

Los derechos de autor frente a la Inteligencia Artificial emergente: una revisión bibliográfica

**Copyright faced with emerging Artificial Intelligence:
a literature review.**

José Esteban Gabarda Balaguer / OSICO Sociedad del Conocimiento | Avda. Poeta Llorente, 148, 46185, La Pobla de Vallbona – Valencia | España | <https://orcid.org/0009-0003-9251-426X> | gabarda@osicoplatform.com

Isabel Iniesta Alemán / Universidad Internacional de La Rioja | Av. de la Paz, 137, 26006, Logroño, La Rioja, España | <https://orcid.org/0000-0002-0127-3487> | isabel.iniesta@unir.net

Ángel Fernández Fernández / The Core School | Av. de España, 4, 28760 Tres Cantos, Madrid, España | <https://orcid.org/0000-0003-4264-7788> | amfernandez@thecoreschool.com

Resumen

La irrupción de tecnologías emergentes como la inteligencia artificial está reconfigurando, entre otras muchas cosas, el concepto de la autoría en las obras intelectuales o artísticas poniendo en tela de juicio las fronteras entre las creaciones puramente humanas y las generadas algorítmicamente. La investigación que aquí se presenta cumple el objetivo de ofrecer una síntesis narrativa del conocimiento actual sobre la conflictiva relación entre autores y desarrolladores de soluciones IA mediante la metodología de la revisión bibliográfica. Reúne, de esta forma, una visión panorámica de lo más reciente e interesante de entre lo publicado tanto en el entorno académico como en el profesional sobre este incipiente fenómeno. Esta selección, por consenso entre el equipo, ofrece una propuesta útil para comprender el estado actual de la cuestión que puede justificar y contextualizar nuevos proyectos sobre el tema. Como principal resultado destaca la diferente protección legal en Europa en comparación con la ambigüedad que se encuentra en los EEUU. Múltiples demandas han llevado a establecer acuerdos de colaboración o estrategias colaborativas para el entrenamiento de los modelos IA. También se ha de tener en cuenta los dos grandes modelos de desarrollo, propietarios o de código abierto, y sus implicaciones éticas. Es evidente interés por el uso de esta tecnología ha llegado para quedarse. El mundo académico ya lo está incorporado tanto en la docencia como en la investigación. Es necesario reflexionar para dar respuesta a los desafíos globales que plantea el desarrollo de la inteligencia artificial.

Palabras clave: inteligencia artificial, derechos de autor, creatividad, comunicación, innovación.

Abstract

The emergence of emerging technologies such as artificial intelligence is reshaping, among many other things, the concept of authorship in intellectual or artistic works, calling into question the boundaries between purely human creations and those generated algorithmically. The research presented here aims to provide a narrative synthesis of current knowledge on the conflictive relationship between authors and developers of AI solutions using the bibliographic review methodology. It thus brings together a panoramic view of the most recent and interesting publications in both the academic and professional spheres on this emerging phenomenon. This selection, agreed upon by the team, offers a useful proposal for understanding the current state of the issue, which can justify and contextualize new projects on the subject. The main finding is the difference in legal protection in Europe compared to the ambiguity found in the US. Multiple lawsuits have led to the establishment of collaboration agreements or collaborative strategies for training AI models. The two major development models, proprietary and open source, and their ethical implications must also be considered. There is clear interest in the use of this technology, which is here to stay. The academic world has already incorporated it into both teaching and research. Reflection is needed to respond to the global challenges posed by the development of artificial intelligence.

Keywords: artificial intelligence, copyright, creativity, communication, innovation.

1. Introducción

La irrupción de tecnologías emergentes, en particular la inteligencia artificial (IA), está generando un profundo impacto en múltiples dimensiones del desarrollo social, cultural, económico y tecnológico. La IA está transformando positivamente sectores clave como el sanitario, el industrial, el educativo, el medio ambiental o el de la cultura y el entretenimiento. Sin embargo, estas transformaciones también plantean importantes desafíos éticos y legales, especialmente en lo que respecta a la protección de cuestiones como la privacidad, la equidad algorítmica o el impacto de estas tecnologías sobre el empleo (Floridi y Cowsls, 2019). La utilización masiva de sistemas de IA compromete derechos fundamentales como la protección de los datos personales, la no discriminación y la autonomía individual. En este contexto, el despliegue de estas tecnologías debe abordarse bajo principios de gobernanza institucional sólidos, que promuevan una inteligencia artificial abierta, colaborativa, transparente y sostenible. Se requiere un marco regulatorio que garantice que los beneficios de estas tecnologías no comprometan los derechos de las personas, especialmente en contextos donde las garantías legales todavía son incipientes o inadecuadas.

El desarrollo acelerado de tecnologías emergentes, como la propia IA, el *blockchain*, los gemelos digitales, la biotecnología o la computación cuántica, está configurando un escenario inédito de transformación digital. Uno de los mayores desafíos derivados de este nuevo entorno es la redefinición del concepto de autoría. Mientras que el régimen tradicional de *copyright* se fundamenta en la protección de las expresiones originales de una idea, el uso creciente de sistemas de IA, capaces de generar textos, imágenes, música o incluso código informático, ha diluido las fronteras entre las creaciones humanas y las generadas algorítmicamente. Esto dificulta no solo la atribución de la autoría, sino también la determinación de los titulares de los derechos sobre dichas obras. Si bien la ley protege la forma en que se expresa una idea, no la idea en sí, la capacidad de la IA para replicar estilos, temáticas o estructuras previamente existentes sin intervención humana directa plantea un reto sin precedentes para los sistemas de propiedad intelectual. Uno de los principales focos de controversia es el uso de datos protegidos por derechos de autor para entrenar estos modelos, sin contar en muchos casos con el consentimiento explícito de los titulares de estos derechos.

La ausencia de un marco legal claro y actualizado ha provocado una oleada de demandas judiciales contra compañías desarrolladoras de IA, acusadas de infringir la legislación sobre propiedad intelectual (Barcott, 2023). La irrupción de estas tecnologías ha intensificado el debate sobre el equilibrio entre innovación y protección de la propiedad intelectual. Por un lado, la IA ha demostrado ser una herramienta útil para detectar infracciones de derechos de autor, automatizar procesos de supervisión y mejorar la gestión de contenidos. Por otro lado, la propia tecnología plantea nuevos desafíos, especialmente cuando se generan obras cuya originalidad y autoría resultan difíciles de atribuir.

