



Facultad de Ciencias Económicas y Empresariales

Implicaciones éticas y legales del uso de la inteligencia artificial en la gestión de la protección de datos personales en las empresas

Autor: Mónica de Miñón García

Director: Víctor Pérez Segura

MADRID | Abril 2026

Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado

ADVERTENCIA: Desde la Universidad consideramos que ChatGPT u otras herramientas similares son herramientas muy útiles en la vida académica, aunque su uso queda siempre bajo la responsabilidad del alumno, puesto que las respuestas que proporciona pueden no ser veraces. En este sentido, NO está permitido su uso en la elaboración del Trabajo fin de Grado para generar código porque estas herramientas no son fiables en esa tarea. Aunque el código funcione, no hay garantías de que metodológicamente sea correcto, y es altamente probable que no lo sea.

Por la presente, yo, Mónica de Miñón García, estudiante de E3-Analytics de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado “Implicaciones éticas y legales del uso de la inteligencia artificial en la gestión de la protección de datos personales en las empresas” declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación:

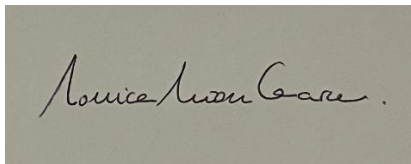
1. **Brainstorming de ideas de investigación:** Utilizado para idear y esbozar posibles áreas de investigación.
2. **Referencias:** Usado conjuntamente con otras herramientas, como Science, para identificar referencias preliminares que luego he contrastado y validado.
3. **Metodólogo:** Para descubrir métodos aplicables a problemas específicos de investigación.
4. **Interpretador de código:** Para realizar análisis de datos preliminares.
5. **Constructor de plantillas:** Para diseñar formatos específicos para secciones del trabajo.
6. **Corrector de estilo literario y de lenguaje:** Para mejorar la calidad lingüística y estilística del texto.
7. **Generador previo de diagramas de flujo y contenido:** Para esbozar diagramas iniciales.
8. **Sintetizador y divulgador de libros complicados:** Para resumir y comprender literatura compleja.

9. **Revisor:** Para recibir sugerencias sobre cómo mejorar y perfeccionar el trabajo con diferentes niveles de exigencia.
10. **Traductor:** Para traducir textos de un lenguaje a otro.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: 23/04/2026

Firma: Mónica de Miñón García.

A rectangular box containing a handwritten signature in black ink. The signature is written in a cursive style and reads "Mónica de Miñón García".

INDICE

1. INTRODUCCIÓN	8
1.1. Relevancia y contexto del estudio de investigación	8
1.2. Problema de investigación y preguntas de investigación	9
1.3. Objetivos generales y específicos del estudio	10
1.4. Metodología.....	10
1.5. Estructura del trabajo	13
2. MARCO CONCEPTUAL.....	14
2.1. Concepto de inteligencia artificial	14
2.2. La ética en las empresas y la responsabilidad corporativa.....	15
2.3. El valor de los datos en la empresa	16
2.4. El ciclo de cambio de los datos en la empresa	17
2.5. La protección de datos personales en la empresa y sus desafíos en el entorno digital: cómo se relaciona la privacidad con el uso de IA.....	17
3. FUNDAMENTOS ÉTICOS DE LA IA Y LA PROTECCIÓN DE DATOS	19
3.1. Conceptos clave: transparencia algorítmica y privacidad.....	19
3.2. Principios éticos aplicados a la IA	19
3.3. El papel de la ética empresarial en el uso de la IA	20
3.4. La gobernanza ética de la IA en el contexto internacional.	21
4. EL DILEMA ÉTICO DEL USO DE LA IA EN LA GESTIÓN DE LOS DATOS PERSONALES EN LAS EMPRESAS	22
4.1. El uso de datos masivos y sus implicaciones éticas.....	22
4.2. Sesgos algorítmicos y discriminación automatizada.....	22
4.3. La caja negra: falta de transparencia y dificultad para atribuir responsabilidades	23
4.4. Pérdida de privacidad y vigilancia corporativa	24
4.5. Responsabilidad moral en la toma de decisiones automatizada	24
5. RESULTADOS	25
5.1. Caracterización bibliométrica del corpus	25
5.2. Análisis de tendencias temáticas.....	30
5.3. Análisis temático mediante topic modelling.....	30
5.4. Principales conclusiones del análisis de contenido.....	34
6. GOBERNANZA Y BUENAS PRÁCTICAS ÉTICAS EN LAS EMPRESAS	36
6.1. Modelos de gobernanza de datos con IA	36
6.2. Herramientas de auditoría ética y transparencia algorítmica	36
6.3. Casos reales de empresas tecnológicas	37
7. PROPUESTAS FUTURAS.....	39
7.1. Recomendaciones para la gestión ética de los datos personales con IA	39

7.2. Tendencias emergentes para la gobernanza ética empresarial: hacia una IA responsable	39
7.3. ¿Equilibrio entre privacidad, innovación y derechos humanos?.....	40
8. CONCLUSIONES	42
8.1. Respuesta a las preguntas de investigación: principales hallazgos	42
8.2. Limitaciones del estudio y futuras líneas de trabajo	43
BIBLIOGRAFÍA.....	45
Literatura científica.....	45
Normativa y fuentes oficiales	47
Otras fuentes	48

RESUMEN

El presente trabajo trata de analizar las implicaciones éticas y legales del uso de la inteligencia artificial en la gestión de datos personales en el ámbito empresarial. Es un campo de creciente relevancia ante la irrupción de los sistemas de IA en los últimos años y las recientes legislaciones del marco normativo europeo, representado por el RGPD y el AI Act.

La investigación del trabajo se ha desarrollado mediante una revisión bibliográfica sistemática basada en 146 artículos científicos extraídos de la base de datos Scopus. Estos se han seleccionados a través de una ecuación de búsqueda que combina los conceptos principales a tratar: inteligencia artificial, privacidad de datos, ética y empresa. El análisis combina una fase bibliométrica, realizada con Python, y una fase de análisis de contenido cualitativo de los artículos más relevantes. Adicionalmente, se aplica un modelo de topic modelling sobre los abstracts del corpus, con el objetivo de identificar los bloques temáticos que predominan y preocupan en la literatura científica.

Los resultados demuestran que la documentación científica sobre este tema ha experimentado un crecimiento muy significativo en el período que ha sido analizado, especialmente a partir de 2023. El análisis temático identifica seis bloques principales: gobernanza y regulación, sesgos algorítmicos y discriminación, transparencia y explicabilidad, IA generativa, privacidad y protección de datos, y ética empresarial. Los hallazgos del análisis de contenido revelan que el conflicto entre la lógica de maximización del valor de los datos y los principios de protección de derechos es de naturaleza estructural, y que los dilemas más graves son la discriminación algorítmica, la opacidad de los sistemas y el profiling predictivo. Asimismo, la evidencia empírica disponible demuestra que las empresas que cuentan con un compromiso organizativo formal con la ética en el uso de la IA presentan un mayor nivel de cumplimiento normativo.

El trabajo concluye con un conjunto de recomendaciones para avanzar hacia una gobernanza responsable de la IA en el contexto empresarial.

ABSTRACT

The present study analyses the ethical and legal implications of the use of artificial intelligence in the management of personal data in the business environment. This is a field of growing relevance given the rise of AI systems in recent years and the recent legislative developments in the European regulatory framework, represented by the GDPR and the AI Act.

The research has been conducted through a systematic literature review based on 146 scientific articles extracted from the Scopus database. These were selected through a search equation combining the main concepts addressed: artificial intelligence, data privacy, ethics and business. The analysis combines a bibliometric phase, carried out using Python, and a qualitative content analysis phase of the most relevant articles. Additionally, a topic modelling approach was applied to the abstracts of the corpus, with the aim of identifying the thematic clusters that predominate and raise concern in scientific literature.

The results show that scientific output on this topic has grown significantly over the period analyzed, particularly from 2023 onwards. The thematic analysis identifies six main clusters: governance and regulation, algorithmic bias and discrimination, transparency and explainability, generative AI, privacy and data protection, and business ethics. The findings of the content analysis reveal that the conflict between the logic of data value maximization and the principles of rights protection is structural in nature, and that the most serious ethical dilemmas are algorithmic discrimination, system opacity and predictive profiling. Furthermore, the available empirical evidence demonstrates that companies with a formal organizational commitment to ethics in the use of AI show higher levels of regulatory compliance.

The research concludes with a set of recommendations for advancing towards responsible AI governance in the business context.

1. INTRODUCCIÓN

1.1. Relevancia y contexto del estudio de investigación

En las últimas décadas, la inteligencia artificial ((en adelante, IA) ha pasado de ser un campo de investigación académica a convertirse en una tecnología transversal que actúa prácticamente en todos los sectores de la actividad empresarial. Su capacidad para procesar volúmenes masivos de datos y automatizar procesos de toma de decisiones la ha situado en el centro de las estrategias de transformación digital de las organizaciones a escala global. A pesar de esto, este avance tecnológico no ha venido exento de tensiones y dilemas profundos, sobre todo en lo que respecta al tratamiento de los datos personales y al respeto de los derechos fundamentales de las personas.

La irrupción de la IA en el ámbito empresarial ha generado una nueva dimensión en la gestión de la información: las organizaciones no solo recopilan y almacenan datos de sus clientes, empleados y proveedores, sino que los utilizan para alimentar algoritmos que toman decisiones con consecuencias directas sobre las personas, como por ejemplo la selección de personal o la personalización de precios. Estos son solo algunos ejemplos de cómo los sistemas de IA median en relaciones que hasta hace poco eran exclusivamente humanas (Kriebitz y Lütge, 2020; Gerlick y Liozu, 2020).

Este contexto ha propiciado un creciente interés académico y regulatorio por comprender y delimitar las implicaciones éticas y legales de estas prácticas. En el ámbito europeo, la aprobación del Reglamento General de Protección de Datos (en adelante, RGPD) en 2016 supuso un hito fundamental que por primera vez establecía un marco jurídico para la protección de los datos personales. Más recientemente, la entrada en vigor del Reglamento de Inteligencia Artificial de la Unión Europea (AI Act) ha reforzado este marco regulatorio al introducir requisitos específicos de transparencia y supervisión humana para los sistemas de IA de alto riesgo. A pesar de los avances que se dan, persisten importantes lagunas entre la regulación existente y las prácticas reales que toman las empresas, lo que justifica la necesidad de investigación académica que profundice en esta problemática (Milossi et al., 2021; Meszaros y Ho, 2021).

La relevancia de este estudio se fundamenta así en la coexistencia de tres fenómenos que definen el panorama empresarial de nuestros tiempos: la aceleración en la adopción de sistemas de IA por parte de las organizaciones, el incremento exponencial en el volumen y la sensibilidad de los datos personales tratados, y la creciente presión regulatoria para

garantizar que el uso de estas tecnologías respete los principios éticos y los derechos de las personas. De esta forma, este trabajo se conecta con la línea de investigación que busca entrelazar la ética empresarial, la gobernanza de datos y los sistemas de Inteligencia Artificial avanzados. (Vidgen et al., 2020; Wu et al., 2020).

1.2. Problema de investigación y preguntas de investigación

El problema central que aborda este trabajo es la existencia de un dilema ético en el uso que hacen las empresas de la inteligencia artificial para el tratamiento de datos personales. Este dilema surge de la tensión entre dos lógicas; la de la eficiencia y la competitividad empresarial, que impulsa a las organizaciones a extraer el máximo valor de los datos; y la de los derechos y la dignidad de las personas, que exige este tratamiento respete la privacidad, la autonomía y la no discriminación.

