Sage Sphere in Artificial Intelligence

ISSN: 3103-1307

Vol. 4 Num. 1 (2026)

SSAI-Vol.4. N1. 006

Protección de datos, transparencia algorítmica y compliance: una revisión de los desafíos jurídicos de la inteligencia artificial

Data protection, algorithmic transparency, and compliance: a review of the legal challenges of artificial intelligence

Autores:
Francisco Javier Santini Rodriguez
Universidad de Sonora
Hermosillo – México
francisco.santini@unison.mx
https://orcid.org/0000-0001-5437-4591

Luis Carlos Rodriguez Montaño
Universidad de Sonora
Hermosillo – México
luis.rodriguez@unison.mx
https://orcid.org/0000-0003-1245-8943

Mario Isaac Romero Lamadrid
Centro de Investigación en Alimentación y Desarrollo (CIAD)
Hermosillo – México
mromero423@estudiantes.ciad.mx
https://orcid.org/0009-0008-3464-4097

Autor de correspondencia: *Francisco Javier Santini Rodriguez*, francisco.santini@unison.mx

Enviado: 19-01-2026 **Revisado:** 14-03-2026
Aceptado: 17-08-2026 **Publicado:** 19-08-2026

Cómo citar este artículo:
Santini Rodriguez, F. J., Rodriguez Montaño, L. C., & Romero Lamadrid, M. I. (2026). Data protection, algorithmic transparency, and compliance: a review of the legal challenges of artificial intelligence. *Sage Sphere in Artificial Intelligence*, *4*(1), 1-25. https://doi.org/10.63688/b35xe061

ISSN: 3103-1307

Vol. 4 Num. 1 (2026)

RESUMEN

Introducción: La proliferación de sistemas de inteligencia artificial ha agudizado la cuestión jurídica sobre la protección de datos, la transparencia de los algoritmos y la rendición de cuentas de las entidades. Objetivo: Realizar una revisión de la literatura científica sobre los riesgos jurídicos que plantea la inteligencia artificial y el mecanismo de gestión de su gobernanza y compliance. Metodología: Se llevó a cabo una revisión integrativa, jurídica y crítica de un total de 23 contribuciones, de la que se ha dado cuenta en la literatura del período de tiempo comprendido a partir del 2016 hasta el 2026, mayormente en Scopus y SciELO. La revisión ha agrupado los estudios en muy pocas áreas: protección de datos; transparencia; accountability y compliance; y gobernanza de la IA. Resultados: La literatura revisada muestra la progresión de los enfoques centrados en la privacidad y la transparencia a un modelo en el que priman: la explicabilidad, la auditabilidad, la trazabilidad, la supervisión humana y la accountability. Se ha puesto de manifiesto el compliance algorítmico como mecanismo de controles preventivos y de gestión de riesgo. Conclusiones: La regulación de la IA sobrepasa el simple formalismo e incorpora, en las estructuras organizacionales que sean controlables.

Palabras clave: inteligencia artificial; protección de datos; transparencia algorítmica; compliance; gobernanza algorítmica.

ABSTRACT

Introduction: The proliferation of artificial intelligence systems has sharpened the legal issue of data protection, the transparency of algorithms and the accountability of entities. Objective: To conduct a review of the scientific literature on the legal risks posed by artificial intelligence and the mechanism for managing its governance and compliance. Methodology: An integrative, legal and critical review of a total of 23 contributions was carried out, which has been reported in the literature for the period of time from 2016 to 2026, mainly in Scopus and SciELO. The review has grouped the studies into very few areas: data protection; transparency; accountability and compliance; and AI governance. Results: The literature reviewed shows the progression from approaches focused on privacy and transparency to a model in which explainability, auditability, traceability, human supervision and accountability prevail. Algorithmic compliance has been highlighted as a mechanism for preventive controls and risk management. Conclusions: The regulation of AI goes beyond simple formalism and incorporates, in organizational structures, that are controllable.

Keywords: artificial intelligence; data protection; algorithmic transparency; compliance; algorithmic governance.

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

1. INTRODUCCIÓN

La inteligencia artificial (IA) avanza a pasos agigantados y, por lo tanto, los discursos de análisis, predicción y decisión empiezan a cambiar o, en algunos casos, incluso a desaparecer por culpa de las organizaciones públicas y privadas que, ya, utilizan sistemas algorítmicos o "sistemas automatizados" para abordar ciertos problemas que un algoritmo humano no puede realizar como, por ejemplo, los procesos de selección del personal, de evaluación crediticia, educación, salud, seguridad y entrega de servicios públicos. A medida que los sistemas algorítmicos son utilizados y/o extensamente propagados, poco a poco van surgiendo características tales como, por un lado, la automatización y, por el otro, el procesamiento de grandes cantidades de información, pero, al mismo tiempo, también surgen interpelaciones jurídicas que hacen o se plantean, como cómo se pueden conciliar aspectos tales como la protección de datos personales, la opacidad del modelo, la explicación de las decisiones o la identificación de la responsabilidad de las posibles vulneraciones de los derechos humanos (Mittelstadt et al., 2016; Floridi et al., 2018).

Uno de los mayores puntos de discrepancia de la evolución de los sistemas de IA es la protección de los datos como tal. Para ponerlos en marcha y enseñarlos hacen falta datos que pueden ser recopilados, mezclados o reutilizados para cumplir con objetivos distintos a los originalmente previstos. Mendoza Enríquez (2021) señala que esa forma de proceder está muy reñida con principios como el de finalidad, el de proporcionalidad, el de minimización o el de control de los titulares sobre sus datos. La situación se vuelve más compleja aún cuando la IA llega a hacer inferencias sobre comportamientos, preferencias o características personales que no se han aportado. En esta línea, Durand Goyzueta et al. (2026) advierten que la forma en que los tratamientos se realizan de forma algorítmica puede dar lugar a nuevas formas de exposición al tratamiento de información sensible, lo que exige reforzar las garantías de privacidad o de gobernanza digital.

