

Hacia revisiones de la literatura más eficientes potenciadas por inteligencia artificial

Erik Carbajal-Degante^{a,†}, Myrna Hernández Gutiérrez^{a,§},
Melchor Sánchez-Mendiola^{a,b,◊,*}

Facultad de Medicina



Resumen

Las revisiones de la literatura para tareas de investigación y docencia se enfrentan cada vez más a nuevos retos, debido principalmente al crecimiento considerable del material bibliográfico en todas las disciplinas. Esencialmente, la búsqueda y selección de información se tornan exhaustivas y agotadoras debido a la falta de organización en la recopilación sistemática de documentos, así como el manejo incorrecto de los metadatos y la gran variabilidad en los resultados que ofrecen los mecanismos automáticos de consulta. En este sentido, la inteligencia artificial (IA) está revolucionando la revisión de la literatura al emplear sistemas capaces de analizar e interpretar el contenido de muchos documentos en una forma rápida, así como ofrecer resultados con un nivel de precisión mayor que los motores de búsqueda

tradicionales. Este trabajo tiene como objetivo describir los avances recientes en la etapa de acoplamiento inteligente de los métodos para la revisión de la literatura académica, refiriendo al uso de herramientas de aprendizaje automático y aprendizaje profundo asociadas a la optimización en el procesamiento del lenguaje natural. Finalmente se abordan las implicaciones que conlleva emplear sistemas capaces de omitir juicios y tomar decisiones con respecto a la utilidad y relevancia de los datos analizados.

Palabras clave: Revisión de la literatura; inteligencia artificial; aprendizaje automático; aprendizaje profundo.

Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

^a Coordinación de Universidad Abierta, Innovación Educativa y Educación a Distancia (CUAIEED), Universidad Nacional Autónoma de México (UNAM), Cd. Mx, México.

^b División de Estudios de Posgrado, Facultad de Medicina, UNAM, Cd. Mx, México.

ORCID ID:

[†]<https://orcid.org/0000-0002-1532-9322>

[§]<https://orcid.org/0000-0002-0485-9037>

[◊]<https://orcid.org/0000-0002-9664-3208>

Recibido: 2 de abril de 2023. Aceptado 9 de junio de 2023.

* Autor de correspondencia: Dr. Melchor Sánchez-Mendiola. CUAIEED, UNAM. Anexo Centro Ciencias de la Complejidad, 1er Piso, Circuito Cultural S/N, Ciudad Universitaria, Coyoacán, Cd. Mx, México.

Correo electrónico: melchorsm@unam.mx

Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Towards enhanced literature review efficiency powered by artificial intelligence

Abstract

Literature reviews for research and teaching tasks are increasingly facing new challenges, mainly due to the considerable growth of bibliographic material in all disciplines. Primarily, the search and selection of information become exhausting and overwhelming due to the lack of planning in the systematic collection of documents, as well as the improper handling of metadata and the high variability of the results offered by automatic search mechanisms. In this sense, artificial intelligence is transforming the landscape of literature review by employing systems capable of quickly analyzing and interpreting the content of text documents while providing very accurate results,

higher than traditional search engines. This paper aims to describe recent advances in the intelligent integration process of literature review methods, referring to the use of machine and deep learning techniques associated with optimization stages of natural language processing. Finally, we address the implications of employing systems capable of making judgments and making decisions regarding the usefulness and relevance of the analyzed data.

Keywords: Literature reviews; artificial intelligence; machine learning; deep learning.

This is an Open Access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

INTRODUCCIÓN

Es difícil estimar el tamaño total de internet debido a su constante crecimiento y la naturaleza distribuida de su infraestructura. Sin embargo, estudios señalan una cifra cercana a los 50 zettabytes (50 mil millones de terabytes) hasta 2022¹. Este tamaño abrumador se debe en gran medida a la cantidad de contenido generado por los usuarios, plataformas en línea y servicios digitales en todo el mundo. Así mismo, se calcula que una gran proporción del contenido en internet está compuesta por texto, ya sea en forma de código con lo que se generan las páginas web y ciertas aplicaciones, como correo electrónico o como documentos de texto. En el caso de material bibliográfico, se estima una tasa de crecimiento de 1 millón de publicaciones al año, lo que representa que cada 30 segundos un nuevo trabajo es añadido a las filas del internet².