Esta tensión se manifiesta en la práctica a través de una serie de fenómenos que aparecen documentados en la literatura científica: la opacidad algorítmica o caja negra, que dificulta la detección de sesgos y la atribución de responsabilidades (Kim y Routledge, 2022; Mühlhoff, 2023); el riesgo de discriminación algorítmica que proviene del entrenamiento con datos históricos sesgados (Rizinski et al., 2022); y la evidencia empírica de que las organizaciones se involucran sistemáticamente en prácticas contrarias a la normativa de protección de datos por ausencia de mecanismos internos de gobernanza ética (Mendez-Suarez et al., 2024).

A partir de este debate, se formulan las siguientes preguntas de investigación que orientan el análisis desarrollado en este trabajo:

Pregunta de investigación principal:

1. ¿Cuáles son las principales implicaciones éticas que surgen del uso de la inteligencia artificial en la gestión de datos personales por parte de las empresas, y de qué manera pueden estas mantener una gobernanza responsable que garantice su protección efectiva?

Preguntas específicas:

1. ¿Cuáles son los dilemas éticos específicos que enfrentan las empresas en relación con el uso de la IA cuando tratan con datos personales, con especial atención a la caja negra, los sesgos algorítmicos y la vigilancia corporativa?

2. *¿Qué herramientas y estrategias de data governance pueden utilizar las empresas para reducir los riesgos éticos en el uso de la IA con datos personales?*
3. *¿Qué recomendaciones pueden formularse para fortalecer la transparencia en la gestión empresarial de datos personales con IA y avanzar hacia una gobernanza ética que respete los derechos humanos?*

1.3. Objetivos generales y específicos del estudio

El presente trabajo persigue dos objetivos generales complementarios. El primero consiste en analizar las implicaciones éticas que se derivan del uso de la IA en el tratamiento de los datos personales dentro de las empresas, pudiendo identificar los dilemas principales y el riesgo de afectación a la privacidad. El segundo objetivo general busca examinar la forma en que las empresas pueden aplicar marcos de gobernanza responsable, integrados con principios éticos, para garantizar una protección efectiva de los datos personales respetando los derechos de las personas vinculadas con ellos.

De estos objetivos generales se derivan tres objetivos específicos. El primero es identificar y analizar el impacto de los dilemas éticos que enfrentan las empresas cuando tratan datos personales con IA. El segundo objetivo específico es evaluar las estrategias de gobernanza del dato y sus mecanismos para frenar los riesgos éticos en el uso de la IA con datos personales. El tercer objetivo específico consiste en proponer las mejores prácticas que pueden adoptar las empresas para fortalecer la confianza digital en su relación con los datos de las personas.

1.4. Metodología

El presente trabajo adopta un enfoque metodológico de investigación documental, de carácter analítico, cualitativo y descriptivo. Este enfoque permite sistematizar y sintetizar el conocimiento académico existente sobre las implicaciones éticas del uso de la IA en el tratamiento de datos personales en las empresas. Sin duda, es un campo que está en creciente evolución y su comprensión requiere integrar perspectivas procedentes de la ética empresarial, el derecho, la ciencia de datos y la gestión organizacional.

La búsqueda bibliográfica se llevó a cabo en la base de datos Scopus (Elsevier, 2026). Se ha utilizado Scopus por su amplitud de cobertura en las áreas temáticas que son relevantes

para este trabajo, si bien esta elección constituye una limitación que se discute en el apartado de conclusiones.

La selección de las palabras clave de la ecuación de búsqueda contiene los conceptos base del objeto de estudio. Se identificaron cuatro ejes temáticos: (1) la tecnología objeto de análisis ("artificial intelligence" OR AI); (2) el ámbito de los datos ("data privacy" OR "data protection"); (3) la dimensión ética (ethics OR ethical); y (4) el contexto organizativo (business OR enterprise OR company OR corporate). La combinación de estos cuatro ejes mediante el operador "AND" garantiza que los documentos recuperados traten simultáneamente todos los aspectos relevantes para el objeto de estudio. La ecuación completa es la siguiente:

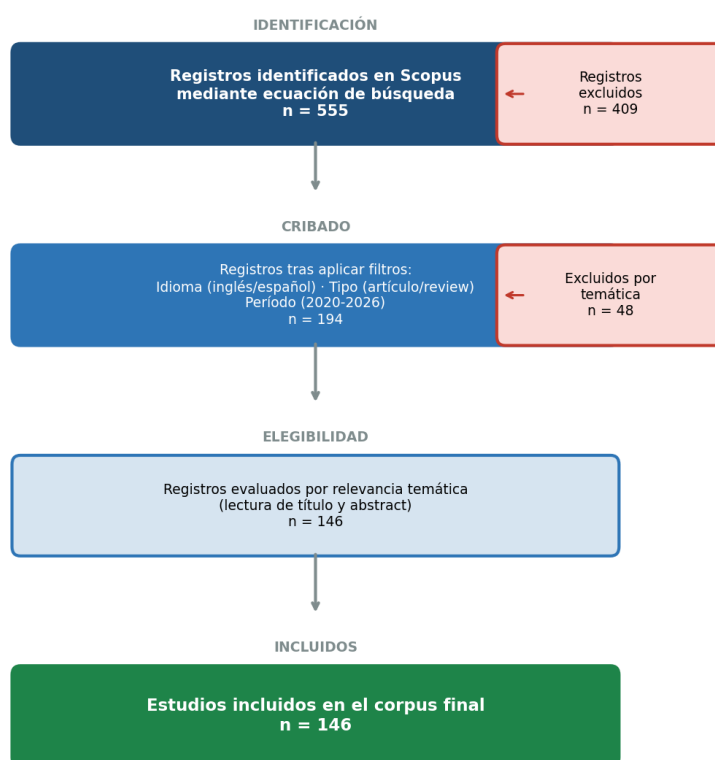
("artificial intelligence" OR AI) AND ("data privacy" OR "data protection") AND (ethics OR ethical) AND (business OR enterprise OR company OR corporate)

Esta búsqueda se aplicó sobre los campos de título, resumen y palabras clave de los documentos indexados en Scopus, obteniendo un total de 555 registros iniciales.

Para delimitar el corpus inicial se aplicaron los siguientes criterios de inclusión: (1) período de publicación comprendido entre 2020 y 2026, con el fin de capturar las tendencias más recientes y garantizar actualidad en las conclusiones del trabajo; (2) documentos redactados en inglés o español; (3) tipo de documento enfocado a artículos científicos originales (Article) y revisiones sistemáticas (Review) publicados en revistas revisadas por pares.

En cuanto a los criterios de exclusión: (1) documentos cuyo abstract no estuviera disponible en la base de datos; (2) artículos cuya temática principal no guardara relación directa con la intersección entre inteligencia artificial, ética empresarial y protección de datos personales; (3) duplicados detectados durante el proceso de revisión. La aplicación de estos criterios redujo el corpus de 555 a 194 registros tras los filtros automáticos, y a 146 documentos tras la revisión manual de relevancia temática. El proceso completo se representa en la Figura 1.

Figura 1. Diagrama de flujo del proceso de selección del corpus (PRISMA)



Fuente: elaboración propia mediante Python a partir de datos de Scopus (2020-2026).

Se ha considerado adecuado el corpus final de 146 documentos porque la ecuación de búsqueda está acotada de forma deliberada para que haya una intersección de los cuatro ejes temáticos específicos. Además, el período de análisis 2020-2026 es relativamente reciente y esto limita el volumen disponible, pero maximiza la actualidad de las conclusiones del objeto de estudio. Y, por último, el análisis de contenido cualitativo de los artículos más citados permite extraer conclusiones sólidas sin necesidad de procesar un volumen mayor de documentos con menos relevancia. Futuras investigaciones podrían ampliar el corpus incorporando Web of Science como fuente complementaria o probando queries alternativas.

El análisis del corpus se realizó en dos fases complementarias. En la primera fase se llevó a cabo un análisis bibliométrico cuantitativo mediante Python, con el objetivo de caracterizar la producción científica seleccionada: evolución temporal de las publicaciones, revistas con mayor presencia, artículos con mayor impacto en términos de citas recibidas, distribución por tipo de documento y análisis de palabras clave (los resultados de este análisis se presentan en el apartado de Resultados). En la segunda fase se realizó un análisis de contenido cualitativo de los documentos más relevantes, orientado a identificar los principales dilemas éticos documentados, los marcos teóricos utilizados, las propuestas de

gobernanza existentes y las tendencias emergentes en la materia. Este análisis constituye la base de los capítulos tres, cuatro, seis y siete.

1.5. Estructura del trabajo

El presente trabajo se estructura en ocho capítulos. Tras esta introducción, el segundo capítulo desarrolla el marco conceptual del estudio. El tercer capítulo analiza los fundamentos éticos aplicados a la IA. El cuarto capítulo constituye el núcleo analítico del trabajo y examina los principales dilemas éticos que emergen del uso de la IA en la gestión de datos personales en las empresas. El quinto capítulo muestra los resultados del trabajo. El quinto capítulo aborda la gobernanza y las buenas prácticas éticas empresariales. El sexto capítulo presenta propuestas de futuro y recomendaciones para avanzar hacia una IA más responsable en el ámbito empresarial. El séptimo, muestra los resultados y finalmente, el séptimo capítulo recoge las conclusiones del trabajo, dando respuesta a las preguntas de investigación planteadas e identificando las limitaciones del estudio.

2. MARCO CONCEPTUAL

2.1. Concepto de inteligencia artificial

La inteligencia artificial puede definirse, según publicó el High Level Group on Artificial Intelligence de la Comisión Europea el 18 de diciembre de 2018: La inteligencia artificial (IA) se refiere a sistemas diseñados por humanos que, ante un objetivo complejo, actúan en el mundo físico o digital percibiendo su entorno, interpretando los datos estructurados o no estructurados recopilados, razonando sobre el conocimiento derivado de estos datos y decidiendo las mejores acciones —de acuerdo con parámetros predefinidos— para lograr el objetivo dado. Los sistemas de IA pueden utilizar reglas simbólicas o aprender un modelo numérico y también pueden adaptar su comportamiento analizando cómo el entorno se ve afectado por sus acciones anteriores. (RAE, s.f.). Se podría resumir en un conjunto de técnicas y sistemas diseñados para realizar tareas que requerirían mayores capacidades cognitivas si fuesen ejecutadas por los humanos. Esta definición, sin embargo, engloba una familia de tecnologías diferentes y es importante distinguirlas para comprender cuales son las implicaciones éticas que pueden tener en el ámbito empresarial.

Desde una perspectiva técnica, los sistemas de IA pueden clasificarse en función de su arquitectura y capacidad. Se van a explicar los tres tipos principales. En primer lugar, el aprendizaje automático (machine learning) constituye la rama más extendida en aplicaciones empresariales: estos programas aprenden patrones a partir de grandes volúmenes de datos de entrenamiento, sin que nadie les programe explícitamente las reglas, y así utilizarlos para hacer predicciones o tomar decisiones sobre datos nuevos. En segundo lugar, y dentro del aprendizaje automático, el aprendizaje profundo (deep learning) es una versión más avanzada que emplea redes neuronales capaces de procesar datos no estructurados (imágenes, texto o audio) con una precisión que podría llegar a superar la humana en muchas ocasiones (Wu et al., 2020). En tercer lugar, y el tipo más recientemente, la IA generativa (representada por modelos como GPT, DALL-E o Gemini) ha ampliado las capacidades de estos sistemas notoriamente. Este último consigue incluir la creación de contenidos nuevos en respuesta a instrucciones en lenguaje de los humanos, lo que ha abierto grandes oportunidades, pero también ha desarrollado nuevos vectores de riesgo ético (Golda et al., 2024).