2. Objetivos y metodología

Con el objetivo de ofrecer una síntesis del conocimiento actual sobre la problemática relación entre los autores y los desarrolladores de soluciones basadas en IA, se optó por la metodología de la revisión bibliográfica. Es este un enfoque muy habitual en las ciencias sociales (García García, 2022) y que permite explorar en profundidad un tema tan específico como es que aquí se desarrolla, proporcionando una síntesis narrativa (Guirao Goris, 2015; Farrukh & Sajjad, 2023) que incorpora tanto fuentes académicas como profesionales. De esta forma, el resultado ofrece una panorámica de lo publicado por los investigadores académicos contrastado con el punto de vista de los profesionales y de los organismos que están trabajando en la regulación de este fenómeno tan incipiente. La búsqueda de fuentes académicas indexadas en la *Web of Science Intelligence* se planteó con la siguiente *query*: *Artificial (Topic) or Artificial Intelligence (Topic) or Generative AI (Topic) or IA generative (Topic)*. De esta forma se llegó a 907,568 resultados desde el año 1900 a 2024. Refinando esta búsqueda para centrarnos en el tema que aquí se presenta y acotando para recoger únicamente las más recientes publicaciones, se llega a 302 documentos, de los cuales 169 son artículos. De ellos, tres son artículos de revisión, aunque ninguno de ellos aborda el tema de la forma en como aquí se plantea. Conscientes del riesgo de sesgo, la selección inicial de las fuentes se efectuó por separado entre los miembros del equipo de investigación, actuando así mutuamente como evaluadores de la calidad e idoneidad de estas para el objetivo de esta investigación. Finalmente, mediante consenso, se llegó a una selección de artículos en Open Acces. Fueron estas fuentes, junto con las publicaciones profesionales más recientes, la base del estudio que aquí se presenta. La utilidad de esta propuesta es la de integrar la información esencial sobre la problemática de los derechos de autor al ser utilizada en los sistemas de Inteligencia Artificial, con una perspectiva unitaria y coherente, ofreciendo una visión global del estado actual de la cuestión (Guirao Goris, 2015). De esta forma se ofrece una base para justificar y contextualizar nuevos proyectos de investigación (van Wee & Banister, 2023; Dhobi, 2024) al efectuar un análisis detallado y crítico de la literatura existente.

3. Resultados

3.1. Privacidad y protección de datos en la IA

La economía digital, con la aparición de la IA, ha transformado la creación, producción, distribución y explotación de obras y contenidos protegidos por derechos de autor. Alrededor de estas tecnologías han surgido nuevos modelos de negocio, pero también han aparecido nuevos actores. En esta constante evolución diversas normativas tratan de adaptar estos cambios que permitan avanzar en el proceso de desarrollo económico y social. Diferentes medidas de protección van surgiendo con el fin de proteger los derechos de autor en la era de las tecnologías emergentes.

La gobernanza digital puede definirse como el establecimiento de la coyuntura necesaria para alcanzar los objetivos, estrategias y acciones de gobierno con el fin de contribuir a un desarrollo económico, social y medioambiental sostenible y resiliente mediante el uso de las tecnologías emergentes. (Gabarda Balaguer, 2024a, p.250)

En el caso de España, el Real Decreto-ley 24/2021 (Gobierno de España, 2021) traspone la Directivas sobre derechos de autor y derechos afines en el mercado único digital. Se intenta buscar una regulación que encuentre el equilibrio entre el acceso a contenidos con la protección de la propiedad intelectual al mismo tiempo se trata de preservar la privacidad de las actuaciones en la era digital. El derecho a la privacidad y a la protección de datos personales es un derecho ampliamente reconocido por diferentes organismos internacionales, como la UE y la ONU, entre otros, que intentan llevar estos derechos a Internet. Con la irrupción de la inteligencia artificial, asistimos a un proceso de transformación sin precedentes, impulsado por el uso masivo de datos en los procesos de entrenamiento de modelos. Muchos de estos datos pueden contener información personal, lo que plantea importantes desafíos en materia de privacidad y protección de datos. En la Unión Europea (2018), la salvaguarda de los datos personales constituye una prioridad. De hecho, la Comisión Europea ha reafirmado su compromiso con la protección de la privacidad y los derechos fundamentales de los ciudadanos. La salvaguarda de la información personal se erige como un pilar fundamental en el desarrollo e implementación de sistemas de inteligencia artificial. Esta premisa encuentra su fundamento jurídico en el marco normativo de la Unión Europea, específicamente en el artículo 8, apartado 1, de la Carta de Derechos Fundamentales

y el artículo 16, apartado 1, del Tratado de Funcionamiento de la Unión Europea (Unión Europea, 2010; Hustinx, 2017). El ecosistema regulatorio europeo en materia de protección de datos se ha fortalecido significativamente mediante la implementación del Reglamento General de Protección de Datos (RGPD) y la Directiva (UE) 2016/680, instrumentos que establecen salvaguardas específicas para el procesamiento automatizado de datos personales en el contexto de la IA (De Hert & Papakonstantinou, 2016).

Según Wachter et al. (2017), estas normativas introducen disposiciones cruciales que abordan los desafíos emergentes en la intersección entre la IA y la privacidad. No obstante, diversos académicos señalan la necesidad de mayor precisión y desarrollo en el marco regulatorio. Como argumentan Veale y Edwards (2018), existe una urgencia por clarificar el alcance y la interpretación de las disposiciones legales relacionadas con la toma de decisiones automatizada, particularmente en contextos donde la IA desempeña un papel predominante. Un ejemplo de cómo se puede proteger los derechos de privacidad lo tenemos en lo que se conoce como «aprendizaje federado» (Nelson, 2020) a través de la inteligencia colaborativa. Así, esta técnica de aprendizaje automático permite entrenar modelos de IA en varios dispositivos o servidores manteniendo los datos en su ubicación original. Esta técnica que combina privacidad y descentralización garantiza la privacidad y evita la exposición de información confidencial. Una solución novedosa para empresas y organizaciones que requieren procesar grandes cantidades de datos sin violar las regulaciones derivadas de la privacidad.

3.2. Privacidad o interés legítimo en el entrenamiento de la IA

El rápido desarrollo de las herramientas IA generativa, en constante transformación, están generando una gran preocupación en relación con la privacidad y la seguridad. La evolución contemporánea de las tecnologías de inteligencia artificial ha suscitado un debate significativo en torno a la utilización de datos personales para el entrenamiento de modelos algorítmicos. En este contexto, las principales corporaciones tecnológicas han implementado estrategias de recopilación y procesamiento de datos que plantean interrogantes fundamentales sobre la privacidad y el consentimiento informado (Zuboff, 2019).