El gran interés en la inteligencia artificial (IA) para transformar las economías se refleja en la escala del gasto global donde la Corporación Internacional de Datos predice casi \$98 mil millones de dólares en 2023, más del doble de lo que se invirtió en 2019³. A pesar de que la IA ofrece numerosos beneficios, su adopción en el campo educativo es más lenta en comparación con otros campos como la medicina, las finanzas y la industria tecnológica⁴. Sin embargo,

en los últimos años, en el periodo comprendido entre 2018 a 2022, el interés de adoptar esta disciplina para la educación ha crecido un 48%⁵, lo cual se refleja en el informe Horizon-2018 con un estimado de tiempo para la adopción de importantes desarrollos tecnológicos educativos de alrededor de tres años⁶.

Entre las principales aplicaciones de la IA en el ámbito educativo se encuentran el modelado predictivo en la analítica del aprendizaje, la tecnología de asistencia (como los sistemas de tutoría o asistentes de aprendizaje) y el análisis automático de contenido⁷; esta última herramienta ayuda a extraer y procesar información de grandes volúmenes de datos de manera eficiente, precisa y automatizada, facilitando la presentación de resultados contextualizados y personalizados con lo que se mejora también la productividad de quienes lo usan.

En este sentido, el panorama complicado que enfrenta la comunidad investigadora y docente en las tareas tradicionales de revisión de la literatura al navegar en un océano de datos que poseen un nivel de aportación poco relevante a sus trabajos, puede verse impulsado por el uso de las nuevas tecnologías basadas en IA, ya que en los procesos de revisiones sistemáticas su uso ha mostrado gran utilidad, cuando se emplean correctamente a través de herramientas de procesamiento inteligente de la información,

lo que convierte a la IA en un aliado potencial en la optimización de los recursos invertidos en esta tarea⁸.

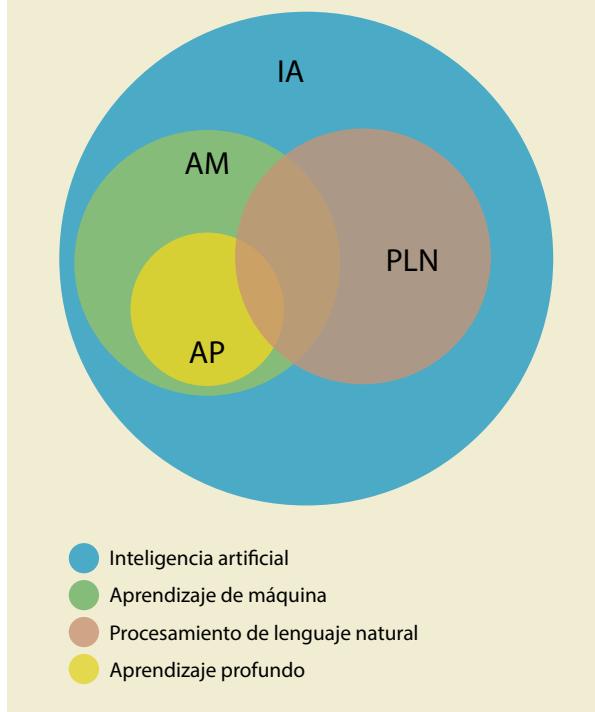
LOS MAYORES DESAFÍOS

Efectuar un proceso de revisión de la literatura académica implica abordar varios elementos reproducibles, como la identificación de todas las publicaciones relevantes de una manera estandarizada, la extracción de datos y la síntesis de los resultados. Dicho proceso tradicionalmente se realiza de forma manual, lo que conlleva recolectar una cantidad excesiva de documentos y supone un desafío para las capacidades cognitivas humanas al procesar grandes volúmenes de información en un tiempo corto asociado a la duración de una investigación. El trabajo de Wallace y cols.⁹, señala que incluso para un académico experimentado, la tarea de lectura y etiquetado de un *abstract* puede tomar hasta 7 minutos y dado que la información relevante se encuentra en una proporción menor al 5% de la información extraída, es sin duda un problema complejo al que es difícil encontrar un balance.

Es de suma importancia notar que la búsqueda de material bibliográfico incluso sobre un tema en específico puede arrojar resultados que exceden los 10,000 artículos¹⁰, lo que representa también un reto en la integración de criterios de selección y filtrado para los usuarios e incluso para los mismos sistemas de adquisición de información, permitiendo un riesgo notable de excluir un estudio potencialmente útil cuando se realizan las llamadas búsquedas rápidas o estrechas.