Existe un factor que caracteriza en gran manera a los sistemas de IA. A diferencia de otras tecnologías, esta no afecta únicamente a un sector concreto si no a todos al mismo tiempo,

definiéndola, así como una tecnología disruptiva de carácter transversal con capacidad de transformar simultáneamente múltiples industrias y funciones organizativas. Esta capacidad de vinculación generalizada es precisamente lo que hace que sus implicaciones éticas sean tan relevantes y complejas. (Kriebitz y Lütge, 2020). Por lo tanto, se trata de un conjunto de desafíos que afectan a la totalidad del mundo empresarial y, a través de él, a los derechos e intereses de todas de personas.

Es importante señalar que la IA no es una tecnología neutral. Los sistemas de IA son diseñados por personas y entrenados con datos generados por personas. Esta construcción social de la tecnología implica, inevitablemente, los sistemas de IA quedan incorporados por el resultado de los sesgos, las desigualdades y los valores de quienes diseñan y despliegan los sistemas. (Anshari et al., 2023). Este entendimiento constituye el punto de partida de cualquier análisis ético sobre el uso de la IA en el contexto empresarial.

2.2. La ética en las empresas y la responsabilidad corporativa

En el contexto del uso de IA en las empresas, la ética adquiere una dimensión muy relevante a través del concepto de responsabilidad corporativa. Kriebitz y Lütge (2020) lo argumentan es específicamente con su posición conocida como “beyond compliance”. Según esta, las empresas que desarrollan o utilizan sistemas de IA no pueden limitarse a cumplir con la ley vigente: tienen una responsabilidad moral que va más allá de eso, deben de garantizar que sus sistemas respeten los derechos humanos fundamentales, incluso cuando la ley no lo exige expresamente. Es decir, ser ético no significa no infringir el RGPD, implica no limitarse al mero cumplimiento normativo.

Gerlick y Liozu (2020) identifican tres dimensiones éticas concretas toda empresa debe de considerar al implementar algoritmos para la toma de decisiones: la justicia, entendida como la ausencia de discriminación; la transparencia, que implica que los afectados por la decisión del sistema IA puedan conocer y estén justificados los criterios que la determinaron; y la responsabilidad, que exige la existencia de mecanismos lo suficientemente claros como para poder identificar, y posteriormente corregir, los errores del sistema y tomar medidas con quienes se hayan visto perjudicados indebidamente por ellos.

Más recientemente, se ha introducido un nuevo concepto que supone una evolución de la responsabilidad social corporativa adaptada a la era digital: la Responsabilidad Digital Corporativa (CDR). Gursoy et al. (2025) definen la CDR como el compromiso de las

organizaciones de utilizar las tecnologías digitales, incluida la IA, de manera que genere valor no solo para los accionistas sino para el conjunto de la sociedad, respetando los derechos de las personas, protegiendo el medio ambiente y contribuyendo al bienestar colectivo. Este nuevo término facilita el análisis de las implicaciones éticas del uso de IA en datos personales en las empresas ya que conecta las decisiones de negocio con consecuencias sociales más amplias.

2.3. El valor de los datos en la empresa

El valor de los datos para las empresas puede tener distinta naturaleza y objetivo. Por un lado, utilizan para mejorar la eficiencia de estas: los sistemas de IA que estén entrenados con datos históricos pueden optimizar procesos de producción, logística, atención al cliente o gestión de recursos humanos con una precisión difícilmente alcanzable con otro método. Por otro lado, los datos tienen valor estratégico: permiten a las empresas comprender mejor a sus clientes, anticipar sus necesidades, personalizar ofertas y diseñar estrategias de precios y marketing más personales. (Truby y Brown, 2021).

Sin embargo, existe otro punto de vista que es problemática éticamente, donde el valor de los datos entra en tensión directa con los derechos de las personas. Mühlhoff (2023) ha denominado "privacidad predictiva" al conjunto de problemas éticos que emergen cuando las empresas utilizan datos anónimos o aparentemente no sensibles de grandes colectivos para inferir información íntima y sensible sobre individuos concretos. A través de técnicas de analítica predictiva, las empresas pueden inferir información íntima y sensible de una persona, a partir de datos que, considerados individualmente, parecen inocuos (por ejemplo, patrones de navegación o historial de compras). Este fenómeno implica que el partido que extraen las empresas de los datos puede generarse violando la privacidad y la autonomía de las personas, puesto que en muchas ocasiones estas no han facilitado conscientemente la información sensible.

La comprensión de esta dinámica es esencial para el análisis ético que desarrolla este trabajo, pues pone de manifiesto que el problema de la protección de datos en la era de la IA no se reduce a la custodia segura de información declarada como sensible, sino que abarca la totalidad del ecosistema de datos que las empresas gestionan, incluidos aquellos que a priori podrían parecer inofensivos.

2.4. El ciclo de cambio de los datos en la empresa

La primera fase del ciclo es la de captura o recopilación. Las empresas obtienen datos personales de acuerdo con tres formas. La primera es de forma directa, cuando el usuario proporciona sus datos conscientemente. La segunda de forma indirecta, cuando el usuario proporciona los datos indirectamente de forma que se obtienen los datos rastreando el comportamiento, por ejemplo, los movimientos en plataformas digitales o aplicaciones móviles. La tercera de forma inferida, cuando los datos los genera el propio sistema a partir del análisis de un comportamiento observado. Cada una de estas modalidades plantea interrogantes éticos distintos. (Wellbrock et al., 2025).

La segunda fase es la de almacenamiento y procesamiento. Los datos recopilados son almacenados, en algunos casos durante períodos prolongados, por lo que esta fase plantea cuestiones de seguridad y en torno a la limitación del plazo de conservación, que son principios que están recogidos en el RGPD. La tercera fase, es la de procesamiento y análisis. Esta es la más relevante para el área de investigación de trabajo ya que es aquí donde los datos se convierten en insumo para los sistemas de IA, y es precisamente en esta etapa del ciclo donde se materializan los principales riesgos éticos: sesgos algorítmicos, discriminación automatizada y violaciones de la privacidad (Mühlhoff, 2023).

La tercera fase es la de eliminación. Aunque el RGPD establece que los datos personales no deben conservarse más allá del período necesario para el fin para el que fueron recogidos, en la práctica empresarial el cumplimiento de este principio es frecuentemente difícil de constatar. Parte de esta falta de deber de cumplimiento se debe a que los modelos de IA entrenados con determinados datos retienen información sobre esos datos incluso después de eliminarlos. Como consecuencia, se han creado nuevos desafíos que permitan conseguir una aplicación efectiva del derecho al olvido reconocido en la legislación. (Meszaros y Ho, 2021).

2.5. La protección de datos personales en la empresa y sus desafíos en el entorno digital: cómo se relaciona la privacidad con el uso de IA

Milossi et al. (2021) explican las tensiones existentes entre los principios del RGPD y las lógicas inherentes que tienen los sistemas de IA. El RGPD defiende el principio de minimización de datos que, por ejemplo, entra en contradicción directa con la necesidad de los sistemas de IA de ser entrenados con el mayor volumen de datos posible para maximizar su eficiencia. El principio de transparencia que defiende el Reglamento choca con la

opacidad que caracteriza por naturaleza a los modelos de aprendizaje. Y el derecho a no ser objeto de decisiones basadas únicamente en el tratamiento automatizado, reconocido en el artículo 22 del RGPD, no tiene apenas cabida en entornos empresariales donde las decisiones algorítmicas se han convertido en el modus operandi.

Meszaros y Ho (2021), al examinar las diferencias en la aplicación del RGPD entre la investigación académica y el desarrollo comercial de productos y servicios basados en IA, añade otro punto de vista. Explica que las empresas tienen un régimen más estricto pero que con lagunas interpretativas sobre la base jurídica del tratamiento, el alcance del consentimiento y los requisitos de evaluación de impacto.

Mendez-Suarez et al. (2024) analizaron una muestra de 34 empresas sancionadas bajo el RGPD por el uso de aplicaciones de IA prohibidas o contrarias a sus principios éticos, concluyendo que los factores que mejor predicen el incumplimiento son la ausencia de una declaración ética explícita sobre el uso de la IA, la insuficiente ciberseguridad y operar en mercados con estándares éticos más laxos.

Estos hallazgos demuestran que la protección efectiva de los datos personales cuando se utilizan sistemas de IA requiere un compromiso con la ética empresarial y la gobernanza responsable, no es solo una cuestión de cumplimiento normativo.

3. FUNDAMENTOS ÉTICOS DE LA IA Y LA PROTECCIÓN DE DATOS

3.1. Conceptos clave: transparencia algorítmica y privacidad

La transparencia algorítmica y la privacidad constituyen los dos pilares fundamentales sobre los que se desarrolla el análisis ético del uso de la IA en el tratamiento de datos personales.

La transparencia algorítmica hace referencia a la capacidad que deben de tener todos los actores implicados (usuarios, reguladores, auditores y la propia organización) de comprender cómo funciona un sistema de IA, qué datos se utilizan, qué criterios se aplican y cómo llega el sistema a determinados resultados y decisiones. Kim y Routledge (2022) argumentan, en su artículo publicado en *Business Ethics Quarterly*, que la transparencia no es solo un valor técnico sino un derecho: las personas que se ven afectadas por decisiones algorítmicas tienen un derecho legítimo a recibir una explicación de los factores que determinaron esa decisión, y aún más si dicha decisión es perjudicial. Ya reconocido en el artículo 22 del RGPD, este derecho constituye uno de los puntos de mayor tensión entre las capacidades de la IA y las exigencias del marco normativo europeo, puesto es complicada su practicidad en los modelos más potentes que son opacos incluso ara quienes los crean. Este debate se posiciona como el central para los académicos actuales y se evidencia en los resultados del análisis temático que confirman que la transparencia y la explicabilidad suponen el bloque mas representado del corpus, con un 39% del total de los documentos.

La privacidad ha evolucionado a un concepto más amplio desde la concepción que la conocemos. Ahora, no se centra únicamente en la confidencialidad de la información si no que incluye controlar por parte del individuo sus propios datos y la protección frente a las consecuencias del análisis de esos datos por terceros. Mühlhoff (2023) explica otra forma de violación de la privacidad no queda cubierta por los marcos de protección de datos tradicionales, que se centran en la información declarada, si no en la que se da cuando a apartir de datos que parecen inconexos se llega a información intima de las personas.

3.2. Principios éticos aplicados a la IA

Wu et al. (2020) realizaron un análisis comparativo de las iniciativas de gobernanza ética de la IA desarrolladas en China. Identificaron cinco principios éticos que también están presentes en los marcos europeos: justicia, transparencia, privacidad, seguridad, que obliga a garantizar la robustez y fiabilidad de los sistemas y responsabilidad.

Rizinski et al. (2022) añade un sexto principio la explicabilidad. La diferencia en la transparencia es sutil pero importante: la transparencia supone que el sistema pueda ser comprendido, la explicabilidad que sus decisiones puedan ser justificadas en términos comprensibles para los no expertos, de forma que los afectados puedan ejercer efectivamente sus derechos. Sin embargo, este término supone un dilema técnico y muy debatido en la actualidad cuando hablamos de los modelos más precisos ya que estos tienden a ser los más opacos.

Vidgen et al. (2020) presentan una herramienta práctica para aplicar estos principios en el contexto empresarial: el "Business Ethics Canvas". Se trata de un marco estructurado, o plantilla, que permite a las organizaciones identificar las implicaciones éticas de sus proyectos de analítica de datos e IA antes de lanzarlos. Integra principios del utilitarismo (maximizar el bienestar general) y la ética deontológica (respetar los derechos independientemente de las consecuencias).

3.3. El papel de la ética empresarial en el uso de la IA

Gerlick y Liozu (2020) analizan concretamente el caso de la fijación algorítmica de precios, un ámbito en el que la IA permite a las empresas personalizar los precios en tiempo real en función del comportamiento y las características de cada usuario. De esta forma analizan como la ética empresarial influye en las decisiones sobre el uso de datos. Llegan a la conclusión de que, las empresas que cuentan con códigos éticos explícitos y mecanismos de supervisión interna tienden a establecer límites más claros sobre el uso de datos personales y son menos propensas a tener prácticas cuestionables desde la perspectiva de la justicia.