La transparencia algorítmica, junto a la protección de datos, ha pasado a ser uno de los aspectos clave del debate jurídico. Como analiza Burrell (2016), la opacidad de los algoritmos puede provenir de la complejidad técnica de los muchos modelos, de las restricciones organizacionales e incluso de la capacidad de interpretar cómo una entrada generó una salida. Por ello, saber que una entidad utiliza un algoritmo no equivale, en sí

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

mismo, a conocer el modo en que opera ni a poder determinar cuáles fueron los elementos que jugaron un papel determinante en una decisión dada. De hecho, como argumentan Ananny y Crawford (2018), puede pasar que la mera transparencia no sea suficiente en un contexto en el cual no existe un modo de poder interpretar, cuestionar y supervisar sus sistemas automatizados.

Este problema ha propiciado la discusión acerca de la explicabilidad y del derecho a obtener información clara concerniente a las decisiones automatizadas. Por un lado, Wachter et al. (2017) impugnan que derecho general a la explicación pueda ser hallado en el régimen europeo de protección de datos tanto si se hubiese desarrollado alguna práctica en tal sentido. Por otro lado, Selbst y Powles (2017) defienden una interpretación ampliamente más favorable a las obligaciones de información declaradas en tales procedimientos. En la proposición de Malgieri y Comandé (2017) se encuentra la inclusión de la naturaleza lectora como condición para poder alcanzar la comprensión y la controversia de decisiones producidas por sistemas automatizados.

Recientemente, el debate ha pasado de reflexionar solamente sobre la transparencia a cuestiones más amplias de la «accountability», de la «auditabilidad» y de la «gobernanza algorítmica». Por ejemplo, Lepri et al. (2018) advierten de que los sistemas automatizados tienen que ser evaluados no sólo por los resultados materiales o incluso por la eficacia técnica de su funcionamiento, sino también por su capacidad para ofrecer decisiones justas, que sean transparentes y que puedan ser atribuibles en caso de errores. De manera similar, Busuioc (2021), en este mismo sentido, quiere advertir de que la accountability implica que haya responsables, y que haya supervisión y consecuencias calificadas después de los errores o de los efectos no deseados que puedan llegar como consecuencia de un mal uso. En esta misma línea de pensamiento, Novelli et al. (2024) afirman que la accountability (rendición de cuentas) relativa a la inteligencia artificial significa contar con estructuras organizacionales que puedan justificar, supervisar y corregir el funcionamiento de aquellos sistemas.

Esta transformación está en línea con el compliance, pues los compliance se han centrado en evitar infracciones a través de políticas internas, gestión de riesgo, controles y seguimiento. Las herramientas, para poder ser usadas por la IA en procesos organizativos, deben ser generalizadas hacia lo tecnocientífico. La evaluación del impacto, la trazabilidad, las auditorías algorítmicas, la documentación de modelos y bajo el control humano,

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

son elementos que permiten transformar los principios jurídicos en procedimientos verificables. Mantelero (2018) señala que para que el compliance no quede en una mera verificación formal, las evaluaciones de IA deben incluir dimensiones jurídicas, sociales y éticas.

La gobernanza en IA se genera ante el auge de normas internacionales para garantizar sistemas confiables y con responsabilidad. Jobin et al. (2019) encontraron que había mucha convergencia en los principios como la transparencia, la justicia, la privacidad, la responsabilidad y la no maleficencia. El que haya principios comunes no los convierte en un compromiso, el reto está en convertir los principios en obligaciones, en controles internos, en mecanismos de supervisión.

Como ocurre en América Latina, la discusión que a continuación se expone resulta pertinente por el auge de la digitalización y la protección de los datos. Sánchez Acevedo (2022) afirma que el empleo de la inteligencia artificial (IA) en el ámbito público debe respetar derechos fundamentales y Gómez Rodríguez (2026) se hace eco de la necesidad de modelos de gobernanza que contemplen la innovación tecnológica y el control institucional. La literatura científica sobre el tema ha aumentado, pero la mayoría de la literatura escribe la protección de datos, la transparencia, la explicabilidad y la obligación algorítmica de un modo separado y por ende no es posible incorporar estas dimensiones en modelos de gobernanza y cumplimiento. Puede ser interesante exponer estas dimensiones desde una óptica jurídica y de organización. El artículo expone la literatura científica sobre los retos jurídicos en protección de datos, transparencia algorítmica, cumplimiento en inteligencia artificial, riesgos, tendencias regulatorias y propuestas de gobernanza y obligación.

2. METODOLOGÍA

La investigación fue elaborada mediante una revisión integrativa de la literatura, de un enfoque jurídico-crítico dirigido a identificar y analizar los principales retos asociados a la protección de datos personales y a la transparencia algorítmica, así como incorporar el compliance en el uso de sistemas de inteligencia artifical. Dicha revisión integrativa de la literatura permite articular aportaciones del derecho digital, la ética en el seno de la inteligencia artificial, la gobernanza tecnológica y la protección de datos, intentando

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

construir una interpretación articulada de las principales tendencias que pueden encontrarse en la literatura por especialistas.

La búsqueda bibliográfica se redujo a las bases de datos Scopus y SciELO, por la extensa cobertura de publicaciones científicas arbitradas, pero también por su idoneidad para las ciencias jurídicas y sociales. Complementariamente se seleccionó un artículo de revista, el cual se publicó en la revista NeuroData, por la relación directa que existe entre tal artículo con la privacidad cognitiva, el tratamiento de información a través de inteligencia artificial y la gobernanza algorítmica. El periodo temporal para las publicaciones se considera entre 2016 y 2026, cuya elección de intervalo viene dada por el hecho de que en estos años se explicó el debate académico y normativo sobre decisiones automatizadas, explicabilidad, responsabilidad algorítmica y regulación de la inteligencia artificial.

