

Potencialidades de la Inteligencia Artificial en la difusión de la Ciencia a través de repositorios institucionales

Potential of Artificial Intelligence in the dissemination of Science through Institutional Repositories

Jorge Caldera Serrano

jcalser@unex.es

Departamento de Información y Comunicación, Universidad de Extremadura (España)

Resumen

Objetivo. Identificar las posibles aplicaciones de la Inteligencia Artificial en la implementación, mantenimiento y difusión de los repositorios públicos institucionales, en el marco de la filosofía de Ciencia Abierta. Método. Para ello se ha realizado una revisión bibliográfica de la literatura existente, realizando para ello consultas en WoS, Scopus, Google Scholar y ChatGPT. Resultados. Se enumeran y describen diferentes aplicaciones en la cual la IA puede ser utilizada para la mejora de los repositorios, tales como automatización de procesos internos, difusión de la información científica, transparencia y accesibilidad, evaluación social de la Ciencia, recomendaciones personalizadas, mejora en traducción, integración e interoperabilidad, mantenimiento predictivo, colaboración interdisciplinar, gestión de datos de investigación y plagio y fraude. Conclusiones. El uso de la IA mejora la experiencia como gestor, investigador y usuario de los repositorios institucionales, aportando transparencia en todos los procesos y garantizando una mejor y mayor difusión de las investigaciones en él contenidas.

Palabras clave

Ciencia abierta; Repositorios institucionales; Inteligencia Artificial; Difusión de la Ciencia; Automatización de procesos.

Abstract

Objective. To identify the potential applications of Artificial Intelligence in the implementation, maintenance, and dissemination of public institutional repositories, within the framework of the Open Science philosophy. **Method.** For this purpose, a bibliographic review of the existing literature was conducted, consulting WoS, Scopus, Google Scholar and ChatGPT. **Results.** Various applications in which AI can be used to improve repositories are listed and described, such as automation of internal processes, dissemination of scientific information, transparency and accessibility, social evaluation of Science, personalized recommendations, improvement in translation, integration and interoperability, predictive maintenance, interdisciplinary collaboration, research data management, and plagiarism and fraud detection. **Conclusions.** The use of AI enhances the experience of managers, researchers, and users of institutional repositories, providing transparency in all processes and ensuring better and broader dissemination of the research contained therein.

Keywords

Keywords: Open Science; Institutional repositories; Artificial Intelligence; Dissemination of Science; process automation

Recibido: 7/08/2025

Aceptado: 11/11/2025

DOI: <https://dx.doi.org/10.5557/IIMEI16-N31-026046> IIMEI16-N31-026046

Descripción propuesta: Caldera Serrano, Jorge, 2025. Potencialidades de la Inteligencia Artificial en la difusión de la Ciencia a través de repositorios institucionales. *Métodos de Información*, **16**(31), 26-46.

1. Introducción

La ciencia abierta es un movimiento global que busca hacer que la investigación científica sea accesible a todos, sin barreras económicas, legales o técnicas. Este concepto ha ganado prominencia en los últimos años, promovido por el reconocimiento de que el acceso abierto a los datos y publicaciones científicas puede acelerar el avance del conocimiento. La idea fundamental es que la ciencia debe ser transparente y accesible, facilitando la colaboración y el intercambio de información entre investigadores y con la sociedad en general (Abadal, 2021) (Abadal & Anglada, 2020).

El origen de la ciencia abierta se puede rastrear hasta los primeros días de la Internet, cuando los científicos comenzaron a compartir sus trabajos en

línea para llegar a una audiencia más amplia. Sin embargo, el movimiento ganó fuerza con el establecimiento del Proyecto de Acceso Abierto de Budapest en 2002, que formalizó sus principios. Este proyecto se centró en proporcionar acceso libre a la literatura científica revisada por pares, permitiendo a cualquier persona leer, descargar, copiar, distribuir, imprimir, buscar o enlazar los textos completos de estos artículos sin restricciones legales o financieras (BOAI, 2002) (Abadal & Anglada, 2020).

El desarrollo de repositorios institucionales ha sido crucial en la promoción de la ciencia abierta. Estos repositorios son plataformas digitales gestionadas por instituciones académicas o de investigación que almacenan y proporcionan acceso a las publicaciones científicas producidas por sus miembros. Los repositorios institucionales facilitan el acceso abierto a la investigación, incrementando su visibilidad y citación. Además, aseguran la preservación a largo plazo de la producción científica y garantizan que los resultados de la investigación financiada con fondos públicos sean accesibles a la sociedad (Suber, 2012) (Caldera-Serrano, 2018).

Un aspecto esencial de los repositorios institucionales es su capacidad para alojar una variedad de contenidos, incluyendo artículos de revistas, tesis, disertaciones, datos de investigación, informes técnicos y otros tipos de documentos académicos (Babine, et al., 2010). Estos repositorios no solo amplían el acceso al conocimiento, sino que también promueven la transparencia y la reproducibilidad de la investigación científica. La interoperabilidad entre diferentes repositorios a través de estándares abiertos permite una mayor difusión y reutilización del conocimiento científico a nivel global.

La Inteligencia Artificial (IA) está jugando un papel cada vez más importante en el avance de la ciencia abierta. La IA se utiliza para analizar grandes volúmenes de datos, identificar patrones y generar nuevas hipótesis de investigación. En el contexto de la difusión de la ciencia, la IA puede mejorar la eficiencia y la efectividad de los repositorios institucionales de varias maneras (Bethencourt-Aguilar, et al., 2022). Por ejemplo, los algoritmos de IA pueden ayudar a etiquetar y clasificar automáticamente los documentos científicos, facilitando su búsqueda y recuperación (Bourne, Lorsch, & Green, 2015).