Anshari et al. (2023) examinan el rol de la ética empresarial en el uso de la IA durante la pandemia de COVID-19. Se trata de un ejemplo extremo para demostrar como la ética empresarial funciona bajo momentos de presión. El COVID-19 supuso un contexto que aceleró extraordinariamente la adopción de sistemas de vigilancia, seguimiento de contactos y toma de decisiones clínicas basadas en IA. Su conclusión principal fue que las organizaciones que lograron mantener un equilibrio entre la eficacia sanitaria y el respeto a los derechos individuales fueron aquellas con marcos éticos y órganos de gobernanza preexistentes.

3.4. La gobernanza ética de la IA en el contexto internacional.

En el plano global, Wu et al. (2020) documentan cómo, aunque en todo el mundo se comparten los cinco principios éticos fundamentales, cada país tiene una forma de interpretarlos y aplicarlos. Explican cómo China enfatiza de manera diferente valores como la armonía social y el beneficio colectivo, y se esfuerza menos en enfatizar en los derechos individuales, como si hace el modelo europeo. Esta diferencia tiene implicaciones prácticas significativas para las empresas que operan en múltiples jurisdicciones, y es relevante para aquellas que operan entre marcos regulatorios y expectativas éticas que no siempre son compatibles entre sí. Además, es también centro de debate esta divergencia internacional y, se evidencia en el corpus analizado, donde el bloque de gobernanza y regulación es el segundo más representado. Por lo tanto, gobernanza de la IA es uno de los desafíos más complejos para la ética empresarial global.

4. EL DILEMA ÉTICO DEL USO DE LA IA EN LA GESTIÓN DE LOS DATOS PERSONALES EN LAS EMPRESAS

4.1. El uso de datos masivos y sus implicaciones éticas

Mühlhoff (2023) identifica la paradoja central del big data: cuanto más útil y valiosa es la información que las empresas extraen de los datos para sus modelos de negocio, más invasiva se vuelve para la privacidad de las personas. Añade, además un grave problema: los mecanismos de consentimiento tradicionales que estaban diseñados para un contexto en el que las personas pudieran comprender qué información estaban cediendo y para con que fin, resultan inadecuados cuando el valor de los datos se genera a través de predicciones que los propios usuarios no pueden anticipar en el momento que otorgan su consentimiento. El nuevo modelo de los servicios digitales gratuitos a cambio de datos personales esconde unas implicaciones éticas que los usuarios no comprenden en totalidad y nos conscientes de la gravedad.

Truby y Brown (2021) llevan este análisis al extremo. Exploran el concepto de "clones digitales de pensamiento": la posibilidad de construir modelos computacionales que representen con alta fidelidad la personalidad, las preferencias y el comportamiento de un individuo concreto, es decir, quien eres, como piensas y como te vas a comportar. La propia existencia de este modelo plantea interrogantes éticos y jurídicos de primer orden la identidad y la autonomía.

4.2. Sesgos algorítmicos y discriminación automatizada

Los sesgos algorítmicos constituyen del dilema ético más documentado y con mayores consecuencias prácticas en el uso empresarial de la IA. Un sesgo algorítmico se define como una generación de un resultado, por parte de un sistema de IA, que resulta desfavorable para determinados grupos de población. No se dan por una intención discriminatoria en sí misma, sino como consecuencia de los datos históricos con los que fue entrenado o de las decisiones de diseño de sus creadores.

Rizinski et al. (2022) analizan el problema de los sesgos en el sector financiero. Demuestran con evidencia que los modelos entrenados con datos históricos tienden a reproducir y amplificar las desigualdades preexistentes en este sector. Por ejemplo, los grupos históricamente más desfavorecidos que hayan tenido menor acceso a productos financieros serán asignados con puntuaciones de riesgo más elevadas, perpetuando así la exclusión

financiera de manera aparentemente objetiva y técnica, lo que dificulta enormemente su cuestionamiento y corrección.

Kriebitz y Lütge (2020) defiende que esto no es un problema que sea únicamente un fallo técnico si no que supone un problema en el contexto de los derechos humanos. Explican que se trata de una violación de derechos fundamentales como la igualdad y la no discriminación. Destacan que las empresas no deben de respaldarse en que la decisión fue del algoritmo, y no intencionada por su parte, puesto que la responsabilidad, tanto moral como jurídica, sigue siendo suya y no se les puede eximir de ella.

4.3. La caja negra: falta de transparencia y dificultad para atribuir responsabilidades

La "caja negra" hace referencia a la opacidad de muchos sistemas de IA que lleva consigo la imposibilidad de comprender por qué un sistema tomó una decisión concreta en lugar de otra. Esta opacidad puede tener dos vertientes que deben distinguirse: la opacidad técnica, que es inherente a la arquitectura de ciertos modelos, y la opacidad organizativa, que resulta de decisiones que han sido deliberadas por las empresas de no revelar el funcionamiento de sus sistemas por razones de confidencialidad comercial o estratégica.

Kim y Routledge (2022) ilustran las consecuencias prácticas de esta opacidad explicando, con dos ejemplos concretos, como esta deteriora la confianza de los usuarios y su capacidad de ejercer sus derechos de manera efectiva. El primero, explica que si a un solicitante se le deniega un crédito no puede impugnar esa decisión ni demostrar que fue fruto de un sesgo discriminatorio. El segundo, demuestra que si un candidato es rechazado en un proceso de selección y no recibe una explicación comprensible de los motivos del rechazo, no puede saber si el sistema aplicó criterios legítimos o si incurrió en discriminación indirecta. La ausencia de transparencia y la presencia de la opacidad convierten las decisiones algorítmicas en una forma de poder que no tiene control y que no resulta compatible con los principios de un Estado de derecho.

Milossi et al. (2021) abordan este problema desde la perspectiva del RGPD. Por un lado, explican el "determinismo algorítmico" es la tendencia de los sistemas de IA a tratar sus outputs como veredictos objetivos e inapelables, lo cual es indefendible con el principio de autodeterminación informativa del RGPD. Para resolver este problema, sus conclusiones reflejan que se requiere no solo de avances técnicos en la interpretación de los modelos,

sino también cambios culturales en las organizaciones en la manera que conciben y utilizan sus sistemas de IA.

4.4. Pérdida de privacidad y vigilancia corporativa

Golda et al. (2024) documentan en su análisis sobre privacidad y seguridad en la IA generativa como se ha multiplicado el potencial de vigilancia de las empresas: la posibilidad de sintetizar información procedente de múltiples fuentes para construir perfiles detallados de individuos o de predecir comportamientos futuros precisos, constituyen herramientas que, en manos de empresas sin escrúpulos, pueden usarse para manipular o discriminar a las personas de formas que resultan muy difíciles de detectar.

Mühlhoff (2023) introduce el concepto de "desigualdad social de datos". Este término lo utiliza para poner de manifiesto como la IA amplifica las desigualdades preexistentes. Al final y al cabo, las personas con mayor vulnerabilidad social que tienen menos recursos, menor educación digital y mayor dependencia de servicios públicos y privados, son también las que tienen menor capacidad de proteger sus datos y de resistir a las consecuencias de estas decisiones algorítmicas.

4.5. Responsabilidad moral en la toma de decisiones automatizada

Gerlick y Liozu (2020) explican este fenómeno de "difusión de la responsabilidad": cuando una decisión es tomada por un algoritmo, nadie se siente del todo responsable ante ella. Lo definen como uno de los riesgos éticos más significativos del uso empresarial de la IA: si nadie siente dicha responsabilidad, los incentivos para garantizar que esas decisiones sean justas, precisas y respetuosas con los derechos de las personas disminuyen. Esta situación se agrava por la tendencia de muchas empresas a pensar que los outputs de la IA son datos objetivos y neutrales, olvidando que son el resultado de decisiones de diseño cargadas de valores y susceptibles de error.

Mendez-Suarez et al. (2024) aportan evidencia empírica de las consecuencias de esta dilución de la responsabilidad en su análisis de empresas sancionadas bajo el RGPD por el uso de aplicaciones de IA prohibidas. Este estudio muestra que el factor que mejor predice el incumplimiento es la ausencia de compromiso organizativo formal con la ética. Este entendimiento sugiere que la responsabilidad moral en el uso de la IA no puede dejarse a la discreción individual de los empleados o técnicos, debe de estar institucionalizada a través de políticas y estructuras de gobernanza formales.

5. RESULTADOS

5.1. Caracterización bibliométrica del corpus

El presente apartado recoge los principales resultados del análisis bibliométrico realizado sobre el corpus de 146 documentos seleccionados. El objetivo de este análisis es caracterizar la producción científica que existe sobre la interacción entre inteligencia artificial, ética empresarial y protección de datos personales, identificando las tendencias de publicación, las fuentes más relevantes y los trabajos de mayor impacto en el campo.

La Tabla 1 recoge los indicadores bibliométricos generales del corpus. El conjunto de documentos analizados acumula un total de 2.180 citas, con una media de 14,9 citas por artículo, lo que refleja un campo de investigación activo y con un nivel de impacto considerable. El 67,8% de los documentos son de acceso abierto, lo que ha facilitado la consulta de los textos completos para el análisis de contenido.

Tabla 1. Indicadores bibliométricos generales del corpus

Indicador	Valor
Total de documentos analizados	146
Período de cobertura	2020 – 2026
Número de revistas distintas	119
Total de citas acumuladas	2.180
Media de citas por artículo	14,9
Artículo más citado (nº citas)	259
Documentos en Open Access	99 (67,8%)
Artículos originales (Article)	122 (83,6%)
Revisiones sistemáticas (Review)	24 (16,4%)

Fuente: elaboración propia mediante Python a partir de datos de Scopus (2020-2026).

La Tabla 2 muestra la evolución temporal de las publicaciones. El patrón de crecimiento es muy pronunciado: si en 2020 y 2021 el número de artículos era modesto (7 y 4 respectivamente), a partir de 2023 se observa una aceleración notable que culmina en 2025 con 66 artículos, lo que representa el 45,2% del total del corpus. Este crecimiento refleja el creciente interés académico por la intersección entre IA, ética y protección de datos en el

ámbito empresarial, probablemente impulsado por la consolidación del RGPD, el debate en torno al AI Act europeo y la irrupción de la IA generativa.

Tabla 2. Distribución temporal de las publicaciones (2020-2026)

Año	Nº publicaciones	% del total
2020	7	4,8%
2021	4	2,7%
2022	9	6,2%
2023	13	8,9%
2024	32	21,9%
2025	66	45,2%
2026	15	10,3%
TOTAL	146	100%

Fuente: elaboración propia mediante Python a partir de datos de Scopus (2020-2026).

En cuanto a las fuentes de publicación, la Tabla 3 recoge las diez revistas con mayor presencia en el corpus. IEEE Access encabeza el ranking con 7 artículos, seguida del Journal of Data Protection and Privacy con 4. La amplia diversidad de fuentes —119 revistas distintas para 146 documentos— evidencia el carácter marcadamente interdisciplinar del campo, que convoca a investigadores de ingeniería, derecho, gestión empresarial y filosofía.

Tabla 3. Top 10 revistas por número de publicaciones

#	Revista	Nº artículos
1	IEEE Access	7
2	Journal of Data Protection and Privacy	4
3	Applied Sciences (Switzerland)	3
4	Intl. J. Advanced Computer Science and Applications	3
5	Applied Marketing Analytics	3
6	AI and Society	3
7	Discover Sustainability	2
8	IEEE Engineering Management Review	2
9	IEEE Transactions on Engineering Management	2

10	Frontiers in Artificial Intelligence	2
----	--------------------------------------	---

Fuente: elaboración propia mediante Python a partir de datos de Scopus (2020-2026).