Para la recuperación de los documentos se utilizaron términos en español e inglés relacionados con los ejes centrales de la investigación, entre ellos: “inteligencia artificial”, “protección de datos”, “transparencia algorítmica”, “decisiones automatizadas”, “explicabilidad”, “responsabilidad algorítmica”, “compliance”, “gobernanza de inteligencia artificial”, “artificial intelligence”, “data protection”, “algorithmic transparency”, “automated decision-making”, “right to explanation”, “algorithmic accountability” y “AI governance”. Estos términos fueron combinados mediante operadores booleanos AND y OR para identificar publicaciones que abordaran simultáneamente aspectos tecnológicos y jurídicos.

Como criterios de inclusión se consideraron artículos científicos relacionados directamente con inteligencia artificial y, al menos, uno de los siguientes componentes: protección de datos personales, transparencia o explicabilidad algorítmica, decisiones automatizadas, rendición de cuentas, gobernanza tecnológica o mecanismos de cumplimiento normativo. Se priorizaron estudios publicados en revistas indexadas, documentos con DOI o enlace electrónico verificable y trabajos que aportaran elementos conceptuales, regulatorios o críticos relevantes para el objetivo de la revisión. También se consideraron investigaciones internacionales y latinoamericanas con el propósito de contrastar diferentes aproximaciones jurídicas al fenómeno.

Se excluyeron publicaciones centradas exclusivamente en el rendimiento técnico de algoritmos, desarrollo de modelos computacionales o aplicaciones de inteligencia

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

artificial sin implicaciones jurídicas o regulatorias. Asimismo, se descartaron documentos duplicados, trabajos sin relación suficiente con los objetivos de la investigación y fuentes cuyo contenido no permitiera analizar alguno de los ejes definidos. Después de aplicar estos criterios y evaluar la pertinencia temática de las publicaciones, se conformó un corpus final de 23 documentos considerados relevantes para el análisis.

La información fue organizada mediante una matriz de análisis bibliográfico que contempló autor y año de publicación, contexto de estudio, problema jurídico abordado, concepto central, principales riesgos identificados, mecanismos de respuesta propuestos y aporte al debate sobre gobernanza de la inteligencia artificial. Posteriormente, los hallazgos fueron sometidos a un proceso de síntesis temática y comparación jurídico-conceptual.

El análisis permitió establecer cuatro categorías principales: protección de datos personales y privacidad frente a sistemas de inteligencia artificial; transparencia, explicabilidad y derecho a una explicación; accountability, compliance y responsabilidad jurídica; y gobernanza de la inteligencia artificial y tendencias regulatorias. La interpretación no se limitó a describir los documentos seleccionados, sino que buscó identificar coincidencias, divergencias, vacíos y transformaciones en la literatura, especialmente respecto del tránsito desde modelos centrados en la transparencia hacia enfoques de gobernanza y compliance algorítmico. Este procedimiento permitió integrar los diferentes aportes en una perspectiva común orientada a comprender cómo los mecanismos de prevención, supervisión, auditabilidad y rendición de cuentas pueden contribuir a una utilización jurídicamente responsable de la inteligencia artificial.

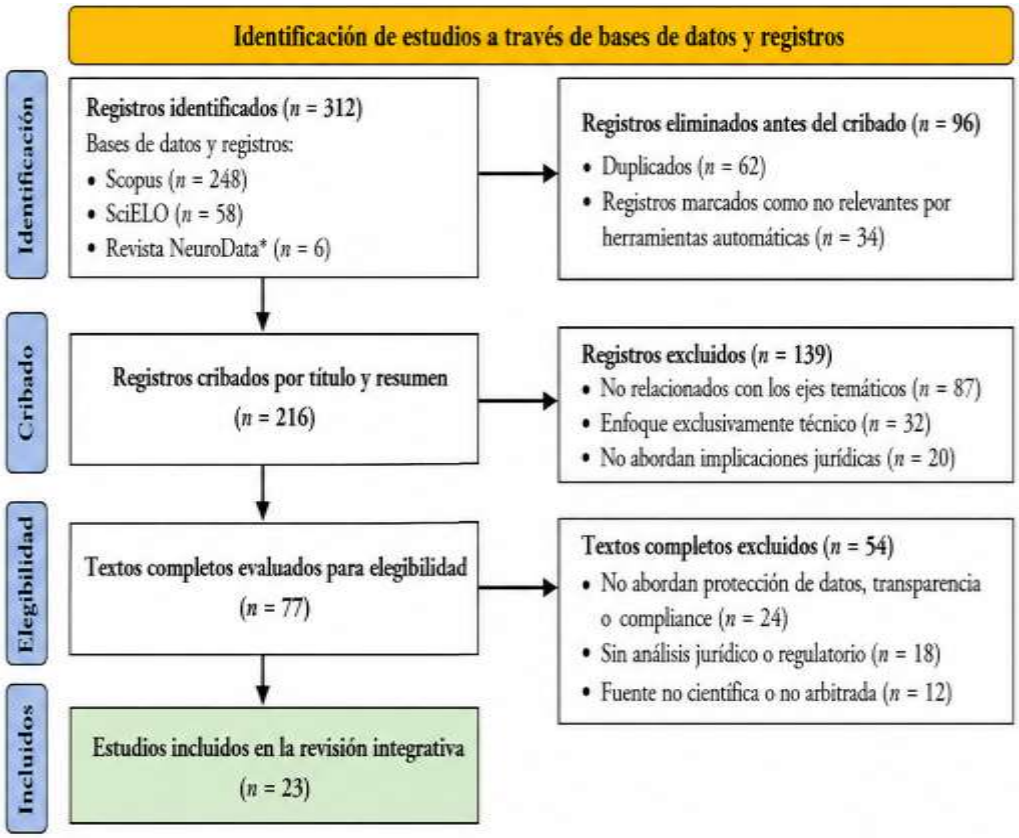
Figura 1.