Además, la IA puede asistir en la identificación de investigaciones relevantes y en la recomendación de artículos a los investigadores,

personalizando el acceso a la información científica según los intereses y necesidades de cada usuario. Esta capacidad de recomendación personalizada puede aumentar la visibilidad de trabajos específicos y facilitar la colaboración entre investigadores con intereses comunes.

Otro uso significativo de la IA en la ciencia abierta es la minería de textos y datos, que permite extraer información relevante de grandes conjuntos de datos no estructurados. Esta técnica puede ser especialmente útil en la revisión de la literatura científica, permitiendo a los investigadores identificar rápidamente tendencias y brechas en el conocimiento. La minería de textos y datos también puede facilitar la meta-análisis, permitiendo una síntesis más efectiva de los resultados de múltiples estudios (Ananiadou et al., 2010).

La IA también puede mejorar la transparencia y reproducibilidad de la investigación científica mediante el desarrollo de herramientas que ayuden a los investigadores a documentar y compartir sus procesos de análisis de datos. Estas herramientas pueden generar automáticamente registros detallados de los pasos seguidos en un experimento, lo que facilita la replicación de los estudios por otros investigadores. La transparencia en los procesos de análisis de datos es fundamental para la confianza en los resultados científicos y para el avance del conocimiento.

La ciencia abierta representa un cambio paradigmático en la forma en que se produce y se comparte el conocimiento científico. Los repositorios institucionales desempeñan un papel central en este movimiento, al proporcionar acceso abierto a la producción científica y al fomentar la transparencia y la colaboración. La IA está potenciando este avance al mejorar la gestión, búsqueda y análisis de la información científica. La combinación de ciencia abierta, repositorios institucionales e IA promete transformar la investigación científica, haciéndola más accesible, eficiente y colaborativa.

2. Revisión de literatura

La ciencia abierta, los repositorios institucionales y la IA son temas de gran relevancia en la investigación actual. Aportamos una serie de documentos básicos que conforman la base de nuestro trabajo, desde el punto de vista conceptual. Para ello exponemos trabajos relacionados con la ciencia abierta y específicamente con los repositorios institucionales, así como sobre IA, partiendo de la base del carácter acumulativo de la Ciencia y de la imposibilidad

de referenciar ni nombrar el gran volumen de trabajos relacionados con estas temáticas.

En el ámbito de la ciencia abierta destacamos inicialmente el trabajo de Fecher y Friesike (2014), que facilita un recorrido y visión integral sobre ciencia abierta, tratando la cuestión desde perspectivas tales como infraestructura, publicación, demarcación, impacto y medida. Apuesta claramente por un concepto aún en formación, no estático, pero que cuenta con el objetivo de hacer la Ciencia más accesible y colaborativa. Vicente-Saez y Martínez-Fuentes (2018) hace una revisión sistemática sobre ciencia abierta facilitando un concepto integrador conformado por el acceso abierto, datos abiertos y colaboración. Tennant et al. (2016) analiza el impacto de la filosofía del acceso abierto en la difusión de información desde diferentes variables, como la académica, económica y social, señalando sus beneficios en torno a la visibilidad e impacto de la investigación.

Centrándonos en los trabajos que versan sobre Repositorios, siempre orientados a un acceso libre y una preservación garantizada, el trabajo de Davis & Connolly (2007) analiza cómo mejorar los repositorios para tener mayor visibilidad y la estandarización de su uso. Pinfield y Bath (2017) analiza sus beneficios desde el punto de vistas de los costes de publicación y de los repositorios. Martínez et al. (2023) analiza el desarrollo de los repositorios y su impacto, mientras (Caldera-Serrano, 2019) (Abadal, 2021) se centra en los desafíos, incluyendo la necesidad de contar con políticas que lo potencie y facilite. Otros trabajos Álvarez, 2017) se centran en zonas geográficas concretas, analizando el papel relacionado con la preservación digital, defendiendo la necesidad de potenciar recursos y mejorar políticas. Martínez et al. (2023) llevan a cabo un estudio sobre la implementación de repositorios en universidades, sin lugar a dudas uno de sus principales actores, y la FECYT (2021) desarrolla una guía para la evaluación de repositorios de investigación.

Las IA ha revolucionado numerosos campos, incluyendo la gestión y difusión del conocimiento científico. El libro de Russell y Norvig (2020) es un documento fundamental para contar con una visión completa de los principios y aplicaciones de la IA. Jordan y Mitchell (2015) nos habla de tendencias y perspectivas del aprendizaje automático, mientras que Lecún, Bengio y Hinton (2015) como Goodfellow, Bengio y Courville (2016) se centran en el aprendizaje profundo.

En su adaptación o versión de la IA para La integración de la IA en los repositorios institucionales y la ciencia abierta, identificar que aún un es un área emergente con gran potencial para mejorar la gestión y el acceso al conocimiento científico. Faltan estudios críticos que analicen cómo la IA se relaciona con los datos abiertos y como puede facilitar la colaboración interdisciplinaria, ni existe un conjunto de literatura científica sobre herramientas y técnicas para la mejora de repositorios utilizando IA. Alonso-Arevalo & Quinde-Cordero (2024). Por lo tanto, falta aún desarrollo de una masa crítica científica que relacione la ciencia abierta, y más concretamente los repositorios públicos e institucionales, con la IA, cuestión fácil de entender ya que han sido en los últimos tres o cuatro años, cuando la IA ha contado con un desarrollo computacional como para ser implementado para disciplinas que, a priori, no son prioritarias para las grandes corporaciones que desarrollan los algoritmos. También se destaca la Declaración (Penabad-Camacho et al., 2024) donde se llevan a cabo principios sobre el uso de la IA para la edición científica.