La Tabla 4 recoge los diez artículos más citados del corpus, que constituyen las referencias más influyentes del campo. El trabajo de Coorey et al. (2022) sobre gemelos digitales en salud encabeza el ranking con 259 citas, aunque su temática principal es la medicina de precisión. Para los objetivos específicos de este trabajo destacan especialmente los de Kriebitz y Lütge (2020) sobre IA y derechos humanos (57 citas), Mühlhoff (2023) sobre privacidad predictiva (59 citas) y Kim y Routledge (2022) sobre el derecho a la explicación algorítmica (56 citas).

Tabla 4. Top 10 artículos más citados del corpus

Autores	Año	Título	Citas
Coorey et al.	2022	The health digital twin to tackle cardiovascular disease	259
Golda et al.	2024	Privacy and Security Concerns in Generative AI: A Comprehensive Survey	179
Minssen et al.	2020	Regulatory responses to medical machine learning	96
Wu et al.	2020	Ethical Principles and Governance Technology Development of AI in China	85
Vo et al.	2023	Multi-stakeholder preferences for the use of AI in healthcare	83
Mikołajewska et al.	2025	Generative AI in AI-Based Digital Twins for Fault Diagnosis	70
Mühlhoff	2023	Predictive privacy: Collective data protection in the context of AI	59
Kriebitz y Lütge	2020	Artificial Intelligence and Human Rights: A Business Ethical Assessment	57
Kim y Routledge	2022	Why a Right to an Explanation of Algorithmic Decision-Making Should Exist	56
Gerlick y Liozu	2020	Ethical and legal considerations of AI and algorithmic decision-making	54

Fuente: elaboración propia mediante Python a partir de datos de Scopus (2020-2026).

Las Figuras 4 y 5 muestran la distribución del corpus por tipo de documento y modalidad de acceso. El corpus está compuesto mayoritariamente por artículos originales (83,6%), con un 16,4% de revisiones sistemáticas, lo que garantiza la solidez empírica del análisis.

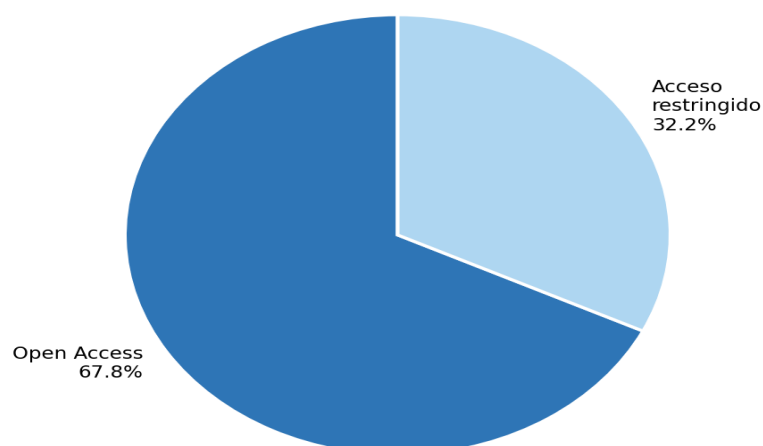
En cuanto al acceso, el 67,8% de los documentos son Open Access, lo que ha facilitado la consulta directa de los textos completos.

Figura 4. Distribución por tipo de documento

Tipo de documento	Nº documentos	% del total
Artículo (Article)	122	83,6%
Revisión sistemática (Review)	24	16,4%
TOTAL	146	100%

Fuente: elaboración propia mediante Python a partir de datos de Scopus (2020-2026).

Figura 5. Distribución por modalidad de acceso



Fuente: elaboración propia mediante Python a partir de datos de Scopus (2020-2026).

Finalmente, la Tabla 5 y la Figura 6 muestran las veinte palabras clave más frecuentes del corpus. "Artificial intelligence" encabeza ampliamente el ranking con 67 apariciones, seguida de "data privacy" (22) y "machine learning" (21). La presencia destacada de términos como "GDPR", "AI ethics", "algorithmic bias", "data protection" y "transparency" confirma la pertinencia temática del corpus seleccionado en relación con los objetivos de este trabajo y anticipa los grandes bloques temáticos que se desarrollan en los capítulos siguientes.

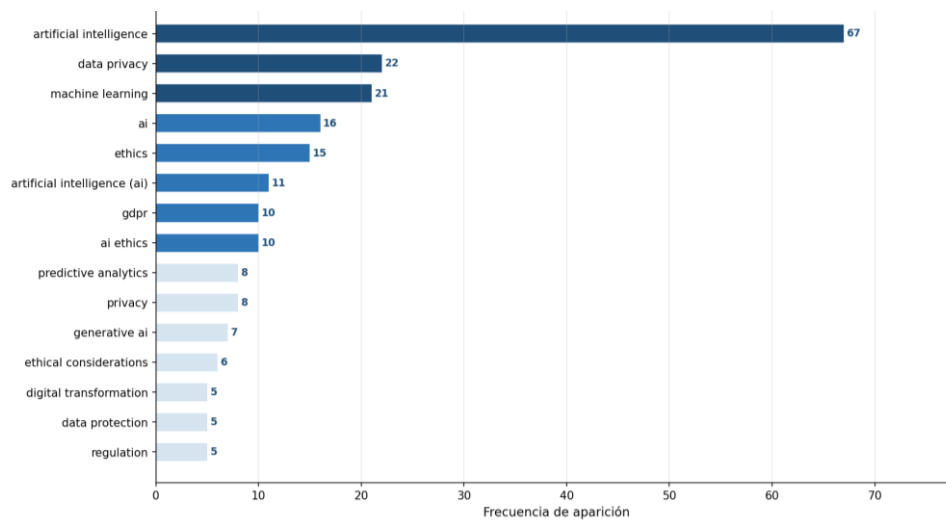
Tabla 5. Top 20 palabras clave más frecuentes del corpus

#	Palabra clave	Frecuencia
1	Artificial intelligence	67

2	Data privacy	22
3	Machine learning	21
4	AI	16
5	Ethics	15
6	Artificial intelligence (AI)	11
7	GDPR	10
8	AI ethics	10
9	Predictive analytics	8
10	Privacy	8
11	Generative AI	7
12	Ethical considerations	6
13	Digital transformation	5
14	Data protection	5
15	Regulation	5
16	Digital marketing	5
17	Ethical AI	5
18	Algorithmic bias	5
19	Data security	4
20	Transparency	4

Fuente: elaboración propia mediante Python a partir de datos de Scopus (2020-2026).

Figura 6. Top 15 palabras clave más frecuentes



Fuente: elaboración propia mediante Python a partir de datos de Scopus (2020-2026).

5.2. Análisis de tendencias temáticas

La primera y más numerosa es la gobernanza y regulación, que agrupa los estudios centrados en el RGPD, el AI Act, los marcos de cumplimiento normativo y las políticas de gobernanza de datos (56 artículos, 38,4% del corpus). La elevada presencia de esta temática refleja el impacto de la regulación europea en la agenda investigadora y la preocupación de la comunidad académica por traducir los principios normativos en herramientas prácticas para las organizaciones.

La segunda área es la de sesgos algorítmicos y discriminación, con 43 artículos (29,5%). Esta es la temática con mayor crecimiento en términos de producción científica reciente, lo que refleja la creciente conciencia sobre los riesgos discriminatorios de los sistemas de IA entrenados con datos históricos sesgados. Los estudios de este bloque se concentran especialmente en los sectores financiero, sanitario y de recursos humanos.

La tercera área es la IA generativa y nuevas tecnologías, con 25 artículos (17,1%). Este es el bloque temático más reciente del corpus, con la práctica totalidad de sus documentos publicados a partir de 2023, lo que refleja la irrupción de los grandes modelos de lenguaje (LLMs) y sus implicaciones específicas para la privacidad y la ética empresarial. La cuarta área es la privacidad y protección de datos, con 15 artículos (10,3%), que complementa el bloque de gobernanza con estudios más centrados en los derechos individuales y los mecanismos técnicos de protección. Finalmente, la transparencia y explicabilidad agrupa 6 artículos (4,1%), un número relativamente reducido que sin embargo incluye algunas de las referencias más influyentes del corpus.

5.3. Análisis temático mediante topic modelling

Para identificar los temas principales de los 146 artículos presentes del corpus, se ha empleado la técnica de LDA (Asignación Latente de Dirichet). Es un tipo de topic modelling que se ha aplicado sobre los abstracts usando Python con la librería sklearn. Lo que permite esta técnica es hacer una lectura de todos los textos e identificar cuáles son los temas que más aparecen a lo largo de ellos, lo que tienen mayor frecuencia, asignándole a cada artículo una probabilidad de pertenecer a un tema.

El modelo se decidió para buscar seis temas después de probar varias opciones y comprobar cual es la que daba los resultados más coherentes. Antes de aplicar dicho modelo, se limpiaron los textos eliminando los stopwords en inglés, la tokenización, de tal forma que

se eliminan todas las palabras que no aportan significado.. Los resultados del análisis se presentan en la Tabla 6 y las Figuras 4, 5 y 6.

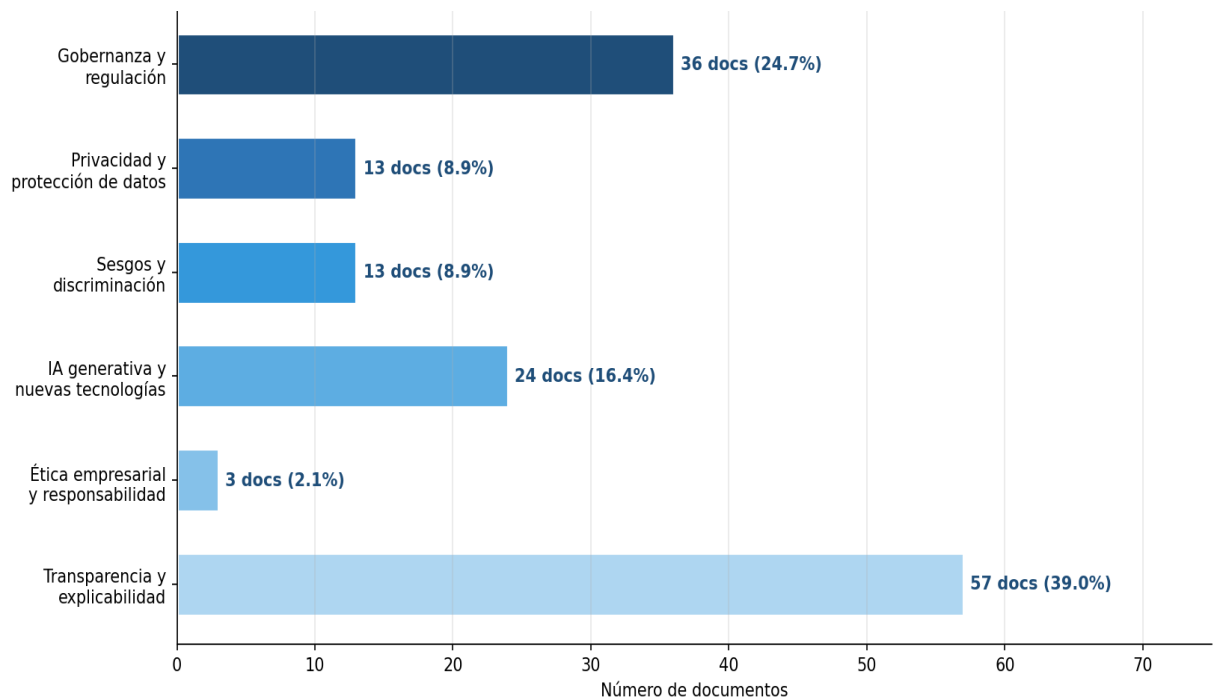
Tabla 6. Temas identificados mediante topic modelling sobre los abstracts del corpus

#	Tema	Palabras clave representativas	Docs	%
1	Gobernanza y regulación (RGPD/AI Act)	governance, privacy, legal, regulation, protection, digital	36	24,7%
2	Privacidad y protección de datos	adoption, automation, business, innovation, technologies, SMEs	13	8,9%
3	Sesgos y discriminación	human rights, law, legal, issues, rights, human	13	8,9%
4	IA generativa y nuevas tecnologías	healthcare, social, health, intelligence, privacy, social media	24	16,4%
5	Ética empresarial y responsabilidad	blockchain, digital, technology, sustainable, education, commerce	3	2,1%
6	Transparencia y explicabilidad	ethical, business, customer, decision, intelligence, machine learning	57	39,0%

Fuente: elaboración propia mediante Python a partir de datos de Scopus (2020-2026).