Diagrama de flujo PRISMA de selección de estudios

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307

Vol. 4 Num. 1 (2026)



Nota. Se incluyó un documento adicional proveniente de la revista *NeuroData* por su pertinencia temática.

El proceso de selección permitió identificar inicialmente 312 registros procedentes de Scopus, SciELO y la revista *NeuroData*. Después de eliminar 62 documentos duplicados y 34 registros considerados no pertinentes en la fase inicial, se sometieron 216 publicaciones al cribado por título y resumen. En esta etapa se excluyeron 139 documentos por no relacionarse directamente con los ejes de la revisión, presentar un enfoque exclusivamente técnico o no abordar implicaciones jurídicas. Posteriormente, se evaluaron 77 textos completos, de los cuales 54 fueron excluidos por no desarrollar suficientemente aspectos de protección de datos, transparencia algorítmica o compliance, carecer de análisis jurídico o regulatorio, o corresponder a fuentes no científicas o no arbitradas. Finalmente, 23 estudios cumplieron los criterios establecidos y conformaron el corpus definitivo de la revisión.

3. RESULTADOS

El análisis de los veintitrés estudios analizados hizo evidente una evolución directa en el tratamiento jurídico de la IA; si los trabajos publicados entre 2016 y 2019 se centraron

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

especialmente en la opacidad de los algoritmos, la ética de los sistemas automatizados, la transparencia o el primer debate sobre el derecho a la explicación, el resto de las investigaciones que han salido a la luz aportan perspectivas que implican la noción de protección de datos inferidos, como la auditabilidad, la trazabilidad, la supervisión humana, la evaluación de impacto, la accountability o la gobernanza de la IA. En cuanto a los hallazgos, se muestra un cambio desde los enfoques centrados en informar sobre el modo de funcionamiento de los algoritmos hacia modelos centrados en la prevención de riesgos, la atribución de responsabilidades o la ordenación de controles organizacionales.

De la misma manera, la revisión de literatura mostró una fuerte presencia del tipo de literatura producida en el marco europeo, estrechamente vinculada con la protección de datos y las decisiones automatizadas, mientras que los estudios latinoamericanos dan muestras de un crecimiento reciente, pero aportan otras nociones relacionadas con derechos fundamentales, administración pública, privacidad o gobernanza tecnológica. La Tabla 1 recoge, como resumen, las características y aportes principales de los estudios revisados.

Tabla 1.

Características generales y principales aportes de los estudios incluidos en la revisión

Autor(es) y año	Contexto principal	Eje de análisis	Fuente	Aporte principal
Ananny y Crawford (2018)	Internacional	Transparencia algorítmica	Scopus	Cuestionan que la transparencia por sí sola garantice accountability y resaltan la necesidad de mecanismos institucionales de control.
Arnesen (2026)	Unión Europea	Derecho a explicación	Scopus	Analiza las obligaciones de información sobre decisiones automatizadas y su relación con el derecho a explicación.

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

Autor(es) y año	Contexto principal	Eje de análisis	Fuente	Aporte principal
Burrell (2016)	Internacional	Opacidad algorítmica	Scopus	Identifica fuentes técnicas, organizacionales y cognitivas de opacidad en sistemas de aprendizaje automático.
Busuioc (2021)	Administración pública	Accountability	Scopus	Propone identificar responsables y establecer mecanismos de supervisión para garantizar rendición de cuentas algorítmica.
Cárdenas Poveda y Sarmiento Salcedo (2025)	Colombia	Gobernanza digital	SciELO	Vinculan gobierno digital, principios éticos de IA y función administrativa.
Demetzou et al. (2023)	Perspectiva comparada	Protección de datos	Scopus	Reorientan la protección de datos hacia los efectos concretos de las decisiones automatizadas.
Durand Goyzueta et al. (2026)	Educación superior	Privacidad cognitiva	NeuroData	Analizan riesgos asociados con inferencias algorítmicas y tratamiento de información en plataformas digitales.
Felzmann et al. (2019)	Unión Europea	Transparencia	Scopus	Relacionan las exigencias jurídicas de transparencia

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

Autor(es) y año	Contexto principal	Eje de análisis	Fuente	Aporte principal
				con las condiciones contextuales de aplicación.
Floridi et al. (2018)	Internacional	Ética y gobernanza	Scopus	Proponen principios para una IA socialmente beneficiosa, responsable y confiable.
García-Llorente y Olmeda (2026)	Unión Europea y Estados Unidos	Transparencia y auditabilidad	Scopus	Vinculan inteligibilidad, evidencia técnica y auditabilidad con la exigibilidad jurídica de la transparencia.
Gómez Rodríguez (2026)	Administración pública	Gobernanza de IA	SciELO	Plantea estructuras institucionales para articular innovación tecnológica, control y garantías jurídicas.
Jobin et al. (2019)	Internacional	Ética de IA	Scopus	Identifican convergencias internacionales en privacidad, transparencia, justicia y responsabilidad.
Lepri et al. (2018)	Internacional	Equidad y accountability	Scopus	Proponen evaluar decisiones algorítmicas a partir de justicia, transparencia y responsabilidad.
Li et al. (2026)	Perspectiva comparada	Datos para entrenamiento de IA	Scopus	Analizan las bases jurídicas utilizadas para legitimar el tratamiento de datos

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

Autor(es) y año	Contexto principal	Eje de análisis	Fuente	Aporte principal
				destinados al entrenamiento de IA.
Malgieri y Comandé (2017)	Unión Europea	Legibilidad	Scopus	Defienden la legibilidad como condición para comprender y cuestionar decisiones automatizadas.
Mantelero (2018)	Internacional	Evaluación de impacto	Scopus	Integra dimensiones jurídicas, sociales y éticas en la evaluación de riesgos de IA y Big Data.
Mendoza Enríquez (2021)	México	Protección de datos	SciELO	Examina los desafíos de la IA frente a finalidad, minimización y control sobre datos personales.
Mittelstadt et al. (2016)	Internacional	Ética algorítmica	Scopus	Sistematizan riesgos relacionados con opacidad, sesgos, discriminación y responsabilidad.
Novelli et al. (2024)	Internacional	Accountability	Scopus	Conceptualizan la rendición de cuentas como un sistema de obligaciones, supervisión, justificación y corrección.
Sánchez Acevedo (2022)	Sector público	Derechos fundamentales	SciELO	Examina límites jurídicos del uso de IA por las administraciones públicas.