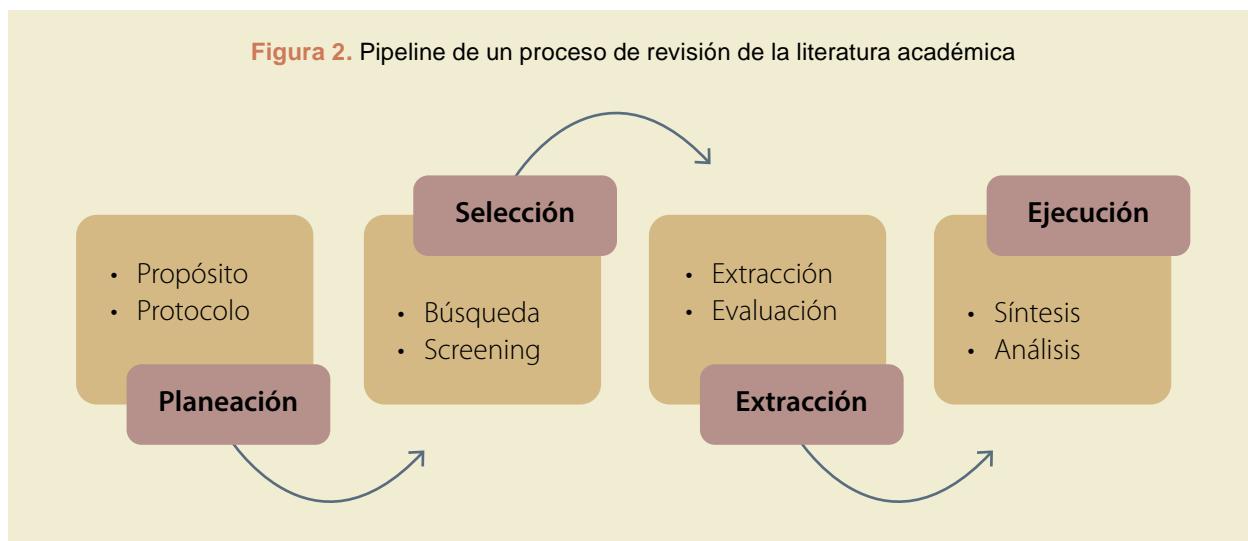
Además, es común toparse con información bibliográfica ocluida, haciendo que los sistemas automáticos operen sobre datos débilmente estructurados, incompletos, difusos y confusos. La presencia de metadatos dañados inhibe directamente la precisión en las consultas, afectando directamente la calidad de los resultados. Superar este desafío implica la utilización de métodos de imputación de datos, técnicas de minería de texto y análisis de patrones, para inferir y completar la información faltante, lo que obliga a los usuarios a hacer uso de herramientas tecnológicas para mitigar los efectos de una navegación desorientada sobre las bases de datos.

Figura 1. Esquema de las inmediaciones de la inteligencia artificial



EL IMPULSO DE LA IA

En general, los sistemas basados en IA, así como el enfoque que brinda la disciplina del procesamiento de lenguaje natural (PLN), utilizan mecanismos de aprendizaje con el propósito de entender y replicar las decisiones que los humanos ejercen para priorizar y ordenar la información por relevancia y utilidad. En el contexto del procesamiento de la información, los mecanismos de aprendizaje de máquina (aprendizaje supervisado, no supervisado, auto-supervisado y reforzado) juegan un papel fundamental al permitir el aprendizaje a partir de un conjunto de datos (etapa de entrenamiento) y mejorar su rendimiento a medida que se les proporciona más información, ofreciendo así adaptabilidad y personalización de acuerdo con el impacto de la aplicación y las necesidades de uso. La figura 1, muestra la muy estrecha relación entre la disciplina del PLN y los enfoques de aprendizaje de máquina, cada campo posee su propia base teórica, sin embargo, todos ellas convergen al darle un sentido aplicado.