Meta Platforms Inc. inició, a mediados de 2024, la integración de datos multimodales provenientes de sus diversas plataformas sociales para el desarrollo de Meta AI, fundamentando esta práctica en el principio de «interés legítimo». Esta base jurídica, aunque reconocida en marcos regulatorios, presenta complejidades interpretativas significativas en el contexto del procesamiento automatizado de datos. La aplicación del «interés legítimo» como fundamento para el procesamiento masivo de datos personales requiere una evaluación exhaustiva de proporcionalidad y necesidad. Paralelamente, la transformación de Twitter en X Corp. y su implementación de políticas de recopilación de datos para el entrenamiento de Grok ejemplifica una tendencia hacia la automatización del consentimiento que podría contravenir principios fundamentales de autodeterminación informativa. Esta práctica se encuentra en un terreno jurídicamente controvertido, especialmente en jurisdicciones con marcos regulatorios robustos como la CCPA y el RGPD.

El ecosistema más amplio de desarrolladores de IA, incluyendo organizaciones como OpenAI, enfrenta desafíos similares en la búsqueda de equilibrio entre innovación y cumplimiento normativo. La ausencia de mecanismos efectivos de consentimiento y transparencia en el procesamiento de datos podría resultar en sanciones regulatorias significativas y en la erosión de la confianza pública. La cuestión es dilucidar si entrenar a la IA supone una invasión a nuestra privacidad o realmente es un uso justificado y amparado por el concepto de interés legítimo, como señala el Reglamento General de Protección de Datos para el tratamiento de estos.

Muchas cuestiones están surgiendo alrededor de cómo abordar los derechos de autor en el contexto de las IA. Desde cuáles deberían ser las condiciones idóneas para el uso de contenidos generados por IA, así como hacer que este abordaje sea adaptable y escalable en una tecnología en constante evolución, hasta como establecer un servicio de monitoreo que garantice que los datos de esa IA en el proceso del entrenamiento sean verdaderos y transparentes, pasando por conocer la fuente y la autoría de los mismos con condición que garantice la legalidad de la utilización de la IA en el contexto de los derechos de autor. En esta dirección las licencias para la utilización de la IA están siendo clave para encontrar un espacio que armonice y equilibre los derechos de autor entre los diferentes actores que participan en este proceso. Las licencias en este sentido deben ser no sólo justas para todas las partes involucradas en el contexto de las IA sino garantizar los derechos básicos de los derechos humanos,

colocando a las personas y a la tecnología en el sitio que le corresponde sacando partido en un desarrollo próspero y brillante. Estas licencias deben tener en cuenta que el proceso de entrenamiento de los datos sea correcto con el fin de que ayude a comprender de la mejor forma la fuente y la autoría de los contenidos que sean generados como consecuencia de esta tecnología. De ahí que la transparencia de estos contenidos recae en los programadores y las empresas que hay detrás de la IA, pero también obliga que la gobernanza global sobre esta materia sea la que impulse y garantice el despliegue de la eficacia de esta.

3.3. Acuerdos de colaboración

La ambigüedad legislativa en torno al uso de obras protegidas por derechos de autor para el entrenamiento de modelos de inteligencia artificial ha dado lugar a intensos debates jurídicos y a una oleada de demandas judiciales (Harris, 2024). En Estados Unidos, entre 2022 y 2024, se presentaron varias acciones colectivas contra herramientas como CoPilot y Stable Diffusion, por la presunta utilización indebida de imágenes con *copyright* en el entrenamiento de sus modelos generativos. Poco después, en 2024, guionistas y autores literarios interpusieron demandas similares contra OpenAI, Meta y Google, acusándoles de entrenar sus sistemas con obras literarias sin respetar los derechos de autor (Kahveci, 2024). También en la industria discográfica se han planteado este tipo de demandas y así, la Recording Industry Association of América (RIAA) en la que se integran sellos como Universal Music Group, Warner Music y Sony Music Group, ha emprendido acciones legales por vulneración de derechos de autor contra Udio y Suno, dos de los principales generadores de música basados en inteligencia artificial (Knibbs, 2024). Al mismo tiempo, en Europa, la Danske Pressepublikationer Kollektive Forvaltningsorganisation (DPCMO), que representa al 99 % de los editores daneses, exigió a OpenAI negociar de forma colectiva el uso de sus contenidos, advirtiendo de futuras acciones legales y desencadenando incluso la mediación del Ministerio de Cultura danés en la negociación (Rønde, 2024).

Todas estas acciones, que son solo una muestra de la controversia generada en torno a la utilización de contenidos protegidos por derechos de autor para el entrenamiento de modelos de IA generativa, se sustentan en argumentos jurídicos muy similares, basados en reclamaciones históricas sobre la infracción estos derechos, el incumplimiento contractual, la vulneración de marcas y la competencia desleal. En este contexto, las principales compañías desarrolladoras de herramientas de IA generativa están comenzando a establecer acuerdos de colaboración con grupos mediáticos, agencias de noticias y medios de comunicación a fin de obtener licencias para el uso de los contenidos que publican. Estos acuerdos permiten a estas compañías entrenar a sus modelos de IA e integrar contenido de calidad en sus productos, garantizando su atribución y, en algunos casos, retribuyendo a los autores.

Gemini, el modelo de lenguaje generativos de Google, apenas ofrece información pública sobre este tipo de acuerdos. No obstante, históricamente Google ha contado con acuerdos y relaciones de larga data con medios y agencias de noticias a través de servicios como Google News y del propio motor de búsqueda. Este marco contractual, complementado con prácticas propias de indexación y acceso a contenidos de fuentes autorizadas, le permite alimentar sus modelos con datos actualizados sin necesidad de establecer acuerdos de colaboración específicos para sus modelos generativos. Esto, no resulta contradictorio con el patente el interés de Google por mejorar la calidad de las respuestas mediante el uso de contenidos verificados y de calidad, así como su voluntad de desarrollar mecanismos estructurados de atribución, compensación y transparencia asociados al tratamiento de la información por parte de sus algoritmos.

De hecho, otras plataformas que forman parte del ecosistema de servicios de Google, como YouTube, también han establecido acuerdos con sellos discográficos, como Universal Music Group, para para investigar acerca de aplicaciones de inteligencia artificial generativa en la composición musical. Un ejemplo relevante de esta colaboración es YouTube Music AI Incubator, una iniciativa conjunta entre YouTube y Universal Music Group que funciona como espacio de experimentación donde artistas y compositores aportan conocimientos que orientan el desarrollo tecnológico. Esta iniciativa busca establecer marcos éticos y creativos para la implementación de sistemas de IA en la producción musical profesional (YouTube, 2023; Universal Music Group, 2023). Sin embargo, la implementación de sistemas de IA en la creación musical todavía enfrenta importantes obstáculos regulatorios. Un buen

ejemplo de ello es Orca, un modelo de inteligencia artificial generativa desarrollado en 2024 a partir de la colaboración entre Google DeepMind y YouTube, que permitía producir obras musicales completas a partir de *prompts* que incluyan artistas, letras o un género musical concreto. Este proyecto, que destacó por su tecnología de síntesis vocal, fue discontinuado tras disputas legales no resueltas acerca de la retribución de los derechos de autor (Business Insider, 2024).