La siguiente figura explica los resultados principales del modelo. Muestra la distribución de documentos por tema. El tema más frecuente es el de Transparencia y explicabilidad, con 57 documentos (39,0% del corpus), lo que refleja que la ética de la IA es la preocupación central del campo de investigación. El segundo es el tema de Gobernanza y regulación con 36 documentos (24,7%), que pone de manifiesto el impacto de la nueva regulación del RGPD y el AI Act. El tercero es el tema de IA generativa y nuevas tecnologías, que agrupa 24 documentos (16,4%), que se concentra en el tiempo, muy significativamente, en los años 2024 y 2025, lo que refleja la irrupción tan reciente que existe sobre estos modelos y el debate ético empresarial acerca de ellos.

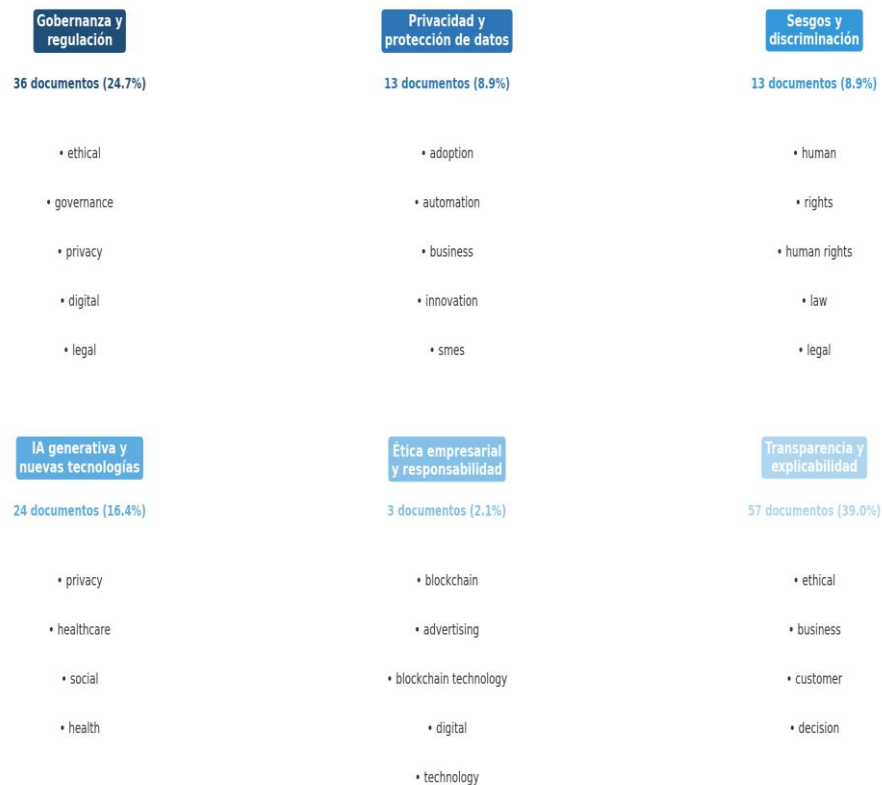
Figura 7. Distribución del corpus por temas identificados mediante topic modelling



Fuente: elaboración propia mediante Python a partir de datos de Scopus (2020-2026).

La siguiente figura demuestra las palabras clave, las más representativas de cada tema. De esta forma, el lector entiende mejor de que trata cada bloque puesto que caracteriza con mayor precisión el contenido de cada bloque temático. El tema de Transparencia y explicabilidad incluye conceptos como decision, ethical, business, customer y machine learning, reflejando, una vez más, la preocupación que existe por hacer comprensibles las decisiones algorítmicas en contextos empresariales. El tema de Gobernanza y regulación combina términos jurídicos como regulación o legal con conceptos de gestión como governance o digital, lo que refleja su el carácter interdisciplinar de este tema.

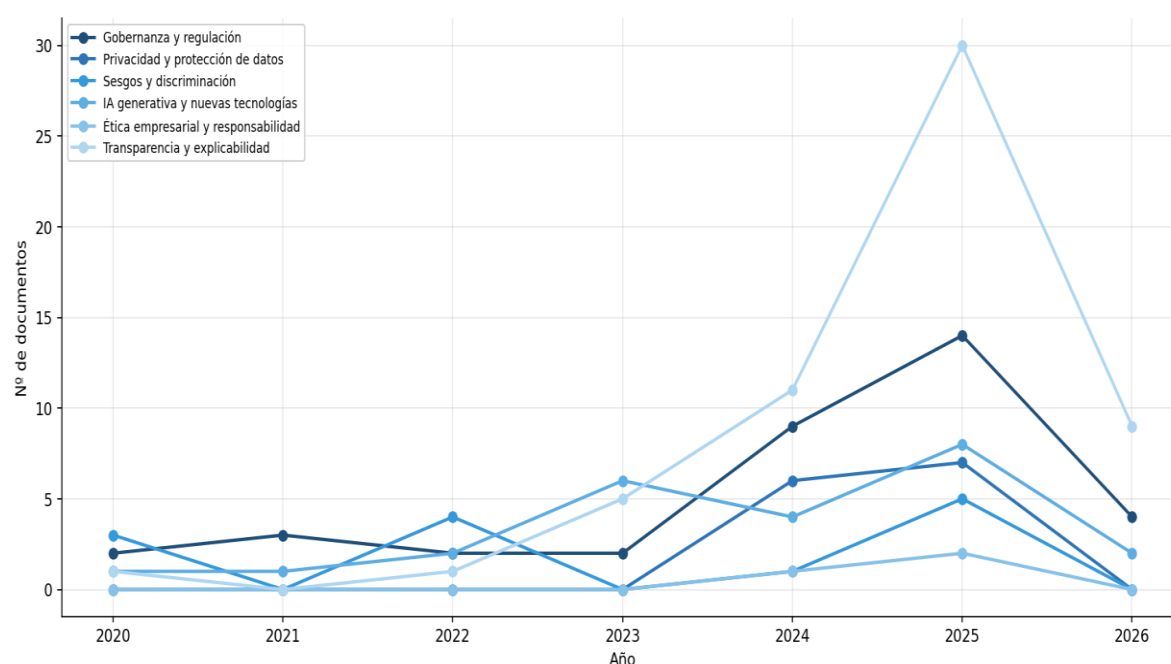
Figura 8. Palabras clave principales por tema identificado



Fuente: elaboración propia mediante Python a partir de datos de Scopus (2020-2026).

La última figura analiza como han ido evolucionando los determinados temas a lo largo de los años, detectando la diferencia de tres patrones. En primer lugar, el tema de Transparencia y explicabilidad mantiene un crecimiento constante a lo largo de todo el período, con un pico significativo en 2025. En segundo lugar, el tema de IA generativa refleja un crecimiento bastante acelerado a partir de 2023, que coinciden con la llegada masiva de modelos de IA generativa como ChatGPT. En tercer lugar, los temas de Sesgos y discriminación y Privacidad y protección de datos muestran una presencia más estable en el tiempo, lo que significa que son preocupaciones y debates que están consolidados en la literatura científica más que tendencias modernas.

Figura 9. Evolución temporal de los temas identificados (2020-2026)



Fuente: elaboración propia mediante Python a partir de datos de Scopus (2020-2026).

Relacionando los resultados del topic modelling con los del análisis bibliométrico, se observa una coherencia suficiente para matizar algunas de las conclusiones. En particular, la identificación de Transparencia y explicabilidad como el tema más representado en el corpus, por encima incluso de Gobernanza y regulación, demuestra que la preocupación académica por comprender como funcionan los algoritmos de los sistemas de IA va más allá del cumplimiento legal. Se puede observar que tiene una dimensión ética propia que merece atención personalizada.

5.4. Principales conclusiones del análisis de contenido

Las conclusiones que se van a dar, están contextualizadas en los apartados anteriores y en los resultados del análisis bibliométrico y del topic modelling. Estos permiten reforzar los hallazgos mencionados a continuación, que han sido sobretodo identificados a partir de un análisis cualitativo de los artículos más relevantes del corpus.

La primera evidencia es la existencia de una tensión estructural que existe entre la lógica de maximizar el valor de los datos y los principios de protección de datos y respeto de los derechos fundamentales que caracteriza al marco normativo europeo. Lo que se puede percibir es que hay un conflicto entre querer tener gran cantidad de datos para explotarlos con ánimo de lucro y querer por otro lado respetar la privacidad. Cabe señalar que, al final, esta forma de actuar representa el modelo de negocio de la economía digital, que esta

tensión no es fruto de una mala gestión de una empresa puntual, si no que el sistema, de alguna forma, está diseñado así y los incentivos están empujando en esta dirección y que forman el comportamiento de las organizaciones actuales. (Kriebitz y Lütge, 2020; Mühlhoff, 2023; Mendez-Suarez et al., 2024).

El segundo hallazgo son los tres dilemas más graves que aparecen en los artículos. En primer lugar la discriminación algorítmica: se puede observar cómo los algoritmos tratan peor a ciertos grupos simplemente porque aprendió de datos históricos sesgados. En segundo lugar, la opacidad de los sistemas (black box): en realidad, nadie sabe exactamente porque el algoritmo toma ciertas decisiones e incluso quien lo creó. En tercer lugar, la erosión de la privacidad a través del profiling predictivo: las empresas tienen información íntima del personal que estos nunca revelaron que sabe por los comportamientos de estos. Estos tres dilemas están conectados entre sí: la opacidad dificulta la detección de la discriminación porque si no sabes cómo funcionan el algoritmo tampoco puedes detectar si se está discriminando, y además, si se hacen perfiles predictivos sin que lo sepa el personal no se puede dar un consentimiento real ante este, violando así la privacidad de formas que los marcos de consentimiento tradicionales no son capaces de capturar ni de prevenir (Kim y Routledge, 2022; Rizinski et al., 2022; Mühlhoff, 2023).

El tercer hallazgo es la poca efectividad a la hora de adoptar, por parte de las empresas, las herramientas organizativas y técnicas de las que disponen para tener una gobernanza responsable de la IA. Méndez-Suarez et al. (2024), explican que el factor más determinante del cumplimiento normativo no es el tamaño de la empresa ni el sector de actividad, si no si la empresa tiene o no una política explícita sobre IA y estructuras de gobernanza reales, y con poderes efectivos. (Mendez-Suarez et al., 2024; Wu et al., 2020; Gursoy et al., 2025).

El cuarto hallazgo es la convergencia global que existe hacia los principios éticos, sin embargo, la discrepancia en la interpretación y aplicación de estos. A nivel mundial, prácticamente todos los países coinciden en que la IA debe de garantizar justicia, transparencia, privacidad, seguridad y responsabilidad, sin embargo, se observan diferencias significativas en su implementación según el contexto cultural, jurídico y político. Por tanto, y así también lo demuestran en su estudio Wu et al. comparando los principios éticos de la IA en China y el modelo europeo, existe un consenso global sobre los valores que deben orientar el desarrollo de la IA, pero las vías para hacerlos efectivos siguen siendo objeto de debate (Wu et al., 2020; Anshari et al., 2023).

