166. <https://editorial.tirant.com/mex/libro/gobierno-abierto-en-iberoamerica-estudios-de-caso-en-gobiernos-subnacionales-volumen-1-luis-f-aguilar-villanueva-9788413971063>

Muller, S. & Barendrecht, M. (2013). *The Justice Innovation Approach: How Justice Sector Leaders in Development Contexts Can Promote Innovation*. The World Bank Legal Review: Legal Innovation and Empowerment for Development, Vol. 4, 17-30. https://doi.org/10.1596/9780821395066_CH02

National Center for State Courts (2020). *National Open Court Data Standards*. <https://www.ncsc.org/consulting-and-research/areas-of-expertise/data/national-open-court-data-standards-nods>

Nyholm, S. (2024). What Is This Thing Called the Ethics of AI and What Calls for It? En D.J. Gunkel (ed.). *Handbook on the Ethics of Artificial Intelligence*. Edward Elgar Publishing. <https://doi.org/10.4337/9781803926728.00006>

OCDE (2024a). *Shaping a Rights-Oriented Digital Transformation*, OECD Digital Economy Papers, No. 368. OECD Publishing. <https://doi.org/10.1787/86ee84e2-en>

OCDE (2024b). *AI, Data Governance and Privacy: Synergies and Areas of International Cooperation*. OECD Artificial Intelligence Papers, No. 22. OECD Publishing. <https://doi.org/10.1787/2476b1a4-en>

OCDE (2024c). *Governing with Artificial Intelligence: Are Governments Ready?* OECD Artificial Intelligence Papers, No. 20. OECD Publishing. <https://doi.org/10.1787/26324bc2-en>

OCDE (2023). *Recommendation of the Council on Access to Justice and People-Centered Justice Systems* (OECD/LEGAL/0498). <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0498>

Open Data Institute (2024). *Building a Better Future with Data and AI: A White Paper*. ODI Research. <https://theodi.org/insights/reports/building-a-better-future-with-data-and-ai-a-white-paper/>

- Pathfinders for Peaceful, Just and Inclusive Societies (2019). *The Roadmap for Peaceful, Just and Inclusive Societies. A Call to Action to Change our World*. Center on International Cooperation. <http://www.sdg16.plus>
- PCPD (2024). *Artificial Intelligence: Model Personal Data Protection Framework*. Office of the Privacy Commissioner for Personal Data. https://www.pcpd.org.hk/english/resources_centre/publications/files/ai_protection_framework.pdf
- PNUD (2024). *Digital Inclusion Playbook 2.0: From Access to Empowerment in a Dynamic World*. UNDP Policy Center Singapore. <https://www.undp.org/policy-centre/singapore/publications/digital-inclusion-playbook-20>
- Programa Justicia Abierta (2022). *Protocolo técnico de datos y procesos, versión III*. <https://acortar.link/T5OJoi>
- Red Internacional de Justicia Abierta (2021). *Mapa de portales de datos abiertos de justicia*. https://public.tableau.com/app/profile/justicia.abierta/viz/Mapasitioswebjusticia_15958628213070/maparegional
- Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, del 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (*Reglamento general de protección de datos*). <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32016R0679#d1e3433-1-1>
- Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, del 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (*Reglamento de Inteligencia Artificial*). <http://data.europa.eu/eli/reg/2024/1689/oj>
- Reiling, A.D. (2020). Courts and Artificial Intelligence. *International Journal for Court Administration*, 11(2), <https://doi.org/10.36745/ijca.343>
- Rocha, C. & Carvalho, J. (2023). *Artificial Intelligence in the Judiciary: Uses and Threats*. CEUR. <https://ceur-ws.org/Vol-3399/paper17.pdf>
- Rostain, T. (2024). *Access to Justice as Access to Data*. Georgetown Law Faculty Publications and Other Works. <https://scholarship.law.georgetown.edu/facpub/2619>
- Santiso, C. (2022). *Trust with Integrity: Harnessing the Integrity Dividends of Digital Government for Reducing Corruption in Developing Countries*. UN Department of Economic and Social Affairs (DESA) Working Papers. <https://doi.org/10.18356/25206656-176>
- Seepma, A.P., de Blok, C. & Van Donk, D.P. (2021). Designing digital public service supply chains: four country-based cases in criminal justice, *Supply Chain Management*, 26(3), 418-446. <https://doi.org/10.1108/SCM-03-2019-0111>
- Sharp Vargas, P. (2020). La defensoría penal pública en el camino del Gobierno Abierto. *Revista 93. La Revista de la Defensoría Penal Pública*, 23, 60-63. <https://revista93.dpp.cl/gobierno-abierto-un-nuevo-camino-de-legitimidad-para-el-estado-y-la-justicia/>
- Socol de la Osa, D.U. & Remolina, N. (9-11 de julio de 2024). *AI at the Bench: Legal and Ethical Challenges of Informing -or Misinforming- Judicial Decision-Making Through Generative AI*. Data for Policy, Londres, Reino Unido. <https://dx.doi.org/10.2139/ssrn.4860853>