Por su parte, Meta ha enfocado su estrategia en el desarrollo interno de modelos de inteligencia artificial basados en datos de dominio público y otras fuentes de información accesibles, lo que le ha permitido eludir la negociación de licencias específicas con medios de comunicación. Sin embargo, actualmente la compañía mantiene conversaciones con Apple para evaluar la posible integración de sus modelos de IA, como LLaMA, en el ecosistema de Apple Intelligence (González, 2024b; Rodríguez et al., 2024). Simultáneamente al desarrollo de estas iniciativas, han surgido distintas herramientas destinadas a salvaguardar los derechos de autor frente al uso indiscriminado de contenidos por parte de los modelos de IA generativa. Entre ellas destaca Nightshade (2024), una aplicación desarrollada por investigadores de la Universidad de Chicago, que convierte cualquier imagen en una muestra de datos no apta para el entrenamiento de modelos de inteligencia artificial. Como consecuencia de ello, cuando dichos modelos procesan estas imágenes, pueden producir salidas erróneas o «alucinaciones», lo que incrementa el coste del uso de datos y hace que recurrir a licencias legítimas resulte una opción más viable y eficiente para ellos.

Es indudable que, en este contexto, el diseño de estrategias colaborativas entre las empresas tecnológicas y los creadores aporta beneficios. A medida que avanzamos hacia el nuevo escenario impulsado por la inteligencia artificial, el respeto a los derechos de autor se vuelve más esencial. Solo reconociendo la propiedad intelectual de los creadores y de los medios y formalizando acuerdos de uso podremos asegurar la inclusión de contenidos veraces y actualizados, frenar la difusión de información falsa y consolidar un ecosistema sostenible basado en la intersección entre tecnología y medios. El caso de DeepSeek presenta particularidades relevantes, ya que se trata de un modelo de inteligencia artificial de código abierto desarrollado en China.

A diferencia de las plataformas occidentales de IA generativa, en el caso de DeepSeek no existen evidencias públicas ni comunicados oficiales que confirmen la firma de acuerdos similares. Esta ausencia de información debe interpretarse en el contexto de la limitada transparencia que caracteriza a muchas compañías tecnológicas chinas, especialmente en lo relativo a prácticas de entrenamiento de modelos y uso de datos protegidos. La falta de mecanismos de rendición de cuentas y la escasa presión regulatoria sobre el uso de obras con derechos de autor en dicho entorno plantean interrogantes adicionales sobre el cumplimiento de estándares internacionales en materia de propiedad intelectual.

Resulta evidente que, en un contexto como este, el desarrollo de estrategias de colaboración entre las compañías tecnológicas y los creadores supone un beneficio común para ambos. A medida que nos adentrarnos en el nuevo escenario que propicia el desarrollo de la inteligencia artificial, resulta más importante el respeto a los derechos de autor. Solo a través del reconocimiento de la propiedad intelectual de los creadores y de la firma de acuerdos podrá garantizarse la inclusión de contenidos veraces y actualizados, evitar la propagación de contenidos falsos y crear una relación más sostenible entre la tecnología y los medios.

3.4. Modelos de desarrollo de la IA

Los sistemas de IA generativa se pueden clasificar en dos grandes categorías: los basados en modelos propietarios y los de carácter abierto. Esta distinción no solo refleja diferencias técnicas, sino que también plantea numerosas implicaciones éticas, económicas y sociales. Los sistemas propietarios son desarrollados, mantenidos y comercializados por entidades privadas o instituciones, generalmente con ánimo de lucro y bajo una lógica de control. Esto implica que el código fuente, los datos de entrenamiento y las arquitecturas internas no están disponibles para el público, lo que limita la transparencia y la capacidad de auditoría externa. El acceso a estos modelos suele estar restringido mediante APIs comerciales o plataformas específicas. Ejemplo de este tipo de modelos son GPT-4 de OpenAI y Bard de Google (Devlin et al., 2018). Este tipo de modelos se caracterizan por sus altos niveles de rendimiento y sofisticación, ya que son el resultado de procesos de entrenamiento altamente costosos, tanto en términos computacionales como de capital humano. Sin embargo, su falta de transparencia genera algunos inconvenientes que es necesario considerar. Por ejemplo, estos sistemas

no ofrecen información detallada acerca de los datos utilizados para su entrenamiento, lo que impide evaluar adecuadamente los sesgos, riesgos de privacidad o posibles vulneraciones de los derechos de autor. Además, el uso de modelos propietarios favorece la dependencia tecnológica respecto de sus proveedores, planteando un riesgo de monopolización y favoreciendo la concentración de poder (Crawford, 2023). Esta forma de dependencia, conocida como *vendor lock-in*, limita la autonomía de los usuarios y restringe las posibilidades de adaptación de los modelos a contextos locales y necesidades específicas (Hintersdorf et al., 2023).

En contraposición, los modelos de código abierto, como DeepSeek, spaCy, TensorFlow y PyTorch, permiten el acceso libre y gratuito al código fuente, a la arquitectura del modelo y, en algunos casos, a los datos de entrenamiento. Esta transparencia favorece la revisión, la modificación y la reutilización de los modelos (Torres, 2020). La disponibilidad pública del código que sustenta a estos modelos permite auditarlos, detectar errores y sesgos y mejorar su robustez mediante el escrutinio colectivo, lo que resulta en el desarrollo de sistemas más confiables y efectivos. Además, al permitir su personalización para contextos locales o sectores especializados, resultan mucho más adaptables que los modelos propietarios).

El código abierto también ofrece ventajas claras en términos de democratización del acceso a la IA. Permite, por ejemplo, que investigadores, desarrolladores independientes y pequeñas empresas puedan experimentar, adaptar y mejorar los modelos sin incurrir en los altos costes asociados al uso de licencias comerciales. Esto dota a estos modelos de una marcada dimensión social. Proyectos como la biblioteca de código abierto *Hugging Face Transformer* son un buen ejemplo de cómo este enfoque fomenta el desarrollo de ecosistemas de innovación colaborativos y descentralizados (Wolf et al., 2020). Sin embargo, la apertura de estos modelos también plantea algunos desafíos. La implementación efectiva de modelos abiertos requiere recursos técnicos y humanos significativos, incluyendo infraestructuras de alto rendimiento (GPUs, clústeres de computación), y equipos altamente capacitados. Sin un modelo de financiación sostenible, estos proyectos pueden enfrentar dificultades para mantenerse a largo plazo.