La disciplina del PLN ofrece ir más allá del tratamiento puramente correctivo del texto en la búsqueda de generar sistemas con la capacidad de abstraer y analizar el significado semántico y sintáctico de la información. El PLN es útil en una amplia gama de tareas, como el resumido automático de documentos, el etiquetado de texto, la interpretación de tópicos, tareas de generación de preguntas y respuestas, el reconocimiento de las partes de una oración (entidades nombradas, verbos, sustantivos, etc.), así como la llamada comprensión lectora de máquina encargada de entender el significado del contexto¹¹. Lograr plasmar estas características en un sistema de cómputo, permite realizar diferentes tareas asociadas al lenguaje por el hecho de exhibir una capacidad cognitiva similar a la del ser humano con mucho menor tiempo de procesamiento¹², esto abre la posibilidad de una exploración más efectiva de la información y desde una perspectiva más nítida.

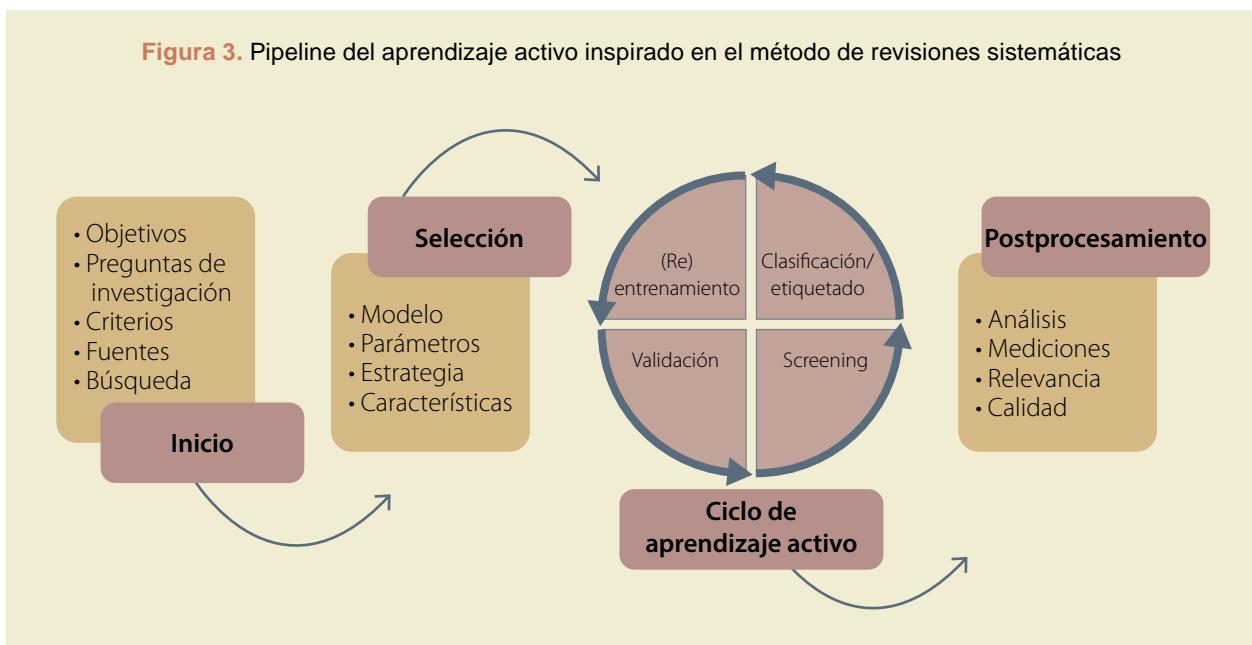
PIPELINE DE UN MECANISMO INTELIGENTE PARA LA REVISIÓN

Las guías presentadas por Okoli y Collins y cols. relatan el proceso de revisiones sistemáticas aplicadas originalmente a la medicina, donde los métodos de revisión difieren conceptualmente de otras áreas. Sin embargo, los retos que enfrenta la investigación actualmente en sentido epistemológico han demandado una búsqueda que pueda generalizar las revisiones sistemáticas con cambios adaptables a cada

disciplina, donde se logre un equilibrio en encontrar eficientemente las metodologías cuantitativas y cualitativas más representativas dentro de los sistemas de información. La calidad de una revisión depende del rigor en la planeación, selección, extracción y el análisis de la información. Se sabe que el objetivo principal de este proceso es hallar tantos trabajos como sea posible, capaces de responder nuestras preguntas de investigación y seguir los protocolos iniciales propuestos en una etapa de planeación.