- Stryker, C. (6 de febrero de 2024). What is responsible AI? *IBM*. <https://www.ibm.com/topics/responsible-ai>
- Susskind, R. (2013). *Tomorrow's Lawyers. An Introduction to your Future*. Oxford: Oxford University Press.
- UK Judiciary (2023). *Artificial Intelligence (AI) Guidance for Judicial Office Holders*. <https://www.judiciary.uk/wp-content/uploads/2023/12/AI-Judicial-Guidance.pdf>
- UNESCO (2024). *Survey on the Use of AI Systems by Judicial Operators*. UNESCO Global Judges' Initiative. <https://unesdoc.unesco.org/ark:/48223/pf0000389786>
- Van Opijnen, M. (2023). *The GDPR and the Reuse of Published Court Decisions; Some Pressing Questions, Illustrated by Developments in The Netherlands*. Presentado en la conferencia Law Via the Internet, Viena, 7 y 8 de noviembre.
- Verhulst, S. (2024, 4 de junio). Toward a Polycentric or Distributed Approach to Artificial Intelligence & Science. *Frontiers Policy Labs*. <https://policylabs.frontiersin.org/content/commentary-toward-a-polycentric-or-distributed-approach-to-artificial-intelligence-science>
- Vial, G. (2019). Understanding Digital Transformation: A Review and a Research Agenda. *The Journal of Strategic Information Systems*, 28, 118-144. <https://doi.org/10.1016/j.jsis.2019.01.003>
- Wang, N. & Tian, M.Y. (2023). Intelligent Justice: human-centered considerations in China's legal AI transformation, *AI Ethics*, 3, 349-354. <https://doi.org/10.1007/s43681-022-00202-3>
- World Justice Project (2023). *Disparities, Vulnerability, and Harnessing Data for People-Centered Justice*. WJP Justice Data Graphical Report II. <https://worldjusticeproject.org/our-work/research-and-data/wjp-justice-data-graphical-report-ii>
- Webster, H. (2024, 28 de junio). Are Risk Assessment Tools Setting the Stage for AI Judges? *The Bail Project*. <https://bailproject.org/learn/are-risk-assessment-tools-setting-the-stage-for-ai-judges/>
- World Economic Forum (2024). *Data Equity Definition*. Global Future Council on Data Equity. https://www3.weforum.org/docs/WEF_Data_Equity_Definition_2024.pdf
- Xue, J.H. (2024). Polycentric Theory Diffusion and AI Governance. En C. Aguerre, M. Campbell-Verduyn & J.A. Scholte, *Global Digital Data Governance. Polycentric Perspectives*. Routledge, 223-237. https://www.researchgate.net/publication/377135755_Polycentric_Theory_Diffusion_and_AI_Governance