Por otro lado, la apertura del código puede exponer vulnerabilidades que podrían ser explotadas por actores malintencionados. Distintos estudios han demostrado que los modelos de código abierto pueden ser susceptibles a ataques de envenenamiento de datos, al introducir información maliciosa destinada a manipular sus resultados (Wan et al., 2023). Estos sistemas también son susceptibles de la inclusión de los llamados «agentes durmientes», comportamientos maliciosos o instrucciones ocultas que permanecen latentes en el modelo y que solo se activan bajo determinadas condiciones específicas, como, por ejemplo, a creación de un *prompt* que incluya determinados términos (Axon, 2024; Edwards, 2024; Hubinger et al., 2024). Además, los modelos abiertos plantean el riesgo de posibles fugas de datos (Brundage et al., 2020) o su uso con fines ilícitos como la generación de desinformación o contenido ofensivo (Obioha-Val et al., 2025; Pearce et al., 2022). Algo que en los sistemas propietarios con un control centralizado resulta mucho más improbable. En realidad, actualmente la mayoría de los modelos abiertos no ofrecen una apertura total de sus sistemas. De hecho, aunque compañías importantes del sector tecnológico como Meta afirman ofrecer modelos de IA de código abierto, en realidad no revelan completamente elementos fundamentales de su estructura y funcionamiento:

No todos los modelos de IA que dicen ser de 'código abierto' son realmente abiertos y aunque muchos de los grandes modelos de lenguaje que impulsan los chatbots afirman ser abiertos, lo cierto es que restringen el acceso al código y a los datos de entrenamiento, además de que las empresas tecnológicas imponen una serie de observaciones de su utilización ya sean relacionadas con contar con algunas restricciones en cuanto a su uso se refiere así como gravar la compra de las licencias que muchas veces llevan aparejadas la restricción de acceso a ciertos servicios, que en caso de que los quieras utilizar, gravan todavía más el costo del servicio de su utilización. (Gibney, 2024a, p.)

Además, desde el punto de vista económico, cabe señalar que incluso los modelos abiertos requieren inversiones considerables. El entrenamiento de modelos de gran escala demanda enormes cantidades de energía, infraestructura computacional y talento técnico, lo que limita su desarrollo a un pequeño grupo de actores. A pesar de su potencial democratizador, es improbable que individuos o comunidades locales puedan desarrollar modelos de gran escala sin apoyo institucional o colaboraciones

internacionales. Desde una perspectiva ética, el debate entre modelos propietarios y abiertos tiene implicaciones complejas. Los modelos propietarios, por su naturaleza cerrada, limitan el acceso al conocimiento y concentran el poder tecnológico en unas pocas empresas. Esto puede reforzar desigualdades existentes y dificultar el control democrático sobre tecnologías que afectan a millones de personas. En cambio, los modelos de código abierto promueven un enfoque más inclusivo y participativo, alineado con principios de ciencia abierta. No obstante, su implementación ética requiere mecanismos de gobernanza adecuados, que garanticen el uso responsable de la tecnología, la protección de los derechos humanos y la rendición de cuentas en caso de abusos.

Cualquier contenido que se publica en internet desde la década de 1990 se convierte en un “software gratuito” que puede ser utilizado o copiado libremente por cualquier usuario. [...] [Gracias a la IA] vamos a reducir el coste de producción del conocimiento a cero. Para la gente esto es muy difícil de asimilar, pero en 15 o 20 años produciremos nuevo conocimiento científico y cultural a un coste marginal casi nulo. Será de código abierto y estará disponible para todos (González, 2024a, p.1).

Sin embargo, la transparencia en los modelos de código abierto no garantiza necesariamente la ética en su uso. En este sentido, es fundamental establecer marcos regulatorios y directrices claras que guíen el desarrollo y la implementación de estas tecnologías, asegurando que se respeten los derechos individuales y se minimicen los potenciales daños derivados de su mal uso. En realidad, podemos considerar la dicotomía entre modelos propietarios y modelos de código abierto en la IA generativa como un reflejo de tensiones más amplias en el desarrollo tecnológico contemporáneo. Si bien ambos enfoques tienen ventajas e inconvenientes, es fundamental establecer marcos éticos, legales y técnicos que aseguren un desarrollo equilibrado y responsable de la inteligencia artificial. Promover la transparencia, fomentar la cooperación global y garantizar el respeto a los derechos fundamentales deben ser objetivos centrales en el diseño y despliegue de estas tecnologías.

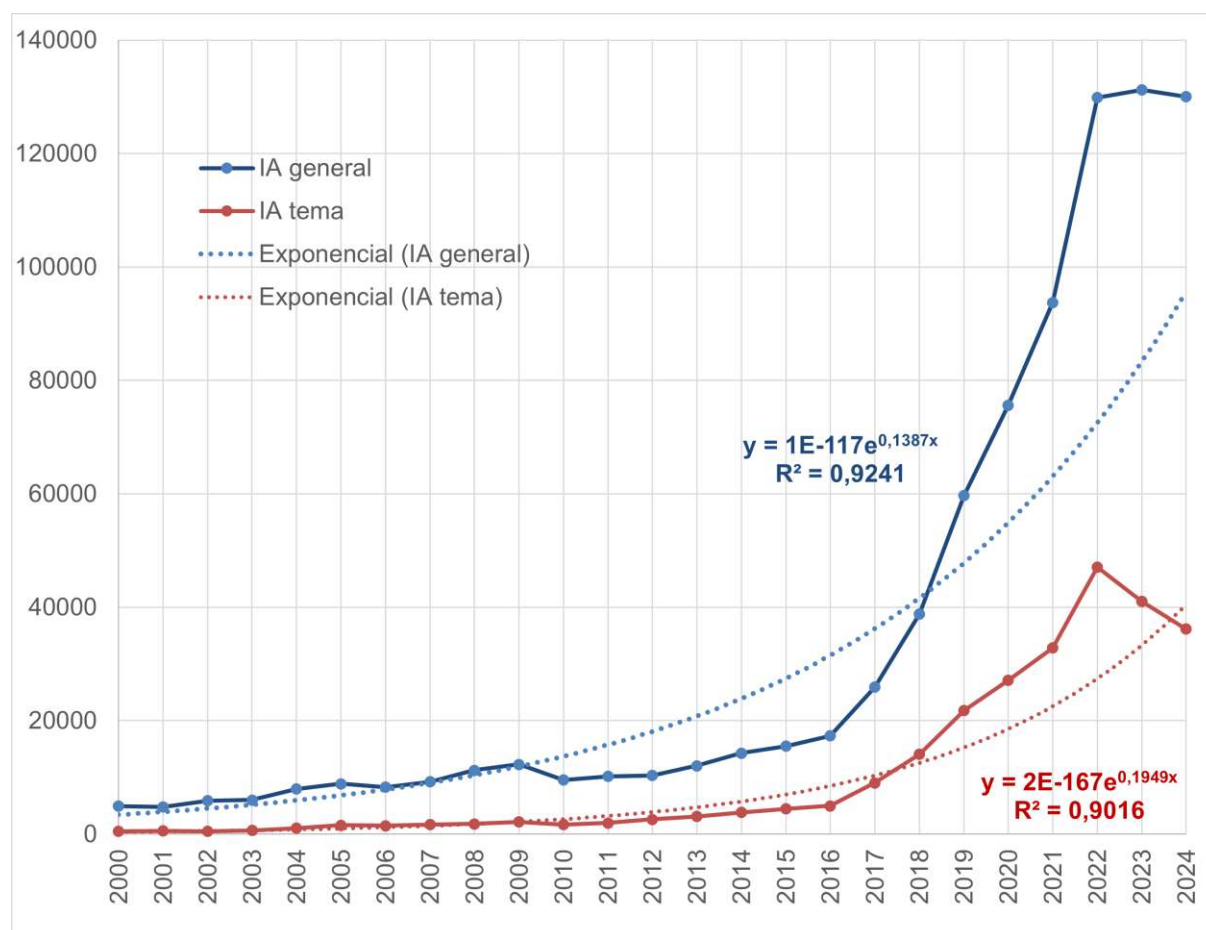
3.5. La IA en la publicación artículos académicos y científicos