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

Autor(es) y año	Contexto principal	Eje de análisis	Fuente	Aporte principal
Selbst y Powles (2017)	Unión Europea	Derecho a explicación	Scopus	Defienden una interpretación amplia de las obligaciones de ofrecer información significativa.
Viollier y Contreras (2026)	Unión Europea y Chile	Decisiones automatizadas	SciELO	Comparan el derecho a explicación en el GDPR y la normativa chilena.
Wachter et al. (2017)	Unión Europea	Decisiones automatizadas	Scopus	Cuestionan la existencia de un derecho general a explicación individualizada bajo el GDPR.

Nota. Los estudios fueron clasificados según su eje principal de análisis, aunque varios abordan simultáneamente protección de datos, transparencia, responsabilidad y gobernanza. A partir de estos estudios se establecieron cuatro categorías analíticas: protección de datos y privacidad; transparencia y explicabilidad algorítmica; accountability, compliance y responsabilidad jurídica; y gobernanza y tendencias regulatorias. La mayor concentración de trabajos se observó en torno a transparencia y explicabilidad, seguida por protección de datos y mecanismos de gobernanza.

Protección de datos y privacidad

La cuestión de la protección de datos se perfila como uno de los problemas jurídicos de mas constante aparición y que han dado lugar a la literatura poniendo de relieve que el riesgo no sólo se da por la simple recolección de datos de carácter personal sino que también por su reutilización, combinación, profilado e inferencia mediante algo que podríamos denominar algoritmos de perfilado; Mendoza Enríquez (2021), traza las tensiones que pueden existir entre el funcionamiento de los sistemas de IA y los principios tradicionales finalidad, minimización y proporcionalidad, diciéndonos que los sistemas de IA suelen necesitar grandes volúmenes de información para su aprenda y su funcionamiento.

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

Los últimos estudios van más allá y trazan una mayor complejidad a esta problemática apuntando a que la IA puede llegar a inferir datos, es decir, la IA es capaz de generar datos inferidos, los cuales pueden mostrar comportamientos, preferencias y capacidades que jamás han sido entregados por el individuo/usuario/a, algo que queda señalado por Durand Goyzueta et al. (2026) mostrándonos que en este caso la problemática puede introducirse en las dimensiones cognitivas incluso en el caso de los espacios digitales donde el comportamiento del usuario es continuamente procesado.

Demetzou et al. (2023) señalan que las normas de protección de datos deben interpretarse no solo desde la legalidad del tratamiento, sino también desde las consecuencias que las decisiones automatizadas producen sobre las personas. En los trabajos más recientes se observa, por tanto, una transición desde la protección formal del dato hacia una protección orientada al impacto y al riesgo.

Transparencia, explicabilidad y derecho a explicación

La transparencia fue el eje que más se desarrolló de entre los estudios que se han revisado. Eso sí, los hallazgos también muestran que la transparencia cambió a lo largo de la investigación. Así, la transparencia se asoció a un principio expuesto en las revistas o artículos de divulgación en los que se informa sobre el funcionamiento de los algoritmos; posteriormente, la literatura empieza a cuestionar si el conocimiento sobre ese funcionamiento sirve de algo al usuario.

Burrell (2016) ya estudiaba distintas fuentes de opacidad, tales como la complejidad técnica y los límites de la interpretación de ciertos procedimientos computacionales, tal y como lo hicieron Ananny y Crawford (2018) al argumentar que visibilizar un sistema no implica necesariamente que haga visible, ni mucho menos cuestionable, el mismo.

A partir de este problema aparece la polémica de la explicabilidad. Wachter et al. (2017) cuestionan que el marco europeo reconozca un derecho general a recibir una explicación individualizada, mientras que Selbst y Powles (2017) defienden una interpretación más amplia de las obligaciones de información. Malgieri y Comandé (2017) hablan de la idea de legibilidad, diciendo que no es suficiente dar información técnica sino que debe permitir entender el proceso y sus consecuencias.

Los trabajos más recientes añaden además la auditabilidad como una prolongación de la transparencia. Según García-Llorente y Olmeda (2026), la inteligibilidad se asocia

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

con la capacidad de poder examinar técnicamente un sistema y comprobar si se puede justificar su comportamiento. La evolución identificada podría resumirse como una evolución desde la simple transparencia hacia la explicabilidad, la legibilidad y la auditabilidad.

Tabla 2.

Principales desafíos jurídicos identificados en la literatura

Dimensión	Desafío principal	Consecuencia jurídica
Protección de datos	Uso y reutilización intensiva de información	Pérdida de control sobre la finalidad del tratamiento
Datos inferidos	Generación automática de nueva información personal	Dificultad para conocer y ejercer derechos sobre las inferencias
Opacidad	Complejidad de modelos y procesos	Imposibilidad de comprender decisiones
Explicabilidad	Información técnica poco accesible	Limitación del derecho a cuestionar resultados
Decisiones automatizadas	Intervención humana reducida	Riesgo de afectación de derechos sin revisión efectiva
Discriminación	Sesgos en datos y modelos	Reproducción de desigualdades
Responsabilidad	Participación de múltiples actores	Dificultad para atribuir consecuencias jurídicas
Compliance	Controles internos insuficientes	Falta de prevención específica de riesgos algorítmicos
Gobernanza	Principios sin mecanismos efectivos	Brecha entre regulación y aplicación práctica

Nota. Las dimensiones fueron sintetizadas a partir de los problemas recurrentes identificados en los 23 estudios revisados.