6. GOBERNANZA Y BUENAS PRÁCTICAS ÉTICAS EN LAS EMPRESAS

6.1. Modelos de gobernanza de datos con IA

La gobernanza de datos se define como una herramienta que ayuda a garantizar la integridad y la seguridad de los datos mediante la definición e implementación de políticas, estándares y procedimientos para la recopilación, propiedad, almacenamiento, procesamiento y uso de datos. El objetivo de la gobernanza de datos es mantener datos seguros y de alta calidad a los que se pueda acceder fácilmente para las iniciativas de descubrimiento de datos y business intelligence (*IBM, s.f*). Es decir, es el conjunto de políticas, procesos, y tecnologías que una empresa usa para gestionar sus datos de forma correcta y alineada con sus objetivos estratégicos y sus valores. En el contexto del uso de la IA con datos personales, este concepto supone una visión que va más allá de la mera gestión técnica porque implica un compromiso ético con el uso responsable de los datos el respeto de los derechos de las personas.

Wu et al. (2020) explican el modelo de gobernanza ética de la IA a través tres niveles complementarios. El primer nivel trata de los principios: la empresa debe comprometerse formalmente con valores éticos que orienten su uso de la IA. El segundo nivel es el de los procesos, deben establecerse procedimientos concretos de evaluación del impacto ético de los proyectos de IA antes de su lanzamiento y de monitorización de su comportamiento una vez lanzados. El tercer nivel es el de las tecnologías de gobernanza: herramientas técnicas que permitan auditar los sistemas de IA y garantizar la trazabilidad de las decisiones.

Golda et al. (2024) defienden que es importante integrar la gobernanza de datos en el ciclo de vida completo de los sistemas de IA siguiendo el principio de "privacidad por diseño" que consagra el RGPD. Este concepto se refiere a que la privacidad no puede ser un añadido posterior al sistema de IA, sino que debe de estar incorporada desde el inicio. Para un cumplimiento efectivo se requiere que los equipos de desarrollo de IA incluyan perfiles con formación en ética, derecho y ciencias sociales, y no solo se dediquen a ello perfiles de ciencias como son la ingeniería o la estadística.

6.2. Herramientas de auditoría ética y transparencia algorítmica

La auditoría ética de los sistemas de IA es un componente esencial de cualquier modelo de gobernanza. Sin embargo, es un campo que todavía no está tan desarrollado como las auditorías financieras o técnicas, que tienen metodologías que están consolidadas por el

paso del tiempo y estándares reconocidos a nivel global. En la auditoría ética todavía no existe un consenso sobre los métodos, los criterios y los actores que deben intervenir.

Rizinski et al. (2022) proponen un marco concreto de auditoría para los modelos de machine learning. Combina tres elementos: interpretabilidad del, análisis de equidad y revisión de los datos de entrenamiento. Con estos se puede entender porque el modelo toma sus decisiones, se puede detectar discriminación y se puede hacer una revisión de los datos de entrenamiento. Lo más importante es que este marco permite identificar no solo si un modelo produce resultados discriminatorios, si no que identifica en que punto del proceso se origina dicha discriminación. Los autores subrayan que las herramientas técnicas de auditoría deben complementarse con mecanismos de gobernanza organizativa sin obviar la supervisión humana.

Vidgen et al. (2020) proponen el uso del Business Ethics Canvas, como se ha mencionado anteriormente, como herramienta de auditoría preventiva que se aplica antes de desarrollar el sistema. De forma que se evalúe el impacto que puede generar sobre los distintos grupos de partes interesadas y, como consecuencia, diseñar salvaguardas específicas para los riesgos que se identifiquen. Este enfoque es más eficiente que la detección y corrección de problemas cuando el sistema ya está en funcionamiento y ya ha ocasionado daños.

6.3. Casos reales de empresas tecnológicas

El análisis de casos reales de empresas que han afrontado problemas éticos en el uso de la IA con datos personales ofrece una perspectiva empírica para comprender los fallos que se dan en la gobernanza y cuáles son las condiciones que favorecen las buenas prácticas.

Mendez-Suarez et al. (2024), en su muestra de análisis de empresas sancionadas bajo el RGPD por el uso de aplicaciones de IA que violaban los principios éticos del reglamento, incluyen empresas que utilizaban sistemas de reconocimiento facial sin base jurídica adecuada, que realizaban perfilado de usuarios sin respetar el derecho a la intervención humana, o que transferían datos personales a terceros para su uso en sistemas de IA sin el consentimiento de los terceros involucrados. Y, de nuevo, se llega a la misma conclusión: existen tres factores que predicen fácilmente el incumplimiento: la ausencia de declaración ética sobre el uso de la IA, un nivel insuficiente de atención a la ciberseguridad, y la operación en países con estándares más laxos sin mecanismos de supervisión

En contraste, Gursoy et al. (2025) documentan casos de empresas del sector de la hospitalidad que si que han logrado implementar sistemas de IA para la personalización de servicios y la gestión operativa de manera éticamente responsable, convirtiéndolo en una ventaja competitiva basada en la confianza del cliente. Lo hacen gracias a la adopción de un marco de Responsabilidad Digital Corporativa, anteriormente explicado, que integra la protección de datos y el bienestar de los empleados como factores inseparables. Estos casos demuestran que la ética en el uso de la IA no es incompatible con la competitividad empresarial, al contrario, puede significar una ventaja competitiva para ellas.

7. PROPUESTAS FUTURAS

7.1. Recomendaciones para la gestión ética de los datos personales con IA

A partir del análisis realizado en los capítulos anteriores, este apartado tiene como objetivo formular recomendaciones dirigidas a las empresas para garantizar una gobernanza ética de sus sistemas de IA en el tratamiento de datos personales. Estas recomendaciones se concretan en tres niveles: estratégico, organizativo y técnico.

En primer lugar, a nivel estratégico, la recomendación principal es que las empresas adopten formalmente una política de IA responsable que defina los mecanismos de gobernanza y que se haga por escrito y públicamente, que sea verificable y que sea aprobada por el consejo de administración o la dirección general. Mendez-Suarez et al. (2024) han demostrado que es la mejor manera para evitar malas prácticas en el uso de la IA.

En segundo lugar, a nivel organizativo, se recomienda crear de comités de ética de la IA con composición multidisciplinar, es decir, que pertenezcan a él ingenieros, juristas, especialistas en ética o representantes de los usuarios. Kriebitz y Lütge (2020) subrayan que estos comités solo pueden ser efectivos si tienen capacidad real de vetar o modificar proyectos que no superen los estándares éticos establecidos, es decir, que no sea un comité únicamente consultivo si no que pueda paralizar los proyectos.

En tercer lugar, a nivel técnico, se recomiendan tres técnicas concretas: la implementación de sistemas de evaluación de impacto sobre la protección de datos para todos los proyectos de IA, la adopción de técnicas de IA explicable para decisiones de alto impacto, y la realización de auditorías periódicas de equidad para detectar sesgos algorítmicos. Rizinski et al. (2022) ofrecen una guía metodológica concreta para implementarlas y, aunque viene del sector financiero, es aplicable a cualquier sector.

7.2. Tendencias emergentes para la gobernanza ética empresarial: hacia una IA responsable

El campo de la gobernanza ética de la IA está evolucionando muy escalonadamente en el tiempo. Está impulsado tanto por los avances técnicos como por los desarrollos de regulación y los cambios en las expectativas sociales. Existen tres tendencias que marcarán el desarrollo de este campo en los próximos años y que tendrán relevantes implicaciones para las empresas.

La primera tendencia es la entrada en vigor del AI Act europeo. Este reglamento, va a crear nuevos estándares obligatorios para los sistemas de IA de alto riesgo. Incluirá requisitos de transparencia, documentación técnica, supervisión humana y gestión de riesgos—. Todas las empresas que operan en Europa tendrán que integrar estos requisitos en sus procesos de gobernanza. Wu et al. (2020) explican que la tendencia hasta esta nueva regulación no es sorprendente pues ya se anticipaba en la literatura desde el 2020 al documentar la creciente convergencia internacional en torno a un conjunto de principios éticos compartidos.

La segunda tendencia es la privacidad por diseño técnico, es decir, el desarrollo de tecnologías de IA que incorporan la privacidad como característica de diseño. Se observan dos tecnologías; el aprendizaje federado, que permite entrenar modelos de IA directamente en los dispositivos de los usuarios sin necesidad de centralizar los datos personales; y el aprendizaje diferencial, que añade ruido estadístico controlado a los datos para que sea imposible identificar a individuos concretos sin perder la utilidad del modelo. Golda et al. (2024) analizan estas tecnologías como soluciones para resolver la tensión entre el rendimiento de los sistemas de IA y la protección de la privacidad, y advierten que ninguna técnica puede resolver por sí sola un problema si no existen medidas organizativas y regulatorias que la complementen.

La tercera tendencia es la IA responsable. Se trata de un marco integrador que va más allá de la ética: integra al mismo tiempo sostenibilidad ambiental, impacto social y gobernanza corporativa. Gursoy et al. (2025), en el artículo más reciente del corpus, argumentan que la Responsabilidad Digital Corporativa representa la mejor estrategia que pueden adoptar las empresas para estar bien posicionadas y generar confianza a largo plazo.

7.3. ¿Equilibrio entre privacidad, innovación y derechos humanos?

Esta pregunta nos conduce al interrogante más importante que subyace a todos los análisis realizados en este trabajo. ¿Es posible aprovechar el potencial de la IA para generar valor económico y social sin comprometer la privacidad de las personas ni violentar sus derechos fundamentales? La respuesta que nace de la literatura analizada es afirmativa, pero sujeta a condiciones muy exigentes.

Desde una perspectiva deontológica, Anshari et al. (2023) argumentan que existen ciertos límites que no se pueden cruzar sea cual sea la situación o el beneficio que se pueda sacar de ella si estos violan derechos fundamentales. Esta posición es coherente con el enfoque

del RGPD y del AI Act, que establecen prohibiciones absolutas para determinados usos de la IA por considerarlos incompatibles con los derechos fundamentales con independencia de sus potenciales beneficios.

Desde una perspectiva más pragmática, Kriebitz y Lütge (2020) sostienen que el conflicto entre innovación y derechos humanos no es inevitable. Si las empresas diseñan correctamente sus sistemas desde el principio y establecen mecanismos de supervisión reales, se pueden tener sistemas de IA eficaces y respetuosos con los derechos de las personas. El equilibrio es posible, pero requiere voluntad política, inversión y un compromiso con la ética que trascienda del cumplimiento de la normativa.

8. CONCLUSIONES

8.1. Respuesta a las preguntas de investigación: principales hallazgos

El presente trabajo tenía como principal objetivo analizar las principales implicaciones éticas que surgen del uso de la inteligencia artificial en la gestión de datos personales por parte de las empresas, y poder estudiar de qué manera pueden las organizaciones mantener una gobernanza responsable y que esté integrada con principios éticos para garantizar la protección efectiva de dichos datos. A continuación, se van a responder a las preguntas de investigación planteadas a través de diferentes hallazgos.