En un terreno de juego donde se busca conocimiento e información de calidad es lógico que los artículos académicos y científicos estén en la diana de las empresas tecnológicas que querrán conseguir dicho conocimiento para entrenar sus modelos de IA. Pero ¿qué es lo que ocurre entonces con los derechos de autor?, ¿qué ocurre cuando los autores no saben lo que está pasando? Y, es más, cuando esto ocurra ¿qué es lo que va a significar después? La respuesta posiblemente sea múltiple, unos no se enterarán, otros querrán participar del beneficio económico juntamente con las editoriales, pero quizá haya otros que puedan plantear cuestiones de litigio sobre esta práctica. Las editoriales académicas están vendiendo el acceso a artículos de investigación a empresas tecnológicas para entrenar modelos de inteligencia artificial (IA).

Es más, ¿qué ocurre cuando ya se ha producido el entrenamiento para ver si esto se puede retrotraer para buscar otra solución? Si un artículo ya se ha utilizado como datos de entrenamiento en un modelo, no hay forma de eliminarlo después de que el modelo haya sido entrenado (Gibney, 2024b). La controversia o conflicto está garantizado: la IA complica la detección del plagio, lo que hará que se tengan que cambiar las reglas de juego de los derechos de autor en la IA en la publicación de artículos académicos y científicos, además de artículos de información en periódicos y otras obras.

Se observa que es a partir de 1990, coincidiendo con la generalización del uso de Internet, cuando el tema cobra mayor interés. Además, es ya en el siglo XXI cuando el interés académico tanto por la Inteligencia Artificial en general (IA general) como por la IA en un contexto de ciencias sociales (IA tema) presenta un crecimiento exponencial (ilustración 1). La Inteligencia Artificial está transformando las prácticas de investigación tradicionales, particularmente en el ámbito de la revisión de la literatura (Pandya, 2023). Y es que el buen uso de la IA puede mejorar la destreza, la eficiencia, la precisión y la profundidad en la revisión de la literatura existente. De hecho, se observa el creciente interés de las universidades por incorporar las destrezas en el uso de la IA en las prácticas de sus estudiantes, quizá conscientes de la imposibilidad de «poner puertas al campo» cuando las empresas están demandando como una competencia deseable la capacidad para su utilización en, cada vez más, los procesos de trabajo.

Ilustración 1 Evolución de las publicaciones sobre Inteligencia Artificial, específicamente en Ciencias Sociales.



Fuente: elaboración propia.

Por otro lado, la inteligencia artificial contribuye a la elaboración de materiales educativos que se caracterizan por su claridad y coherencia, lo cual se evidencia al evaluar el desempeño y las necesidades del estudiantado universitario. Esto permite ajustar los contenidos según su nivel de comprensión y corregir posibles errores. Asimismo, la IA favorece una enseñanza personalizada al adaptar las narrativas a los intereses y estilos de aprendizaje de cada persona, lo que incrementa su implicación y sentido de pertinencia. La integración de recursos basados en IA en entornos virtuales de aprendizaje, como los sMOOC, enriquece el proceso formativo al hacerlo más dinámico e interactivo (Gil-Quintana et al., 2025).

4. Conclusiones

Esta investigación presenta los claroscuros del gran salto tecnológico que se presenta al centrarnos en la autoría de las nuevas obras en las que la IA está involucrada. Todo conocimiento se ha basado en lo que la humanidad ha aprendido de sus ancestros: También es cierto que la tecnología desde que apareció la imprenta, pasando por las calculadoras y los ordenadores, han acelerado el proceso de aprendizaje y lo ha acercado a capas sociales cada vez más amplias. Sin embargo, las bibliotecas y los museos, templos sagrados del conocimiento firmado por los autores de cada obra, han sido sustituidos por algoritmos que se han dado en llamar Inteligencia Artificial.

Es necesario reflexionar, por lo tanto, sobre la capacidad de los marcos de gobernanza existentes para dar respuesta a los desafíos globales que plantea el desarrollo de la inteligencia artificial. Más allá de

la mera gestión de riesgos, una gobernanza eficaz se perfila como un elemento clave para orientar la innovación tecnológica hacia objetivos de desarrollo sostenible y bienestar colectivo. En efecto, la gobernanza puede convertirse en el verdadero motor de una transformación digital responsable, siempre que se fundamente en principios de transparencia, rendición de cuentas, inclusión y cooperación internacional. Resulta imprescindible repensar en profundidad los marcos normativos existentes, especialmente en lo que respecta a la propiedad intelectual.

En definitiva, el desarrollo de sistemas de inteligencia artificial avanzados no solo transforma radicalmente el acceso, creación y distribución del conocimiento, sino que también exige una gobernanza multinivel que combine regulación, ética y participación social. En un escenario en el que los modelos pueden replicar estilos, voces y todo tipo de contenidos con una enorme precisión, el desafío no es únicamente técnico, sino profundamente ético, político y normativo. Hay que asegurar que el avance de la inteligencia artificial beneficie a la sociedad en su conjunto y no solo a los actores con mayor capacidad tecnológica o financiera. Esa es, sin duda, una de las grandes tareas de nuestro tiempo.