Accountability, compliance y responsabilidad jurídica

Otro hallazgo relevante es el desplazamiento desde una concepción exclusivamente reactiva de responsabilidad hacia modelos preventivos. La literatura revisada muestra que no basta

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

con identificar quién debe responder después de ocurrido un daño; también es necesario establecer quién autoriza el uso del sistema, quién supervisa su funcionamiento y quién debe intervenir cuando se detectan resultados anómalos.

Busuioc (2021) explica que la rendición de cuentas o accountability requiere estructuras bien definidas de responsabilidad entre los desarrolladores, las organizaciones usuarias y las autoridades de control. Novelli et al. (2024) comparten la misma visión que el accountability, para ellos, no sólo es determinar al responsable; sino que para ser accountability hay que garantizar que se está en una posición para dar explicaciones, revisar y corregir.

Desde esta perspectiva se demuestra la conexión con el compliance. El término compliance algorítmico aún no aparece de forma homogénea en toda la literatura, aunque toda la serie de mecanismos propuestos responden a la lógica del compliance: identificación de los riesgos, documentación, controles preventivos, seguimiento, trazabilidad, y evaluación continua.

Mantelero (2018) hace hincapié en la utilidad de las evaluaciones de impacto como medio para la anticipación de las consecuencias jurídicas y sociales antes de la implementación de los sistemas de IA. Los descubrimientos demuestran que el compliance aplicado a la IA debe ir más allá de los controles, y hay que tener en cuenta los riesgos específicos, relacionados con la privacidad, la discriminación, la opacidad, la automatización.

Gobernanza y tendencias regulatorias

Las investigaciones relativas a la gobernanza dejan ver que hay una destacable confluencia en cuanto a los principios de privacidad, transparencia, igualdad, responsabilidad y control. Jobin et al. (2019) evidencian esa confluencia de los principios en los marcos internacionales; aunque también indican que hay una confluencia importante cuando se habla de su aplicación. Por su parte, Floridi et al. (2018) dirán que esos principios son dependientes de estructuras institucionales que puedan traducir orientaciones éticas en prácticas que sean verificables; en literatura más reciente, esta urgencia ha podido concretarse en mecanismos de auditoría, evaluación de impacto y supervisión humana.

Respecto a Latinoamérica, las investigaciones revisadas cierran el círculo dedicando la mayor parte de sus esfuerzos al estudio de la relación entre la IA y los derechos fundamentales (relacionados a las libertades fundamentales). Sánchez Acevedo (2022) llama la atención sobre que el uso de los sistemas automatizados por parte del Estado ha de estar siempre vigilado por el cumplimiento de la legalidad y la proporcionalidad. Gómez

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

Rodríguez (2026) indica que deberán desarrollarse sistemas de gobernanza que logren un equilibrio entre innovación, eficiencia administrativa y protección de derechos.

A partir de estos hallazgos se identificaron los principales instrumentos que pueden incorporarse en programas de compliance orientados al uso responsable de inteligencia artificial.

Tabla 3.

Mecanismos de respuesta jurídica y organizacional frente a los riesgos de la IA

Mecanismo	Finalidad	Aplicación dentro del compliance
Evaluación de impacto algorítmico	Detectar riesgos antes de implementar sistemas	Prevención y gestión anticipada
Auditoría algorítmica	Evaluar resultados, sesgos y funcionamiento	Verificación interna o independiente
Documentación de modelos	Registrar propósito, datos y limitaciones	Evidencia del cumplimiento
Trazabilidad	Reconstruir procesos y decisiones	Facilita control y atribución de responsabilidad
Supervisión humana	Revisar decisiones de alto impacto	Permite intervención y corrección
Explicabilidad	Facilitar comprensión de resultados	Favorece transparencia y ejercicio de derechos
Asignación de responsabilidades	Definir roles y obligaciones	Reduce vacíos de responsabilidad
Monitoreo continuo	Detectar cambios o desviaciones	Permite correcciones durante el ciclo de vida del sistema
Protocolos de reclamación	Facilitar cuestionamiento de decisiones	Refuerza tutela de derechos

Nota. Los mecanismos sintetizan las principales propuestas identificadas en los estudios revisados para transformar los principios jurídicos en controles organizacionales verificables.

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

Sumando todos los datos ya expuestos, los resultados evidencian la lenta transformación de la literatura científica. La protección de datos sigue siendo una condición necesaria, pero ya no se considera suficiente para hacer frente a los riesgos de la inteligencia artificial. La transparencia, por su parte, ha pasado a exigir niveles más sofisticados de explicabilidad y de auditabilidad, mientras que la responsabilidad jurídica se articula en mecanismos de supervisión y gobernanza cada vez más preventivos.

Se puede ejemplificar la tendencia general observada con la siguiente secuencia de desarrollo conceptual: protección de datos, transparencia, explicabilidad, accountability, gobernanza y compliance algorítmico. Parece claro que el reto jurídico contemporáneo no es tanto regular el tratamiento de la información, ni siquiera demostrar el funcionamiento de un algoritmo, sino crear estructuras capaces de prevenir riesgos, documentar decisiones, supervisar resultados, corregir desvíos y establecer responsabilidades en todo el ciclo de vida de los sistemas de inteligencia artificial.

4. DISCUSIÓN

Los resultados indagan si el punto de partida del debate jurídico en torno a la inteligencia artificial y su control es la protección de datos y concluyen que sí, aunque ya no es suficiente también para explicar los riesgos de los sistemas automatizados. Así también la literatura revisada establece que el problema no radica únicamente en la materia objeto de la recopilación de información sino también en la reutilización, combinación y formación de nuevas inferencias. Esto se encuentra en consonancia con el artículo de Mendoza Enríquez en 2021, cuando resalta que los principios de la finalidad, minimización y proporcionalidad presentan nuevas tensiones cuando la inteligencia artificial propone sistemas que requieren grandes volúmenes de datos.