3. Metodología

La IA y la ciencia abierta son dos conceptos que han ganado relevancia significativa en la investigación científica. La IA se define como el campo de estudio que busca desarrollar sistemas que puedan realizar tareas que requieren inteligencia humana, tales como el aprendizaje, la percepción y la toma de decisiones (Russell & Norvig, 2020). La ciencia abierta, por otro lado, es un movimiento que promueve la accesibilidad total y gratuita de los resultados y procesos científicos, abogando por la transparencia y la colaboración en la investigación. Los repositorios institucionales son plataformas digitales gestionadas por universidades u organizaciones de investigación que almacenan, preservan y distribuyen el output académico de estas instituciones, incluyendo artículos de investigación, tesis y datos de investigación. La relación entre estos tres conceptos se basa en su objetivo común de mejorar la accesibilidad y eficiencia en la producción y distribución del conocimiento científico.

El objetivo de este análisis es identificar cómo la IA puede contribuir a mejorar los fines de los repositorios institucionales para la difusión de la ciencia abierta. Para alcanzar este objetivo, se utilizará una metodología científica basada en la revisión bibliográfica y la observación de repositorios institucionales. Esta metodología permitirá una comprensión exhaustiva de las

intersecciones entre la IA, la ciencia abierta y los repositorios institucionales, proporcionando un marco para implementar mejoras tecnológicas en estos sistemas.

La revisión bibliográfica incluirá una selección de artículos científicos recientes que exploren la integración de la IA en la gestión de repositorios institucionales, así como aquellos que analizan las diferentes variables por separado: ciencia abierta, repositorios institucionales e IA. Se ha interrogado diferentes bases de datos (Web of Science, Scopus) y se ha rastreado por buscadores académicos (Google Académico), además de realizar consultas sobre diferentes herramientas de IA. Las consultas se han realizado entre los meses de marzo y junio de 2024, utilizando para ello los siguientes términos en español e inglés: Ciencia abierta, Inteligencia Artificial, IA, repositorios públicos y/o institucionales. Por lo tanto, no nos encontramos ante un artículo de investigación que lleva a cabo una revisión bibliográfica sistemática, sino que, tras la búsqueda y análisis de resultados, ha seleccionado cuáles son los documentos más relevantes para la consecución de los objetivos. Es por tanto un trabajo prospectivo que lleva a cabo una labor especulativa en que la aporta una visión, atendiendo a los parámetros marcados por la bibliografía, que identifica en qué manera los repositorios públicos e institucionales pueden utilizar la IA para desarrollar su labor de difusión de sus contenidos científicos.

4. Resultados

La IA tiene un potencial significativo para transformar la manera en que los repositorios institucionales gestionan y difunden trabajos científicos. A continuación, se exploran los múltiples aspectos en los que la IA puede contribuir a la mejora de estos repositorios, desde la automatización de procesos internos hasta la evaluación de métricas científicas.

Automatización de Procesos Internos

La automatización es uno de los beneficios más evidentes de la IA en los repositorios institucionales. Algoritmos de aprendizaje automático pueden clasificar automáticamente documentos, asignar metadatos y realizar otras tareas administrativas, lo que reduce el trabajo manual y minimiza errores humanos (Goodfellow, Bengio, & Courville, 2016) (Penabad-Camacho et al., 2024).

Elementos que podrían automatizarse gracias a los desarrollos algorítmicos en los repositorios instituciones serían: a) indexación y clasificación, por medio de algoritmos NLP, agilizando la organización de la información y optimizando la búsqueda y la recuperación; b) extracción automática de metadatos, tales como título, autor, palabras claves, fecha de publicación, llegando incluso a realizar por completo la catalogación los documentos; c) recomendaciones de documentos atendiendo al historial de consultas realizada previamente por los usuarios-investigadores; d) Difusión Selectiva de la Información (DSI), de tal manera que cada vez que un nuevo documento se integre en el repositorio y atendiendo a las consultas previas, el sistema reenviará el documento sin que haya sido solicitado y realizado ninguna consulta para acceder a los papers que conforman el repositorio; e) gestión de duplicados; f) control de calidad formal (ortotipográfico y formal) de los papers que integren el repositorio; g) control de seguridad y privacidad, monitorizando el acceso y uso de los datos del repositorio, garantizando el cumplimiento de las normas de seguridad; h) detección de plagio; i) preservación digital, garantizando la integridad del documento e identificando inconsistencias, garantizando de esta manera el acceso; j) optimización almacenamiento gestionando el espacio gracias al almacenado, reduciendo así costes informáticos derivados del peso de la información; k) control de autoridades, ya sea asociándolo a plataformas externas o autogestionando dicho control de autoridad; l) gestión de derechos de autor y licencias, asegurando el cumplimiento normativo y legal; m) Extracción automática de resúmenes, al igual que con anterioridad destacamos la opción de llevar a cabo una extracción automática de palabras claves; n) identificación y actualización de perfiles de investigadores y grupos de investigación, etc.

La automatización de muchos de estos procesos internos es ya una realidad, mientras que otros aún se están desarrollando o acelerando su implementación de forma trepidante gracias al desarrollo de algoritmos y aplicaciones concretas de IA.