7. Declaración sobre la contribución específica de cada una de las autorías, según la taxonomía CrediT

Concepción y enfoque del trabajo	Autores 1 y 2
Curación de datos	Autores 2 y 3
Análisis formal	Autores 1, 2 y 3
Adquisición de fondos	No aplica
Investigación	Autores 1, 2 y 3
Metodología	Autor 2
Administración del proyecto	Autor 2
Recursos	No aplica
Software	No aplica
Supervisión	Autor 1
Visualización	Autores 1, 2 y 3
Redacción	Autores 1, 2 y 3

Semblanza de los/as autores/as

José Esteban Gabarda Balaguer, experto en desarrollo local, se ha desempeñado como Director de Agencia de Desarrollo en diversas áreas, destacándose por su liderazgo visionario. Como fundador y director de la Plataforma OSICO, impulsa un espacio innovador para el análisis, la reflexión y la colaboración frente a los principales desafíos sociales actuales. Su enfoque combina la inteligencia colectiva, el pensamiento circular y sistémico, y la innovación social como herramientas clave para generar cambios transformadores en la sociedad, la economía y la política. Su compromiso y dedicación se manifiestan en su habilidad para liderar procesos estratégicos que fomentan un desarrollo sostenible y equitativo.

Isabel Iniesta Alemán, Doctora en Comunicación por la Universidad de Málaga y Licenciada en Ciencias Empresariales por la Universidad de Zaragoza es fundadora de MarketReal Consultores S.L, especialista en Marketing y consultora en emprendimiento homologada por la Diputación General de Aragón. Con un sexenio de investigación y certificada en ANECA como Contratada Doctor, imparte docencia en la Universidad Internacional de La Rioja además de formar parte activa en varios grupos de investigación y de la asociación científica ICONO 14.

Ángel Fernández Fernández es Vicedecano de Profesorado en The Core School. Doctor en Comunicación por la Universidad Europea y licenciado en Bellas Artes por la Universidad Complutense de Madrid, su línea de investigación se centra en el estudio de las narrativas visuales basadas en datos y en el impacto cultural de las tecnologías de la información y de la comunicación. A lo largo de su

trayectoria docente, ha impartido asignaturas relacionadas con la visualización de datos y la estética y cultura digitales.

Referencias bibliográficas

- Apple. (2024). Introducing Apple Intelligence for iPhone, iPad, and Mac. *Apple Newsroom*. <https://www.apple.com/newsroom/2024/06/introducing-apple-intelligence-for-iphone-ipad-and-mac/>
- Axon, S. (2024). Anthropic publicly releases AI tool that can take over the user's mouse cursor. *Ars Technica*: <https://arstechnica.com/ai/2024/10/anthropic-publicly-releases-ai-tool-that-can-take-over-the-users-mouse-cursor/>
- Barcott, B. (2023). AI Lawsuits Worth Watching: A Curated Guide. *Tech Policy Press*. <https://techpolicy.press/ai-lawsuits-worth-watching-a-curated-guide>
- Business Insider (2024). Google DeepMind's Orca AI music tool faces copyright issues. Business Insider: <https://www.businessinsider.com/google-deepmind-orca-ai-music-tool-copyright-issues-2024-11>
- Brundage, M. et al. (2020). Toward trustworthy AI development: Mechanisms for supporting verifiable claims. *arXiv*. <https://arxiv.org/abs/2004.07213>
- Crawford, K. (2023). *Atlas de IA: Poder, política y costes planetarios de la inteligencia artificial* (Vol. 2089). Ned ediciones.
- de Hert, P., & Papakonstantinou, V. (2016). The new General Data Protection Regulation: Still a sound system for the protection of individuals? *Computer Law & Security Review*, 32(2), 179-194. <https://doi.org/10.1016/j.clsr.2016.02.006>
- Devlin, J. et al. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv*. <https://arxiv.org/abs/1810.04805>
- Dhobi, S. (2024). Engaging with Review of Literature in Research. *Patan Prospective Journal*, 110-117. <https://doi.org/10.3126/ppj.v4i01.70207>
- Edwards, B. (2024). Los modelos abiertos de IA podrían contener códigos maliciosos, advierte Anthropic. *Wired*: <https://es.wired.com/articulos/modelos-abiertos-de-ia-podrian-contener-codigos-maliciosos-advier-anthropic>
- Farrukh, A., & Sajjad, A. (2023). A Critical Review of Literature Review Methodologies. Em S. S. Rana, Advancing Methodologies of Conducting Literature Review in Management Domain (Review of Management Literature, Vol. 2) (pp. 103-123). *Emerald Publishing Limited*. <https://doi.org/10.1108/S2754-586520230000002006>
- Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- Gabarda Balaguer, J. E. (2022). Interview with Asker Bryild Staunaes of the Danish Synthetic Party. Plataforma Osico: <https://www.osicoplatform.com/interview-with-asker-bryild-staunaes-of-the-danish-synthetic-party/>
- Gabarda Balaguer, J. E. (2024). *Gobernanza, innovación social y tecnología*. Ediciones OSICO.
- García García, M. (2022). Systematic Literature Review Applied to a Research Project. Em M. V. Carrillo Dura, & M. Pérez Pulido, Cases on Developing Effective Research Plans for Communications and Information Science (pp. 45-60). *IGI Global*. <http://dx.doi.org/10.4018/978-1-6684-4523-5.ch003>
- Gibney, E. (2024). Has your paper been used to train an AI model? Almost certainly. *Nature*, 632(8026), 715-716. *Nature*: <https://www.nature.com/articles/d41586-024-02599-9>