Esta interpretación se refuerza con los estudios sobre privacidad e inferencias algorítmicas. Durand Goyzueta et al. (2026) muestran que los sistemas digitales pueden generar perfiles y conclusiones sobre los usuarios que no fueron proporcionados de manera directa. Los resultados de esta revisión permiten inferir que este tipo de información representa uno de los principales vacíos de los modelos tradicionales de protección, debido a que el individuo puede desconocer tanto la existencia de la inferencia como el uso posterior que se realiza de

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

ella. Por tanto, una regulación efectiva de la IA debería considerar no solo los datos de entrada, sino también los resultados, perfiles y predicciones generados a partir de ellos.

En cuanto a transparencia, los resultados confirman que la divulgación de información sobre un algoritmo no asegura que una decisión pueda ser comprendida o cuestionada. Esta conclusión está de acuerdo con Ananny y Crawford (2018), quienes señalan las limitaciones de la transparencia como solución suficiente a la opacidad. La revisión permitió ver que el debate ha evolucionado a exigencias de explicabilidad, legibilidad y auditabilidad. Ese es un cambio importante. Ahora no debería tratarse de conocer todos los detalles técnicos de un sistema, sino de disponer de información suficiente para entender cómo una decisión puede afectar a una persona y cuáles son los mecanismos para recurrirla.

Las diferencias entre Wachter et al. (2017) y Selbst y Powles (2017) reflejan esta controversia. La primera, en términos de la conciliación de la defensa de un derecho al olvido, y la segunda al afirmar que existe un derecho general a explicación de sus decisiones. Los hallazgos dejan entrever que ambas pueden ser superadas mediante la adopción de un criterio funcional: una explicación es adecuada, en tanto en cuanto regule la identificación de los factores relevantes a la decisión, permite comprender las consecuencias de la misma y que permite la activación de mecanismos de revisión. Desde esta óptica, el valor de la explicación es mucho más alto que el de la cantidad de información técnica.

Otro de los hitos relevantes ha sido que la responsabilidad de los algoritmos avanza desde un enfoque reactivo para irrumpir en un modelo previo. Busuioc (2021) sostiene que la responsabilidad algorítmica necesita identificar a aquellos que son responsables y la necesidad de incluir alguna forma de supervisión de la actividad de la máquina. Eso bien se integra con nuestros resultados - la trazabilidad, la documentación y la delineación de roles aparecen en ese sentido como mecanismos recurrentes.

En este caso, el compliance pasa a ocupar un lugar central. Novelli et al. (2024) conceptualizan la accountability como una estructura compuesta por obligaciones, justificaciones, controles y medidas en caso de que los datos personales no hayan sido procesados de la manera adecuada. Los resultados de esta revisión permiten trasladar esta lógica al terreno organizacional: el compliance algorítmico puede interpretarse como un sistema de control capaz de integrar la protección de datos, la evaluación de impacto, las auditorías, la trazabilidad y el control humano. La función del compliance

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

algorítmico no consistiría únicamente en hacer posible el cumplimiento formal sino en proporcionar prueba de que los riesgos de la IA fueron identificados y gestionados durante todo el ciclo de vida de la elaboración del sistema. Esta conclusión se vincula con la propuesta de Mantelero (2018), la cual sugiere que las evaluaciones de impacto deban extenderse hacia las dimensiones jurídicas, sociales y éticas. De la misma manera, los resultados evidencian que el uso de este tipo de herramientas consiste, sin lugar a dudas, en una de las maneras más eficaces de tratar las afectaciones de antemano.

Se evidencia con la revisión que existe una distancia entre la validez de los principios generales y su aplicación real. Jobin et al. (2019) encontraron una elevada coincidencia internacional sobre la transparencia, privacidad, la justicia, la responsabilidad, la explicación, etc; pero los contenidos de este estudio muestran que el problema actual reside sobre todo en la forma de hacer efectiva esos principios, es decir, en la forma de convertir aquellos principios en procedimientos comprobables. De esta forma, la evolución de la regulación jurídica de la IA puede sintetizarse como un paso de la protección de los datos a la gestión y compliance integrales.

Una respuesta jurídica más adecuada implica integrar la protección de datos, la explicabilidad, la auditoría, la supervisión por el humano y la rendición de cuentas todo dentro de estructuras organizativas de compliance. De este modo, se puede pasar de un modelo reactivo respecto al daño, a uno en el que la información se procese para su prevención, documentación y despeje antes de llegar a tener consecuencias jurídicas relevantes.

5. CONCLUSIÓN

La observación indica que la protección de datos es una pieza esencial de la regulación de la inteligencia artificial, aunque poco se puede sostener cuando se aborda de manera estanca. Los principales peligros también se producen en la producción de inferencias, el perfilamiento, el uso de información derivada que pone en riesgo a las personas sin redescubrir cabalmente de dónde provienen o cómo se utiliza.

La transparencia algorítmica es un prerrequisito aunque en modo alguno se presenta una fuente de eficacia en la protección de los derechos. Los artículos revisados muestran una evolución hacia conceptos con mayores exigencias (explicabilidad, legibilidad y

Doi: https://doi.org/10.63688/b35xe061

ISSN: 3103-1307
Vol. 4 Num. 1 (2026)

auditabilidad) con el fin de que las decisiones automatizadas puedan ser entendidas, cuestionadas y revisadas.

La responsabilidad jurídica en relación con los sistemas de IA ha de abordarse con un sentido preventivo y no sólo de forma posterior al daño que producen. La identificación del responsable, la trazabilidad de las decisiones, la documentación de los sistemas y la verificación humana constituyen mecanismos clave para lograr una reducción de los vacíos de responsabilidad y de incrementar las posibilidades de accountability.

El compliance algorítmico aparece como una medida que articula protección de datos, transparencia, rendición de cuentas y gobernanza en una misma estructura organizacional. Dicho enfoque necesita ser acompañado de evaluaciones de impacto, auditorías, controles internos, monitorización continua y procesos de verificación que permita minimizar riesgos durante todo el ciclo de vida de los sistemas de IA.