Posteriormente, la capacidad de filtrar y seleccionar el contenido de acuerdo con criterios establecidos, propicia una reducción sustancial del material a revisar. Esta etapa reconoce la necesidad de lidiar con los retos derivados del problema del contenido exorbitante de datos. La información adquirida mediante una etapa de extracción obedece el propósito de establecer la medida en que una aportación es relevante para la investigación. La evaluación de la calidad suele ser subjetiva, por la dificultad de asignar un valor numérico en el sentido de utilidad. Finalmente, la síntesis y el análisis permiten organizar, comparar y discutir los elementos que infieren en la comprensión y realización de un estudio. Este desarrollo y sus componentes se pueden apreciar en la **figura 2**.

El concepto de aprendizaje activo (del inglés *active learning*) se inspira en un pipeline generalizado y pretende simular los procesos cognitivos del humano, al plantear un modelo autónomo para el

Figura 3. Pipeline del aprendizaje activo inspirado en el método de revisiones sistemáticas

procesado de la información mediante diversas etapas. Como primera medida, se establece una etapa capaz de elegir un conjunto de datos provenientes de una búsqueda automática a partir de criterios establecidos. Segundo, la etapa de selección contribuye a definir el conjunto de elementos técnicos relacionados con el sistema de IA, dentro de los más relevantes están la elección de un modelo y el ajuste inicial de sus parámetros. Posteriormente, la etapa de entrenamiento logra encontrar diferencias significativas reduciendo drásticamente el número de registros para llevar a cabo un primer cribado y etiquetado. Una característica notable de estos sistemas es que sus procesos van más allá de solo analizar los registros relevantes de una búsqueda por título, palabras clave y resumen, sino que también considera necesario aprender una gama complementaria de conceptos como campo y perfil del usuario, especialidad, referencias bibliográficas, etc., lo que provee de robustez, flexibilidad y precisión en el enfoque del clasificado de documentos personalizado, como se muestra en el esquema de la **figura 3**.

La aplicación del aprendizaje activo a la revisión sistemática se le conoce como RITL, del inglés Researcher-In-The-Loop¹⁵, y es utilizada por algunas aplicaciones enfocadas a resolver prioritariamente algunas cuestiones del manejo de información.

Principalmente, esta metodología es utilizada en el proyecto *ASReview* que ofrece un ecosistema donde incluye detección asistida por IA, infraestructura para personalizar los modelos, así como opciones para ejecutar sistemas de simulación para beneficiar principalmente el *screening*¹⁶. Una herramienta similar es *Rayyan*¹⁷, una plataforma que ofrece un sistema que promete aprender de las decisiones que toman los usuarios para la selección del material, además de ofrecer portabilidad y un ambiente colaborativo. Otra más es *Elicit*, un asistente de investigación parcialmente impulsado por IA que le permite automatizar varios procesos de revisión al hacer desde búsquedas de similitud semántica con resumido textual hasta metaanálisis¹⁸. Este potencial se ha obtenido gracias a los recientes avances de los grandes modelos de lenguaje (del inglés Large Language Models) y en específico de GPT (del inglés Generative Pretrained Transformer). La **tabla 1** muestra algunas de las herramientas más utilizadas en la actualidad y su aportación significativa en alguna etapa interna del proceso de revisión.

Los grandes modelos de lenguaje (LLMs) marcan un importante ciclo en la revolución de la IA de los últimos años para efectuar y perfeccionar las tareas en las que se enfoca el PLN¹⁹. Son básicamente algoritmos de aprendizaje profundo aplicados al

Tabla 1. Herramientas utilizadas en alguna etapa interna del proceso de revisión de literatura académica