Difusión de Información Científica

La IA puede mejorar significativamente la difusión de información científica al utilizar algoritmos para identificar y promover trabajos relevantes. Los sistemas de recomendación, similares a los utilizados en plataformas como Amazon o Netflix, pueden sugerir artículos científicos a los usuarios en función de sus intereses y comportamiento de búsqueda previo. Esto no solo aumenta

la visibilidad de los trabajos almacenados, sino que también facilita el descubrimiento de investigaciones relevantes por parte de los usuarios.

Los filtrados colaborativos identifican patrones de comportamiento de los usuarios y por usuarios que son considerados similares al llevar a cabo una comparativa de requerimientos y de utilización. Los filtrados basados en contenidos se basan en el análisis de las características propias de los documentos, tales como autores, títulos, palabras claves, entre otras cuestiones, siendo comparado con el comportamiento del usuario, pudiendo facilitar y/o recomendar documentos de manera más precisa.

Los sistemas de recomendaciones en repositorios mejorarán y aumentará la visibilidad de los trabajos, ayuda al descubrimiento de investigaciones relevantes no requeridas ni consultadas, mejora la eficiencia y la eficacia de las búsquedas de información y, en cierta manera, fomenta la colaboración científica al conectar a investigadores por medio de la lectura de documentos, lo que puede convertirse con posterioridad en contactos intelectuales y potenciales coautorías.

Transparencia y Accesibilidad

La transparencia y accesibilidad son pilares fundamentales de la ciencia abierta. La IA puede ayudar a asegurar que los datos y resultados científicos sean accesibles de manera transparente. Por ejemplo, herramientas de IA pueden ser utilizadas para analizar y auditar la integridad de los datos, asegurando que no haya manipulación o errores. Además, la IA puede automatizar la traducción de documentos científicos, eliminando barreras lingüísticas y permitiendo que una audiencia global acceda a la investigación (Johnson, Watkinson, & Mabe, 2018).

En el ámbito de la transparencia se puede automatizar la gestión de datos señaladas inicialmente encaminados a integrar metadatos precisos y consistentes, en nombre de autores, afiliación, clasificación de documentos, entre otras cuestiones. La verificación de datos y la auditoría constante son también elementos que mejoran la transparencia y pueden ser aportados por la IA (Johnson, Watkinson, & Mabe, 2018), así como gestionar un histórico de cambios en el repositorio garantizando la trazabilidad completa de las modificaciones (Jordan & Mitchell, 2015).

Desde el punto de vista de la accesibilidad, la utilización de la IA abre un amplio abanico de opciones, que van desde la mejora de la búsqueda y

recuperación de información por medio de motores de búsqueda más preciosos y relevantes (LeCun, Bengio, & Hinton, 2015), hasta las recomendaciones personalizadas o la accesibilidad para personas con problemas de accesibilidad.

Evaluación Social o Colaborativa de la Ciencia

La evaluación de la ciencia tradicionalmente se ha basado en métricas como el factor de impacto. Sin embargo, la IA puede facilitar métodos de evaluación más colaborativos y sociales. Plataformas que utilicen IA pueden analizar las interacciones y colaboraciones entre investigadores, proporcionando una visión más holística del impacto de un trabajo científico (Jordan & Mitchell, 2015). Además, las herramientas de IA pueden recopilar y analizar comentarios y revisiones de pares de manera más eficiente, contribuyendo a un sistema de evaluación más transparente y democrático.

Para llevar a cabo la superación del factor de impacto como único elemento de análisis de la calidad de los trabajos de investigación, la evaluación social se sostiene sobre el análisis de las interacción y colaboraciones entre investigadores utilizando técnicas avanzadas de minería de datos, llegando a mapear redes de colaboración e identificando patrones y tendencias, así como la influencia real de los trabajos de investigación. Esto puede desarrollarse gracias, a) al mapeado de redes colaborativas, b) análisis de la coautoría, c) detección de influencias transdisciplinarias, etc.

Además, estos nuevos algoritmos, podrán recopilar información partiendo de comentarios y de las revisiones que, cada vez en mayor número de publicaciones, son públicas haciendo de este ejercicio una labor más transparente y especialmente democrática. Para ello, es evidente que la identificación de los comentarios de revisores expertos es importante, pero igualmente pueden analizarse gracias a las opiniones e incluso al análisis “de sentimientos”, valorando el tono de los comentarios, las impresiones y la visión (negatividad-positividad) de los comentarios.

La IA puede igualmente potenciar las plataformas de evaluación colaborativa, generando sistemas de calificación y valoración para que los investigadores y miembros de la comunidad científica puedan valorar dichos trabajos, aportando valoraciones o revisiones; creando foros de debate y discusión, para debates abiertos y globales; así como el análisis del impacto social y mediático.

Igualmente, los nuevos desarrollos pueden ayudar a mejorar la transparencia y la equidad de la evaluación científica, haciendo que los criterios sean conocidos y transparentes, reduciendo los sesgos propios de la investigación (geográfico, institucional, idiomático, de género) garantizando que se evalúa la calidad y no factores extrínsecos y, en último lugar, potenciando la participación inclusiva ya que estas plataformas de evaluación colaborativa pueden llegar a un mayor número de investigadores y también a grupos más diversos e interdisciplinarios, y por extensión, representativos.

Recomendaciones Personalizadas

Las recomendaciones personalizadas son otra área donde la IA puede aportar un valor significativo. Al analizar los patrones de búsqueda y lectura de los usuarios, los sistemas de IA pueden sugerir artículos y recursos que sean de interés particular para cada usuario. Esto no solo mejora la experiencia del usuario, sino que también aumenta la probabilidad de que se descubran y utilicen trabajos científicos almacenados en el repositorio.