En relación con la pregunta de investigación principal, el análisis del corpus permite constatar que las implicaciones éticas del uso empresarial de la IA con datos personales son estructurales. Como se menciona en el apartado de resultados, no es algo que pase en empresas aisladas, si no que el propio sistema digital está diseñado para explotar los datos y esto choca con la protección de los derechos de las personas. Esta tensión está inscrita en la arquitectura de los sistemas de IA y en los incentivos que estructuran el comportamiento de las organizaciones. Méndez-Suarez et al. (2024), explican que, al haber analizado 34 empresas que habían sido sancionadas por el RGPD, se puede observar cómo estas lo incumplían por falta de compromiso organizativo con la ética y no por desconocimiento. Por lo tanto, la gobernanza responsable es posible, pero requiere un compromiso que vaya mucho más allá del cumplimiento formal de la normativa, es decir, no basta con no infringir el RGPD, si no de comprometerse activamente con la ética. (Kriebitz y Lütge, 2020; Méndez-Suarez et al., 2024). Tanto los resultados del topic modelling como los del análisis bibliométrico que demuestran que la gobernanza y la transparencia son los bloques más representados en la literatura reciente, esfuerzan esta conclusión

En relación con la primera pregunta específica, los dilemas éticos más frecuentes y graves identificados en la literatura son la discriminación algorítmica, la opacidad de los sistemas (black box) y la erosión de la privacidad a través del profiling predictivo. Estos tres dilemas que están conectados entre sí revelan que se siguen perpetuando desigualdades históricas porque los algoritmos aprenden de datos sesgados. Nos revelan la importancia de que las personas, como derecho, reciban una explicación lógica de las decisiones algorítmicas que les afectan directamente y de que se evite la violación de la privacidad cuando se analiza el comportamiento de las personas para conseguir datos sensibles. (Kim y Routledge, 2022; Rizinski et al., 2022; Mühlhoff, 2023).

En relación con la segunda pregunta específica, las estrategias de gobernanza más efectivas son aquellas que integran los principios éticos desde el diseño, no como un añadido posterior. La gobernanza de la ética en el uso de la IA necesita tanto principios formales, como procesos concretos y tecnologías que la controlen. La evidencia empírica indica que la existencia de una declaración ética explícita sobre el uso de la IA es el factor más determinante del cumplimiento normativo, por encima de factores técnicos o económicos (Mendez-Suarez et al., 2024; Wu et al., 2020).

En relación con la tercera pregunta específica, las mejores prácticas identificadas se mueven en torno a tres características: la transparencia activa hacia las personas afectadas por las decisiones algorítmicas, la existencia de órganos de supervisión con poderes reales y la evaluación continua del impacto ético de los sistemas. Estas prácticas no son incompatibles con la competitividad empresarial. Kriebitz y Lütge refuerzan la idea de que el compromiso ético es compatible con ser competitivo y que, al contrario de evitar la competitividad, la evidencia pone de manifiesto que las empresas que adoptan buenas prácticas generan mayor confianza y están mejor posicionadas ante los desafíos regulatorios que se anuncian con el AI Act (Gursoy et al., 2025; Kriebitz y Lütge, 2020).

Estos hallazgos, en su conjunto, sugieren que haya un cambio en la forma en la que las organizaciones conciben los sistemas de IA, que se trabaje en el equilibrio entre la innovación y la protección de datos como algo que puede, perfectamente, funcionar a la par.

8.2. Limitaciones del estudio y futuras líneas de trabajo

El presente trabajo presenta un conjunto de limitaciones metodológicas y de alcance que se deben tener en cuenta al interpretar las conclusiones y que abren, a su vez, líneas de investigación futura.

La primera, y más relevante, es el uso exclusivo de Scopus como fuente bibliográfica. Aunque Scopus es una de las bases de datos científicas de mayor cobertura, la no inclusión de Web of Science, implica que parte de la cobertura científica relevante sobre el tema puede haber quedado fuera del corpus. Ambas bases de datos presentan similitudes, pero también diferencias en su cobertura por áreas temáticas y revistas; en particular, Web of Science tiene una mayor presencia de publicaciones en ciencias jurídicas y humanidades, áreas relevantes para el análisis ético que desarrolla este trabajo. Futuras investigaciones

podrían replicar el análisis utilizando ambas fuentes de forma combinada para obtener una visión más completa del estado del área de investigación.

La segunda limitación está relacionada con la ecuación de búsqueda utilizada. Si bien la query empleada fue diseñada de forma que pudiera capturar los documentos más relevantes para el objeto de estudio, cualquier ecuación de búsqueda implica decisiones sobre qué términos incluir y cuáles excluir, lo que puede sesgar el corpus resultante. Por ejemplo, la no inclusión de términos como "algorithmic fairness", "explainable AI" o "responsible AI" como términos de búsqueda independientes puede haber dejado fuera documentos relevantes que no utilizan los términos "ethics" o "ethical" en su título, abstract o palabras clave pero que abordan problemáticas directamente relacionadas con el objeto de estudio.

La tercera limitación es la velocidad de evolución del campo analizado. La inteligencia artificial está avanzando muy significativamente, especialmente la IA generativa, lo que puede terminar en conclusiones que se queden parcialmente superadas en un período de tiempo relativamente breve. Esta limitación pone de manifiesto la necesidad de actualización constante de la investigación en esta área.

Por lo tanto, y a pesar de las limitaciones señaladas, el presente trabajo contribuye a definir el conocimiento que existe sobre las implicaciones éticas en el uso de los sistemas IA en la gestión de los datos personales en las empresas. Estas limitaciones, no invalidan las conclusiones que se han alcanzado, ni los resultados de los análisis bibliométricos, topic modelling y análisis cualitativo de la literatura relevante a cerca del tema. Al contrario, abren nuevos caminos de investigación para profundizar y ampliar conclusiones, de forma que se pueda contribuir a un avance en este campo de investigación que, como se ha ido observando a lo largo de trabajo, está en continua expansión y esta ganando mayor relevancia en el contexto empresarial.

BIBLIOGRAFÍA

Literatura científica

Anshari, M., Hamdan, M., Ahmad, N., Ali, E., y Haidi, H. (2023). COVID-19, artificial intelligence, ethical challenges and policy implications. *AI and Society*, 38(2), 707–720.

<https://doi.org/10.1007/s00146-022-01471-6>

Gerlick, J. A., y Liozu, S. M. (2020). Ethical and legal considerations of artificial intelligence and algorithmic decision-making in personalized pricing. *Journal of Revenue and Pricing Management*, 19(2), 85–98. <https://doi.org/10.1057/s41272-019-00225-2>

Golda, A., Mekonen, K., Pandey, A., Singh, A., Hassija, V., Chamola, V., y Sikdar, B. (2024). Privacy and security concerns in generative AI: A comprehensive survey. *IEEE Access*, 12, 48126–48144. <https://doi.org/10.1109/ACCESS.2024.3381611>

Gursoy, D., Başer, G., y Chi, C. G. (2025). Corporate digital responsibility: navigating ethical, societal, and environmental challenges in the digital age and exploring future research directions. *Journal of Hospitality Marketing and Management*, 34(3), 305–324. <https://doi.org/10.1080/19368623.2025.2465634>

Kim, T. W., y Routledge, B. R. (2022). Why a right to an explanation of algorithmic decision-making should exist: A trust-based approach. *Business Ethics Quarterly*, 32(1), 75–102. <https://doi.org/10.1017/beq.2021.3>

Kriebitz, A., y Lütge, C. (2020). Artificial intelligence and human rights: A business ethical assessment. *Business and Human Rights Journal*, 5(1), 84–104. <https://doi.org/10.1017/bhj.2019.28>

Mendez-Suarez, M., De Obesso, M. M., Marquez, O. C., y Palacios, C. M. (2024). Why do companies employ prohibited unethical artificial intelligence practices? *IEEE Transactions on Engineering Management*, 71, 12218–12227. <https://doi.org/10.1109/TEM.2023.3258686>

Meszaros, J., y Ho, C.-H. (2021). AI research and data protection: Can the same rules apply for commercial and academic research under the GDPR? *Computer Law and Security Review*, 41, Artículo 105532. <https://doi.org/10.1016/j.clsr.2021.105532>

- Mikołajewska, E., Mikołajewski, D., Mikołajczyk, T., y Paczkowski, T. (2025). Generative AI in AI-based digital twins for fault diagnosis for predictive maintenance in Industry 4.0/5.0. *Applied Sciences*, 15(6), Artículo 3166. <https://doi.org/10.3390/app15063166>
- Milossi, M., Alexandropoulou-Egyptiadou, E., y Psannis, K. E. (2021). AI ethics: Algorithmic determinism or self-determination? The GDPR approach. *IEEE Access*, 9, 58455–58466. <https://doi.org/10.1109/ACCESS.2021.3072782>
- Minssen, T., Gerke, S., Aboy, M., Price, N., y Cohen, G. (2020). Regulatory responses to medical machine learning. *Journal of Law and the Biosciences*, 7(1), Artículo lsa002. <https://doi.org/10.1093/jlb/lsa002>
- Mühlhoff, R. (2023). Predictive privacy: Collective data protection in the context of artificial intelligence and big data. *Big Data and Society*, 10(1). <https://doi.org/10.1177/20539517231166886>
- Rizinski, M., Peshov, H., Mishev, K., Chitkushev, L. T., Vodenska, I., y Trajanov, D. (2022). Ethically responsible machine learning in fintech. *IEEE Access*, 10, 97531–97554. <https://doi.org/10.1109/ACCESS.2022.3202889>
- Truby, J., y Brown, R. (2021). Human digital thought clones: The Holy Grail of artificial intelligence for big data. *Information and Communications Technology Law*, 30(2), 140–168. <https://doi.org/10.1080/13600834.2020.1850174>
- Vidgen, R., Hindle, G., y Randolph, I. (2020). Exploring the ethical implications of business analytics with a business ethics canvas. *European Journal of Operational Research*, 281(3), 491–501. <https://doi.org/10.1016/j.ejor.2019.04.036>
- Vo, V., Chen, G., Aquino, Y. S. J., Carter, S. M., Do, Q. N., y Woode, M. E. (2023). Multi-stakeholder preferences for the use of artificial intelligence in healthcare: A systematic review and thematic analysis. *Social Science and Medicine*, 338, Artículo 116357. <https://doi.org/10.1016/j.socscimed.2023.116357>
- Wellbrock, W., Malinowska, M., y Ludin, D. (2025). Ethical implications and potential opportunities and risks of artificial intelligence in supply chain management. *Discover Sustainability*, 6(1). <https://doi.org/10.1007/s43621-025-01808->

Wu, W., Huang, T., y Gong, K. (2020). Ethical principles and governance technology development of AI in China. *Engineering*, 6(3), 302–309. <https://doi.org/10.1016/j.eng.2019.12.015>

Normativa y fuentes oficiales

Agencia Española de Protección de Datos (AEPD). (2020). Adecuación al RGPD de tratamientos que incorporan inteligencia artificial: Una introducción. <https://www.aepd.es/documento/adequacion-rgpd-ia.pdf>

Agencia Española de Protección de Datos (AEPD). (2021). Una aproximación a la inteligencia artificial desde la perspectiva de los derechos fundamentales. <https://www.aepd.es/documento/aproximacion-ia-derechos-fundamentales.pdf>

Comisión Europea. (2019). Directrices éticas para una IA fiable. Grupo de Expertos de Alto Nivel en Inteligencia Artificial. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

Comisión Europea. (2021). Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial). COM/2021/206 final. <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52021PC0206>

Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2019). Principios de la OCDE sobre inteligencia artificial. <https://oecd.ai/en/ai-principles>

Parlamento Europeo y Consejo de la Unión Europea. (2016). Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 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 (Reglamento General de Protección de Datos). Diario Oficial de la Unión Europea, L 119, 1–88. <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32016R0679>

Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Reglamento de Inteligencia Artificial). Diario Oficial de la Unión Europea, L, 2024/1689. <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689>

UNESCO. (2021). Recomendación sobre la ética de la inteligencia artificial. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

Otras fuentes

Elsevier. (2026). *Scopus* [base de datos]. <https://www.scopus.com>

IBM. (s.f.). *¿Qué es data governance?* <https://www.ibm.com/es-es/topics/data-governance>

Real Academia Española. (s.f.). *Inteligencia artificial*. Diccionario panhispánico del español jurídico. <https://dpej.rae.es/lema/inteligencia-artificial>