Recurso	Año ↓	Principales etapas de aportación	Características principales
RobotReviewer ²⁵	2015	• Evaluación	Se enfoca en la evaluación del riesgo de sesgo de los ensayos controlados, permite rastrear de forma interactiva las calificaciones en cada dominio de sesgo hasta su origen en el documento de texto completo.
Rayyan ¹⁷	2016	• Evaluación	Predice la relevancia de una cita en una escala de cinco estrellas. El usuario puede ordenar las citas por su relevancia prevista. Plataforma multimedia y colaborativa.
Libaskets ²⁶	2019	• Formulación • Búsqueda	Permite evaluar el volumen potencial de los resultados de la búsqueda, en función de palabras clave y sobre diferentes bases de datos que se pueden ajustar de forma flexible, además de informar sobre el diseño de estrategias de búsqueda.
TheoryON ²⁷	2020	• Búsqueda	Ejecuta búsquedas basadas en ontologías para constructos individuales y construir relaciones a través de teorías conductuales.
ASReview ⁸	2021	• Screening • Clasificación	Software instalable localmente en un ordenador por lo que la información del usuario se encuentra salvaguardada. Se basa principalmente en el aprendizaje activo con modelos adaptables del aprendizaje de máquina.
Iris.AI ²⁸	2022	• Búsqueda • Screening • Clasificación	Plataforma en línea para realizar búsquedas inteligentes, propone una amplia gama de filtros, análisis de lecturas, resúmenes generados automáticamente, extracción autónoma y sistematización de datos.
Assistant by scite.ai ²⁴	2021	• Búsqueda • Clasificación • Evaluación	Se desarrolló analizando más de 25 millones de artículos científicos con SciBERT y actualmente posee una base de datos de más de 800 millones de citas etiquetadas.
Elicit ¹⁸	2022	• Búsqueda • Screening • Clasificación	El sistema hace uso del Semantic Scholar Academic Graph, sobre el cual, el modelo GPT-3 es implementado para realizar la búsqueda y adquisición del corpus.
ChatGPT ²¹	2023	• Búsqueda • Screening • Clasificación • Evaluación	Debido al gran tamaño de sus datos de entrenamiento y su modelo GPT4, esta herramienta es potenciada mediante la ingeniería de instrucciones. La gran demanda de su uso genera que la comunidad científica sume esfuerzos para la creación de plugins de interconexión con otras aplicaciones.

lenguaje humano, con un conocimiento previo obtenido a través de un proceso de entrenamiento con un conjunto masivo de datos. El uso efectivo de los LLMs implica diseñar estrategias de instrucciones (prompt engineering) con el fin de generar respuestas precisas, relevantes y coherentes, así como evitar un comportamiento incongruente o ficticio del tipo alucinaciones²⁰. Dentro del gran abanico de los LLMs, se encuentra el modelo GPT en su versión más reciente (GPT4) el cual produce la herramienta *ChatGPT* de gran popularidad y aplicación para casi todas las áreas del lenguaje natural²¹. Un modelo similar es el llamado BERT (del inglés Bidirectional Encoder Representation from Transformer), que ha sido desarrollado y lanzado como herramienta de código abierto con la intención de ser personalizable

por la comunidad²². Su versión SciBERT aprovecha un conjunto masivo de información científica de múltiples dominios para mejorar el rendimiento precisamente en tareas del PLN con enfoque científico²³. Tal es el caso de la herramienta *Scite* que, con este modelo en particular, afronta la cuestión de evaluación de referencias basado en el diseño de una métrica para contextualizar la información²⁴. Con ello se busca proporcionar evidencia de apoyo a la etapa de *screening*.

INTELIGENCIA ARTIFICIAL RESPONSABLE Y EXPLICABLE

Las herramientas existentes que facilitan los procesos de revisión actualmente tienen tres inconvenientes principales:

1. Las métricas más comunes de validación en IA como *precision*, *recall*, *F1-score*, *AUC*, etc., muestran el desempeño de los algoritmos sobre un conjunto de datos de manera objetiva y cuantitativa. Sin embargo, se sigue careciendo de los procesos de validación y evaluación adecuados que plasmen los principios del rigor metodológico que utilizan los especialistas en su campo, así como ratificar la correcta interpretación del lenguaje científico que emplean las máquinas.
2. El desarrollo y mantenimiento de las aplicaciones basadas en IA suele ser costoso por el poder de cómputo requerido dentro de alguna etapa de su producción. La infraestructura y los recursos humanos encargados de desempeñar la importante labor del diseño, recolección, adaptación, son limitados, por lo que muchas aplicaciones son ofrecidas ya sea como prueba gratuita restringida, o con un costo adicional para un acceso completo.
3. Finalmente, se sigue lidiando con muchas aplicaciones de código oculto, de donde surge el concepto de los llamados sistemas de caja negra. Esto resulta problemático ya que la replicabilidad, la ética, la transparencia y la propiedad de los datos enfrentan riesgos dentro de una era que fomenta la ciencia abierta y responsable.