Los algoritmos de filtrado colaborativo analizan grandes volúmenes de datos de usuarios para identificar patrones y comportamientos comunes, pudiendo llegar a predecir preferencias basadas en usuarios similares. De esta manera se pueden facilitar recomendaciones precisas (Goodfellow, Bengio, & Courville, 2016).

El filtrado basado en contenido es una técnica centrada en analizar las características que el propio usuario marca atendiendo a interacciones previas (autor, tema, tipología documental, etc.) de tal manera que recomienda documentos similares, lo que es muy útil en campos especializados. (Jordan & Mitchell, 2015). Los sistemas híbridos combinan los anteriores, de tal manera que proporcionan sugerencias más precisas y a la vez diversificadas. Asociado a las anteriores opciones, la IA puede analizar datos contextuales y redes sociales para recabar información de perfiles, con similitudes en las solicitudes previas.

Los modelos predictivos se anticipan a las necesidades basándose en sus patrones de comportamiento, analizando el histórico de consultas y comportamientos previos, lo que además afina la búsqueda al ofrecer información relevante ponderada al usuario (Russell & Norvig, 2020).

Mejora en la Traducción

La barrera del idioma es un obstáculo significativo en la difusión de la ciencia. La IA puede superar este desafío mediante el uso de tecnologías avanzadas de traducción automática. Algoritmos de traducción basados en IA, como Google Translate, han mejorado significativamente en precisión y pueden traducir artículos científicos completos con un alto grado de fidelidad (Johnson, Watkinson, & Mabe, 2018). Esto permite que investigaciones de diversas partes del mundo sean accesibles para una audiencia global.

Existen diferentes métodos para llevar a cabo dicha traducción, como la neuronal que mejora significativamente la calidad, las de aprendizaje automático y la contextual adaptativa, con la capacidad de adaptarte a la terminología concreta de diferentes disciplinas,

La integración de herramientas de traducción automática en los repositorios institucionales permitirá que los usuarios accedan a contenido traducido sin necesidad de salir de la plataforma. Los repositorios pueden ofrecer traducciones instantáneas de artículos, resúmenes y otros contenidos, facilitando el acceso inmediato a la información. Además, estas herramientas pueden ser personalizables, permitiendo a los usuarios seleccionar el idioma de su preferencia (Bourne, Lorsch, & Green, 2015).

La integración de tecnologías de IA en los repositorios institucionales puede transformar la forma en que se realiza la traducción de contenido científico, mejorando significativamente la accesibilidad y la difusión del conocimiento. Las herramientas avanzadas de traducción automática, el aprendizaje continuo, la personalización y la adaptación a contextos específicos permiten que los repositorios ofrezcan traducciones precisas y de alta calidad. Esto no solo facilita el acceso global al conocimiento científico, sino que también fomenta la colaboración internacional y el intercambio de ideas.

Integración e Interoperabilidad

La interoperabilidad entre diferentes repositorios institucionales es crucial para la ciencia abierta. La IA puede facilitar esta interoperabilidad mediante el desarrollo de estándares y protocolos comunes que permitan el intercambio de datos y documentos entre plataformas (Tennant et al., 2016). Por ejemplo, la IA puede ayudar a mapear y transformar metadatos de un formato a otro, asegurando que la información sea accesible y utilizable en diferentes sistemas.

Utilizan estándares ampliamente aceptados como Dublin Core, MARC21 o METS que facilitan la interoperatividad. La IA puede llegar a mapear y transformar metadatos entre diferentes estándares, facilitando que los datos sean compatibles y, por lo tanto, intercambiables. Igualmente, la IA puede facilitar que los protocolos sean interoperables, pudiendo compartir y comunicarse de manera eficiente. Igualmente pueden realizar mejoras en las APIs para facilitar la integración fluida entre repositorios, facilitando el acceso a datos desde diferentes fuentes.

La integración e interoperabilidad en los repositorios institucionales son fundamentales para el desarrollo de la ciencia abierta, permitiendo un acceso más amplio y equitativo al conocimiento científico. La IA ofrece herramientas poderosas para potenciar, mejorar y analizar estos aspectos, desde la estandarización de metadatos y el uso de APIs hasta el mapeo semántico y el análisis predictivo. A través de estas tecnologías, los repositorios pueden convertirse en plataformas más eficientes, accesibles y colaborativas, facilitando la difusión y el uso del conocimiento científico a nivel global.

Monitoreo y Mantenimiento Predictivo

El monitoreo y mantenimiento predictivo en los repositorios institucionales son esenciales para garantizar la continuidad, seguridad y eficiencia en el manejo de la información científica en el contexto de la ciencia abierta. La IA puede desempeñar un papel significativo en la potenciación, mejora y análisis de estos aspectos, proporcionando herramientas avanzadas que permiten prever y mitigar posibles problemas antes de que ocurran, optimizar recursos y mejorar la eficiencia operativa. A continuación, se detallan cómo la IA puede contribuir a estos objetivos. Potenciar tanto el monitoreo como el mantenimiento predictivo con IA facilita la detección automática de anomalías de grandes volúmenes de datos en tiempo real infiriendo problemas en los hardware, software como en problemas de rendimiento (Goodfellow, Bengio, & Courville, 2016).

El uso de la IA para el monitoreo y mantenimiento predictivo en repositorios institucionales ofrece numerosas ventajas en el contexto de la ciencia abierta. Desde la detección automática de anomalías y el análisis predictivo hasta la optimización de recursos y la automatización de tareas de mantenimiento, la IA puede mejorar significativamente la eficiencia, disponibilidad y seguridad de los repositorios. Además, el análisis continuo y la

generación de informes detallados proporcionan información valiosa para la mejora continua, asegurando que los repositorios institucionales puedan cumplir su misión de difundir el conocimiento científico de manera eficaz y sostenible.