- Gibney, E. (2024). Not all 'open source' AI models are actually open: here's a ranking. *Nature*. <https://doi.org/10.1038/d41586-024-02012-5>
- Gil Quintana, J., García Blazquez, E., Hueso Romero, J. J., & Cantillo Valero, C. (2025). Diseño y análisis de narrativas enriquecidas con inteligencia artificial. Proyecto de transferencia UNED - UTMACH. *RIED-Revista Iberoamericana de Educación a Distancia*, 28(2). <https://doi.org/10.5944/ried.28.2.43305>
- Gobierno de España. (2021). Real Decreto-ley 24/2021, de 2 de noviembre, de transposición de directivas de la Unión Europea en las materias de bonos garantizados, distribución transfronteriza de organismos de inversión colectiva, datos abiertos y reutilización de la información del. <https://www.boe.es/eli/es/rdl/2021/11/02/24>
- González, F. (2024a). "Cualquiera puede copiar, recrear o reproducir el contenido de la web": director de IA de Microsoft. *Wired*: <https://es.wired.com/articulos/cualquiera-puede-copiar-contenido-de-la-web-director-de-ia-de-microsoft>
- González, F. (2024b). La IA de Meta podría llegar a la nueva Apple Intelligence de los iPhone. *Wired*: <https://es.wired.com/articulos/ia-de-meta-podria-llegar-la-nueva-apple-intelligence-de-los-iphone>
- Guirao Goris, S. J. (2015). Utilidad y tipos de revisión de literatura. *Ene, Revista de enfermería* 9(2). <https://dx.doi.org/10.4321/S1988-348X2015000200002>
- Harris, B. (2024). Legal challenges surrounding generative AI models. *Tech Policy Press*. <https://techpolicy.press/legal-challenges-surrounding-generative-ai/>
- Hintersdorf, D., Struppek, L., & Kersting, K. (2023). Balancing Transparency and Risk: The Security and Privacy Risks of Open-Source Machine Learning Models. *arXiv*. <https://arxiv.org/abs/2308.09490>
- Hubinger, E. et al. (2024). Sleeper agents: Training deceptive LLMs that persist through safety training. *arXiv*. <https://arxiv.org/abs/2401.05566>
- Hustinx, P. (2013). EU data protection law: The review of directive 95/46/EC and the proposed general data protection regulation. University of Tartu. Data Protection Inspectorate, Tallinn. <https://edps.europa.eu/sites/default/files/publication/14-09-15>
- Knibbs, K. (2024). Discográficas demandan a los generadores de música con IA Suno y Udio por violar derechos de autor. *Wired*: <https://es.wired.com/articulos/discograficas-demandan-a-generadores-de-musica-con-ia-suno-y-udio-por-violar-derechos-de-autor>
- Meaker, M. (2024). La prensa de Dinamarca amenaza con demandar a OpenAI por derechos de autor. *Wired*: <https://es.wired.com/articulos/prensa-de-dinamarca-amenaza-con-demandar-a-openai-por-derechos-de-autor>
- Miranda, L. (2024). YouTube negocia con la industria musical para generar canciones con IA. *Hipertextual*: <https://hipertextual.com/2024/06/youtube-negocia-industria-musical-generador-canciones-ia>
- Nelson, D. (2020). ¿Qué es el aprendizaje federado? *Unite.Ai*: <https://www.unite.ai/es/que-es-el-aprendizaje-federado/>
- Nightshade. (2024). What Is Nightshade? Obtido em 30 de 12 de 2024, de *Nightshade*: <https://nightshade.cs.uchicago.edu/whatis.html>
- Obioha-Val, O. A., Lawal, T. I., Olaniyi, O. O., Gbadebo, M. O., & Olisa, A. O. (2025). Investigating the Feasibility and Risks of Leveraging Artificial Intelligence and Open Source Intelligence to Manage Pearce, H. et al. (2022). *Asleep at the keyboard? Assessing the security of GitHub Copilot's code contributions*. IEEE Symposium on Security and Privacy
- Predictive Cyber Threat Models. *Journal of Engineering Research and Reports*, 27(2), 10–28. <https://doi.org/10.9734/jerr/2025/v27i21390>
- Pandya, T. P. (2023). Literature Review as Research Methodology: Aid of Ai (Artificial Intelligence). *Vidhyayana - An International Multidisciplinary Peer-Reviewed E-Journal*, 9(si1), 123-133. <https://doi.org/10.58213/vidhyayana.v9isi1.1589>

- Rebollo, C. (2024). De la regulación de la IA al tratamiento contra el VIH: los eventos que dieron forma a la ciencia en 2024. *Wired*: <https://es.wired.com/articulos/2024-eventos-que-dieron-forma-a-la-ciencia-y-la-tecnologia>
- Rodríguez, S., Tilley, A., & Kruppa, M. (2024). Apple, Meta Have Discussed an AI Partnership. *The Wall Street Journal*: <https://www.wsj.com/tech/ai/apple-meta-have-discussed-an-ai-partnership-cc57437e?st=tdq3k6copeubhe>
- Rønde, K. (2024). DPCMO Threatens to Sue OpenAI. *Danske Pressepublikationer Kollektive Forvaltningsorganisation*. <https://dpcmo.dk/dpcmo-threatens-to-sue-openai/>
- Torres, J. (2020). *Python deep learning: Introducción práctica con Keras y TensorFlow 2*. Alpha Editorial.
- Unión Europea. (2010). Versión consolidada del Tratado de Funcionamiento de la Unión Europea. *Diario Oficial de la Unión Europea*: <https://www.boe.es/doi/2010/083/Z00047-00199.pdf>
- Unión Europea. (2018). Reglamento (UE) 2018/1725 del Parlamento Europeo y del Consejo, de 23 de octubre de 2018, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por las instituciones, órganos y organismos de la Unión, y al. Boletín Oficial del Estado. *Boletín Oficial del Estado*: <https://www.boe.es/buscar/doc.php?id=DOUE-L-2018-81849>
- Universal Music Group. (2023, September 21). YouTube announces AI music principles and launches YouTube Music AI Incubator with artists, songwriters and producers from Universal Music Group. *Universal Music Group*. <https://bit.ly/3ECMkbE>
- van Wee, B., & Banister, D. (2023). Literature review papers: the search and selection process. *Journal of Decision Systems*, 33(4), 559-565. <https://doi.org/10.1080/12460125.2023.2197703>
- Veale, M., & Edwards, L. (2018). Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling. *Computer Law & Security Review*, 34(2), 398-404. <https://doi.org/10.1016/j.clsr.2017.12.002>
- 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://bit.ly/4a4h5BG>
- Wan, A., Wallace, E., Shen, S., & Klein, D. (2023). Poisoning Language Models During Instruction Tuning. *arXiv*. <https://arxiv.org/abs/2305.00944>
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., Drame, M., Lhoest, Q., & Rush, A. (2020). Transformers: Procesamiento del lenguaje natural de última generación. En *Actas de la Conferencia de 2020 sobre métodos empíricos en el procesamiento del lenguaje natural: Demostraciones de sistemas* (pp. 38-45). *Asociación de Lingüística Computacional*. <https://www.aclweb.org/anthology/2020.emnlp-demos.6>
- YouTube. (2023, August 21). Partnering with the music industry on AI. YouTube Blog: <https://blog.youtube/inside-youtube/partnering-with-the-music-industry-on-ai/>
- Zuboff, S. (2019). Surveillance capitalism and the challenge of collective action. *New labor forum*, 28(1), 10-29. <https://doi.org/10.1177/1095796018819461>