La principal tendencia de la literatura es la transición desde modelos de cumplimiento hacia modelos de gestión global de la IA, por lo que el desafío jurídico para el futuro no se centrará tanto en la simple formulación de principios éticos o normativos, sino en la de convertirlos en mecanismos sometidos a verificación que permitan dar cuenta de que los sistemas de inteligencia artificial funcionan o son utilizados de forma legal, transparente, con verificación y con responsabilidad.

REFERENCIAS BIBLIOGRÁFICAS

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society, 20*(3), 973–989. https://doi.org/10.1177/1461444816676645

Arnesen, L. (2026). Comparing the right to information about automated decision-making to the new right to explanation. *International Data Privacy Law, 16*(2), ipag008. https://doi.org/10.1093/idpl/ipag008

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society, 3*(1), 1–12. https://doi.org/10.1177/2053951715622512

Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review, 81*(5), 825–836. https://doi.org/10.1111/puar.13293

Cárdenas Poveda, M., & Sarmiento Salcedo, M. (2025). Principios del gobierno digital, marco ético de la inteligencia artificial y función administrativa en Colombia. *Pro Jure Revista de Derecho, 64*, 305–332. https://doi.org/10.4151/SO2810-76592025064-1458

Demetzou, K., Zanfir-Fortuna, G., & Vale, S. B. (2023). The thin red line: Refocusing data protection law on ADM, a global perspective with lessons from case-law. *Computer Law & Security Review, 49*, 105806. https://doi.org/10.1016/j.clsr.2023.105806

Durand Goyzueta, E. A., Condori Mamani, B., Ramos Chura, E., Avila Inquilla, V. E., Salinas del Carpio, A., Ito Díaz, R. R., & Hilasaca Zea, C. Y. (2026). Cognitive privacy and the use of artificial intelligence in digital platforms for higher education. *NeuroData, 3*, 160. https://doi.org/10.63688/neurodata2026160

Felzmann, H., Fosch-Villaronga, E., Lutz, C., & Tamò-Larrieux, A. (2019). Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. *Big Data & Society, 6*(1), 1–14. https://doi.org/10.1177/2053951719860542

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

García-Llorente, C., & Olmeda, I. (2026). Making algorithmic transparency enforceable: Sufficient intelligibility, technical evidence, and auditability across the EU and the US. *International Journal of Law and Information Technology, 34*, eaag008. https://doi.org/10.1093/ijlit/eaag008

Gómez Rodríguez, J. M. (2026). Hacia la gobernanza de la inteligencia artificial en la Administración pública. *Foro: Revista de Derecho, 45*, 29–47. https://doi.org/10.32719/26312484.2026.45.2

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2018). Fair, transparent, and accountable algorithmic decision-making processes: The premise, the proposed solutions, and the open challenges. *Philosophy & Technology, 31*(4), 611–627. https://doi.org/10.1007/s13347-017-0279-x

Li, W., Zhang, Y., Zheng, Q., & Li, A. (2026). How the legal basis for AI training is framed in data protection guidelines and interventions: Comparative perspectives and the prospect of global convergence. *International Data Privacy Law, 16*(2), ipaf032. https://doi.org/10.1093/idpl/ipaf032

Malgieri, G., & Comandé, G. (2017). Why a right to legibility of automated decision-making exists in the General Data Protection Regulation. *International Data Privacy Law, 7*(4), 243–265. https://doi.org/10.1093/idpl/ipx019

Mantelero, A. (2018). AI and Big Data: A blueprint for a human rights, social and ethical impact assessment. *Computer Law & Security Review, 34*(4), 754–772. https://doi.org/10.1016/j.clsr.2018.05.017

Mendoza Enríquez, O. A. (2021). El derecho de protección de datos personales en los sistemas de inteligencia artificial. *Revista IUS, 15*(48), 179–207. https://doi.org/10.35487/rius.v15i48.2021.743

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 1–21. https://doi.org/10.1177/2053951716679679

Novelli, C., Taddeo, M., & Floridi, L. (2024). Accountability in artificial intelligence: What it is and how it works. *AI & Society, 39*(4), 1871–1882. https://doi.org/10.1007/s00146-023-01635-y

Sánchez Acevedo, M. E. (2022). La inteligencia artificial en el sector público y su límite respecto de los derechos fundamentales. *Estudios Constitucionales, 20*(2), 257–284. https://doi.org/10.4067/S0718-52002022000200257

Selbst, A. D., & Powles, J. (2017). Meaningful information and the right to explanation. *International Data Privacy Law, 7*(4), 233–242. https://doi.org/10.1093/idpl/ipx022

Viollier, P., & Contreras, P. (2026). Derecho a una explicación en decisiones automatizadas: Contrapunto entre el GDPR y la Ley 21.719. *Revista Chilena de Derecho y Ciencia Política, 17*(1), 1–27. https://doi.org/10.7770/rchdcp-v17n1-art527

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. *International Data Privacy Law, 7*(2), 76–99. https://doi.org/10.1093/idpl/ipx005

CONFLICTO DE INTERESES

No existe.

FINANCIAMIENTO

Ninguno.

CONTRIBUCIÓN DE AUTORÍA

Conceptualización: Francisco Javier Santini Rodriguez, Luis Carlos Rodriguez Montaño y Mario Isaac Romero Lamadrid.
Metodología: Francisco Javier Santini Rodriguez y Luis Carlos Rodriguez Montaño.
Investigación: Francisco Javier Santini Rodriguez, Luis Carlos Rodriguez Montaño y Mario Isaac Romero Lamadrid.
Redacción – borrador original: Francisco Javier Santini Rodriguez y Mario Isaac Romero Lamadrid.
Redacción – revisión y edición: Francisco Javier Santini Rodriguez, Luis Carlos Rodriguez Montaño y Mario Isaac Romero Lamadrid.