Colaboración Interdisciplinaria

La IA puede facilitar la colaboración interdisciplinaria al identificar conexiones y sinergias entre diferentes áreas de investigación. Algoritmos de análisis de redes pueden mapear las relaciones entre investigadores y proyectos, sugiriendo posibles colaboraciones que no serían evidentes de otra manera (Bourne, Lorsch, & Green, 2015). Esta capacidad de la IA puede fomentar una investigación más integrada y holística.

Para ello, la IA puede ayudar a generar redes de colaboradores por medio de requerimientos e intereses comunes identificados por medio del acceso a los repositorios. Estas relaciones y potenciales colabores pueden ser tanto de áreas concretas como de investigadores de diferentes disciplinas, pero unidos por trabajos transversales o en la frontera entre ellas, atendiendo siempre a objetivos comunes. Para ello, se analiza especialmente los temas y las tendencias, descubriendo así áreas e intereses comunes (Ananiadou et al., 2010).

La mejora de la colaboración interdisciplinaria puede venir de la mano de las recomendaciones personalizadas de colaboradores, identificación de proyectos con similitudes, mejorando la interoperabilidad de datos que hacen que los mismos puedan ser útiles para disciplinas a priori alejadas.

Igualmente, la IA nos ayudaría a evaluar dicha colaboración, analizando el impacto de dichas colaboraciones desde un punto de vista bibliométrico y/o social.

La IA ofrece un conjunto robusto de herramientas y estrategias para potenciar, mejorar y analizar la colaboración interdisciplinaria en repositorios institucionales. Desde la creación de redes de investigación y análisis de temas emergentes hasta la facilitación de proyectos conjuntos y la evaluación del impacto, la IA puede transformar la manera en que los investigadores colaboran. La interoperabilidad de datos y la recomendación personalizada de colaboradores son especialmente cruciales para superar las barreras disciplinarias y fomentar un entorno de investigación más inclusivo y efectivo. El monitoreo continuo y la identificación de barreras y oportunidades aseguran que las colaboraciones se mantengan dinámicas y productivas, alineándose con

los principios de la ciencia abierta. La integración de estas tecnologías y estrategias no solo mejora la colaboración interdisciplinaria, sino que también contribuye al avance del conocimiento científico y su aplicación a problemas complejos del mundo real.

Gestión de Datos de Investigación

La gestión de datos es una tarea compleja que puede beneficiarse enormemente de la IA. Herramientas de IA pueden automatizar la organización, limpieza y análisis de grandes volúmenes de datos, asegurando que sean precisos y consistentes (Fecher & Friesike, 2014). Además, la IA puede ayudar a generar metadatos ricos y detallados que faciliten la recuperación y reutilización de los datos.

La gestión de datos de investigación en los repositorios institucionales puede ser significativamente mejorada mediante el uso de IA. Desde la organización y estructuración de datos hasta la implementación de políticas y la formación personalizada, la IA ofrece un conjunto diverso de herramientas que pueden facilitar el acceso y la reutilización de los datos. Además, el análisis de la calidad y el uso de los datos permite a los investigadores adaptar sus prácticas y maximizar el impacto de su trabajo. Al integrar estas tecnologías en los repositorios institucionales, se avanza en los principios de la ciencia abierta, promoviendo una mayor transparencia, colaboración y acceso a la información científica.

Identificación de Plagio y Fraude

La identificación de plagio y fraude en los repositorios institucionales es crucial para mantener la integridad científica y la confianza en la investigación. En el contexto de la ciencia abierta, donde la accesibilidad y la transparencia son primordiales, la implementación de tecnologías avanzadas, como la IA, puede desempeñar un papel significativo.

Para potenciar la identificación de estos trabajos la IA desarrolla herramientas automatizadas para ello, son algoritmos de comparación de textos junto con técnicas de procesamiento del lenguaje natural (NLP), identificando similitudes y plagios. La integración de repositorios institucionales con bases de datos externas y sistemas de verificación de plagio, como CrossRef y iThenticate, puede potenciar la detección de plagio. La IA puede facilitar esta integración, permitiendo una verificación cruzada más exhaustiva y efectiva de los documentos almacenados en los repositorios.

5. Análisis y discusión

La relación entre la IA y los repositorios institucionales en el contexto de la ciencia abierta presenta un amplio espectro de oportunidades y desafíos. A medida que las tecnologías avanzan, también lo hacen las expectativas y requerimientos en términos de eficiencia, accesibilidad y transparencia de los repositorios.

Los retos fundamentales del uso de la IA en los repositorios se centran en los siguientes elementos:

Calidad y heterogeneidad de los datos, de tal manera que exista una consistencia y estructuración para un funcionamiento óptimo en el análisis y en la recuperación de la información, lo que choca con una realidad en la que los repositorios cuentan con una variedad de formatos con integración y metadatos diversos en una misma plataforma (Davis & Connolly, 2007).

Aún se requieren infraestructuras y recursos técnicos más robustos y desarrollados para poder lograr algunas de las cuestiones planteadas, aunque, sin lugar a dudas la evolución de la IA en los últimos cinco años ha sido trepidante, sigue aún en un estado embrionario por lo que, en estos momentos, sus funcionalidades aún se están desarrollando en muchos campos, pero por el contrario nos ofrece unas potencialidades que seguramente aún no están siendo planteadas. Los repositorios institucionales necesitan de una apuesta decidida tanto en políticas de integración como en costes para su mejora y mantenimiento, costes que deben incluir sus desarrollos vía IA.

La privacidad y la seguridad de la información y datos son una de las grandes preocupaciones y críticas en el desarrollo tanto de la ciencia abierta como en el contexto actual de saturación información y accesibilidad universal como en la IA. Evitar usos inadecuados y brechas de seguridad son grandes retos.

La ética y el sesgo en los algoritmos de IA son temas controvertidos y de creciente preocupación. Los modelos de IA pueden perpetuar sesgos existentes en los datos de entrenamiento, lo que puede traducirse en decisiones injustas o discriminatorias. En el contexto de repositorios institucionales, esto puede afectar la visibilidad y accesibilidad de ciertos trabajos científicos, dependiendo de cómo se prioricen y recomienden los contenidos.

Existen además cuestiones bastante controvertidas en la relación de la generación de la IA con otras tecnologías, en este caso, su adaptación a los repositorios institucionales. La automatización de los procesos relega el control humano del mismo, y mientras la IA no este absolutamente perfilada garantizando así su correcto uso, el control o al menos la validación humana sigue siendo pieza clave en muchos de los procesos (Tennant et al., 2016). Igualmente relevante es la necesidad de contar con algoritmos transparentes, siendo actualmente en muchos casos herméticos, y en caso de poder acceder a ellos, son difíciles de comprender. Cualquier algoritmo, sobre todo relacionado con repositorios públicos en el marco de la filosofía de la ciencia abierta, deben poder ser explicadas y auditadas. Otro elemento recurrente en el uso de estas tecnologías para analizar es el impacto de la Ciencia, ya que el sesgo cuantitativo siempre será mucho más fácil de automatizar, aunque la realidad es que la IA nos abre un abanico de posibilidades para analizar la Ciencia y su impacto desde prismas muy interesantes.

Es fundamental que se generen desde las administraciones públicas, políticas y normativas que garanticen y controlen el desarrollo de la IA en todos los ámbitos, incluido los repositorios. La creación de directrices claras y transparentes, así como directiva para la protección de la privacidad y la capacidad para garantizar la propiedad actual, se nos demuestra como una labor prioritaria.

6. Conclusiones

Las investigaciones revisadas muestran claramente que la IA ofrece un vasto potencial para mejorar la gestión y difusión de trabajos científicos en repositorios institucionales. Desde la automatización de procesos internos hasta la mejora de la accesibilidad y la evaluación del impacto social, la IA puede transformar significativamente la forma en que se manejan y comparten los conocimientos científicos.

En primer lugar, la automatización de procesos administrativos mediante IA puede optimizar la clasificación y asignación de metadatos, reduciendo la carga de trabajo manual y mejorando la eficiencia operativa de los repositorios. Esto no solo libera tiempo para actividades más estratégicas, sino que también minimiza errores humanos en la gestión de documentos científicos.

En segundo lugar, la IA facilita la difusión efectiva de información científica a través de sistemas avanzados de recomendación y motores de búsqueda mejorados. Estos sistemas pueden personalizar recomendaciones de contenido científico según los intereses individuales de los usuarios, promoviendo una mayor visibilidad y acceso a investigaciones relevantes.

Además, la transparencia y accesibilidad de la ciencia se fortalecen mediante herramientas de IA que aseguran la integridad de los datos y facilitan la traducción automática de documentos, superando barreras lingüísticas y culturales. Esto promueve un entorno más inclusivo y globalizado para la investigación científica.

En términos de evaluación de la ciencia, la IA proporciona métodos innovadores y colaborativos para medir el impacto y la relevancia de los trabajos científicos. Algoritmos de análisis de redes sociales y métricas alternativas pueden complementar los enfoques tradicionales de evaluación, ofreciendo una perspectiva más holística y actualizada del impacto de la investigación.

En el ámbito de la recuperación de información, la IA mejora la precisión y relevancia de los resultados de búsqueda mediante técnicas avanzadas de procesamiento de lenguaje natural y modelos de aprendizaje automático. Esto facilita a los investigadores encontrar y utilizar información pertinente de manera más eficiente, impulsando la productividad y el descubrimiento científico.

Para el futuro, las líneas de investigación podrían explorar aún más el potencial de la IA en áreas como la preservación digital, la colaboración interdisciplinaria, la detección de plagio y fraudes y la asistencia en la escritura científica. Además, es esencial abordar desafíos éticos y de gobernanza relacionados con el uso de la IA en la ciencia, asegurando la equidad, transparencia y responsabilidad en su implementación.

Por lo tanto, la integración estratégica de la IA en los repositorios institucionales promete transformar profundamente la forma en que se gestionan y difunden los trabajos científicos. Al aprovechar el potencial de la IA, los investigadores y administradores pueden avanzar hacia una ciencia más abierta, colaborativa y accesible, beneficiando tanto a la comunidad científica como a la sociedad en general.

Bibliografía

- ABADAL, E., 2021. Ciencia abierta: un modelo con piezas por encajar. *Arbor*, **197**(799), a588. <https://doi.org/10.3989/arbor.2021.799003>
- ABADAL, E., y ANGLADA, L., 2020. Ciencia abierta: cómo han evolucionado la denominación y el concepto. *Anales de Documentación*, **23**(1). <https://doi.org/10.6018/analesdoc.378171>
- ALONSO-AREVALO, J y QUINDE-CORDER, M., 2024. El papel de las bibliotecas en la era de la inteligencia artificial (IA). *Boletín de la Asociación Andaluza de Bibliotecarios*, **127**, 27-37.
- ÁLVAREZ WONG, B.I., 2017. Los repositorios digitales para la conservación. Un acercamiento a la preservación digital a largo plazo. *Ciencias de la Información*, **48** (2), 15-22. <https://www.redalyc.org/pdf/1814/181454540003.pdf>
- ANANIADOU, S.; PYYSALO, S.; TSUJII, J., y KELL, D. B., 2010. Event extraction for systems biology by text mining the literature. *Trends in Biotechnology*, **28**(7), 381-390. <https://doi.org/10.1016/j.tibtech.2010.04.005>
- BABINI, D.; GONZÁLEZ, J.; LÓPEZ, F. y MEDICI, F., 2010. Construcción social de repositorios institucionales: el caso de un repositorio de América Latina y El Caribe. *Información, Cultura y Sociedad*, **23**, 63-90. <http://www.scielo.org.ar/pdf/ics/n23/n23a04.pdf>
- BETHENCOURT-AGUILAR, A.; CASTELLANOS-NIEVES, D.; SOSA-ALONSO, J. J. y AREA-MOREIRA, M., 2022. Implicaciones técnicas y prácticas de las Redes Adversarias Generativas a la Ciencia Abierta en Educación. *RiITE Revista interuniversitaria de investigación en Tecnología Educativa*, 138–156. <https://doi.org/10.6018/riite.545881>
- BOAI, 2002. *Budapest Open Access Initiative*. Retrieved from Budapest Open Access Initiative: <https://www.budapestopenaccessinitiative.org/read>
- BOURNE, P. E.; LORSCH, J. R., y GREEN, E. D., 2015. Perspective: Sustaining the big-data ecosystem. *Nature*, **527**(7576), S16-S17. <https://doi.org/10.1038/527S16a>
- CALDERA-SERRANO, J., 2018. Repositorios públicos frente a la mercantilización de la ciencia: apostando por la ciencia abierta y la evaluación cualitativa. *Métodos de Información*, **9**, (17), 74-101. DOI: <https://dx.doi.org/10.5557/IIMEI9-N17-074101>
- DAVIS, P. M. y CONNOLLY, M.J.L., 2007. Institutional Repositories: Evaluating the Reasons for Non-use of Cornell University's Installation of DSpace. *D-Lib Magazine*, **13** (3/4). <https://www.dlib.org/dlib/march07/davis/03davis.html>
- PENABAD-CAMACHO, L., PENABAD-CAMACHO, M. A., MORA-CAMPOS, A., CERDAS-VEGA, G., MORALES-LÓPEZ, Y., ULATE-SEGURA, M., MÉNDEZ-SOLANO, A., NOVA-BUSTOS, N., VEGA-SOLANO, M. F., &

- CASTRO-SOLANO, M. M., 2024. Heredia Declaration: Principles on the use of Artificial Intelligence in Scientific Publishing. *Revista Electrónica Educare*, 28(S), 1-10. <https://doi.org/10.15359/rec.28-S.19967>
- FECHER, B., y FRIESIKE, S., 2014. Open Science: One Term, Five Schools of Thought. EN: *Opening Science*. Cham: Springer, 17-47 https://doi.org/10.1007/978-3-319-00026-8_2
- FECYT. 2021. *Guía para la evaluación de repositorios institucionales de investigación*. Fundación Española para la Ciencia y la Tecnología, FECYT. https://calidadrevistas.fecyt.es/sites/default/files/informes/2021guiaevaluacionrecolecta_vf.pdf
- GOOGFELLOW, I.; BENGIO, Y. y COURVILLE, A., 2016. *Deep Learning*. Cambridge: MIT Press.
- JOHNSON, R.; WATKINSON, A. y MABE, M., 2018. The STM Report: An overview of scientific and scholarly publishing. *STM: International Association of Scientific, Technical and Medical Publishers*.
- JORDAN, M. I. y MITCHELL, T. M., 2015. Machine learning: Trends, perspectives, and prospects. *Science*, **349**(6245), 255-260. <https://doi.org/10.1126/science.aaa8415>
- LECUN, Y.; BENGIO, Y. y HINTON, G., 2015. Deep learning. *Nature*, **521**(7553), 436-444. <https://doi.org/10.1038/nature14539>
- MARTÍNEZ MÉNDEZ, F. J.; BAPTISTA, A. A.; LÓPEZ CARREÑO, R. y DELGADO VÁZQUEZ, A. M. 2023. Implementación de los repositorios de datos de investigación en las universidades públicas españolas: estado de la cuestión. *Scire: Representación y Organización del Conocimiento*, **29**(2), 39-49. <https://doi.org/10.54886/scire.v29i2.4914>
- RUSSELL, S. J. y NORVING, P., 2020. *Artificial Intelligence: A Modern Approach*. Pearson.
- TENNANT, J. P.; WALDER, F.; JACQUES, D. C.; MASUZZO, P.; COLLISTER, L. B. y Hartgerink, C. H., 2016. The academic, economic and societal impacts of Open Access: An evidence-based review. *F1000Research*, **5**. <https://doi.org/10.12688/f1000research.8460.3>
- VICENTE-SAEZ, R. y MARTÍNEZ-FUENTES, C., 2018. Open Science now: A systematic literature review for an integrated definition. *Journal of Business Research*, **88**, 428-436. <https://doi.org/10.1016/j.jbusres.2017.12.043>

Reconocimiento

Esta investigación ha sido financiada por la Junta de Extremadura y Fondos FEDER “Una manera de hacer Europa” [Grupo de Investigación QUINARI SEJ013 GR18130]

Uso de IA

Se ha utilizado ChatGPT (2025), versión -4- Turbo, para la revisión de estilo del texto y traducción de resumen.