Se sabe que la ciencia abierta tiene el potencial de generar avances significativos en diversos campos, esto conlleva responsabilidades referentes al manejo de la información. Se pretende que los sistemas de IA adquieran un sentido de responsabilidad social y humana, eviten el sesgo y la discriminación garantizando la equidad, promuevan la colaboración y apertura, todo esto con el propósito de hacer válidos los principios éticos con los que se apuesta a desarrollar las nuevas tecnologías²⁹. Un punto medular para afrontar la concientización de la IA responsable es la transparencia algorítmica, la cual tiene como objetivo fomentar la confianza y seguridad al proporcionar la explicabilidad necesaria sobre la autonomía de estos sistemas, además de encontrar un equilibrio entre la protección de los datos y el derecho de acceso a la información para todos los usuarios.

Es evidente que la autonomía juega un papel impor-

tante cuando se habla de contextualizar los elementos de información que se transmiten a medida que se ocupa la IA dentro de un proceso de adquisición de datos, como lo son las revisiones de la literatura académica. La capacidad de resaltar la relevancia de ciertas situaciones en la interacción máquina-humano para generar juicios y tomar decisiones, obedece a un nivel mayúsculo en la autoconciencia y la comprensión del entorno. De esta forma, es posible concebir una IA más general (del inglés AGI, Artificial General Intelligence) cuyas características de aprendizaje, razonamiento, comunicación, creatividad y conciencia son comparables, en teoría, a las capacidades humanas.

Se espera que las futuras investigaciones logren abordar las preguntas clave sobre ¿cómo dar transparencia y explicabilidad a un sistema autónomo? Pues el hecho de hablar de una AGI tangible es sinónimo de complejidad, con miras a un proyecto a largo plazo por el hecho de ser un desafío que involucra a un campo interdisciplinario de estudio. Hasta que estas preguntas logren responderse con cierto grado de satisfacción, es probable que los dilemas éticos con implicaciones en lo social y la educación continúen siendo una interrogante persistente, incluso en los procesos de revisión de literatura.

CONCLUSIONES

El proceso de revisión de la literatura académica involucra tanto tareas creativas como mecánicas, lo que abre la posibilidad para que herramientas basadas en IA emergan, con el fin de reducir los esfuerzos que los usuarios emplean en las tareas repetitivas de este complejo proceso. El objetivo de estas herramientas va más allá de optimizar los procesos, considerando también liberar tiempo para dedicarlo a las tareas que requieren de la interpretación humana, intuición, creatividad y la experiencia académica, por lo que se convierten en un aliado potencial con miras al continuo crecimiento de la información y el conocimiento.

Se prevé un escenario donde la IA siga mejorando su eficiencia en las tareas tradicionales de revisión, en comparación con otras metodologías más creativas. Al mismo tiempo, se necesita un discurso bien informado para identificar las áreas de aplicación apropiadas, así como para abordar los desafíos aso-

ciados al uso responsable de estos sistemas, lo que da lugar a la necesidad de generar un marco teórico que analice los efectos de la autonomía, la explicabilidad, la privacidad, así como sus implicaciones. Finalmente, la aceptación y normalización por parte de la comunidad académica es esencial para brindar un impulso sobresaliente en la investigación y docencia al mantenerse a la vanguardia tecnológica y adquirir las habilidades de innovación demandadas hoy en día en el ámbito educativo.

CONTRIBUCIÓN INDIVIDUAL

- ECD: Concepción, conceptualización, escritura, revisión y edición.
- MHG: Concepción, conceptualización, escritura, revisión y edición.
- MSM: Concepción, conceptualización, escritura, revisión, edición y supervisión

AGRADECIMIENTOS

Ninguno.

PRESENTACIONES PREVIAS

Ninguno.

FINANCIAMIENTO

CONACYT apoyo 625718.

CONFLICTOS DE INTERESES

Ninguno. 

REFERENCIAS

1. Ibáñez A. 2022: Más tráfico en Internet que el acumulado en tres décadas. [Online]. 2019 [citado 2023 Mayo 6]. Disponible en: <https://blog.sarenet.es/trafico-web/>.
2. Extance A. How AI technology can tame the scientific literature. Nature. 2018 Septiembre;561:273-274.
3. Majchrzak A, Markus ML, Wareham J. Designing for Digital Transformation: Lessons for Information Systems Research from the Study of ICT and Societal Challenges. MIS Quarterly. 2016 Junio;40(2):267-277.
4. Castro D, Rodríguez-Gómez D, Gairín J. Exclusion Factors in Latin American Higher Education. Education and Urban Society. 2016 Julio; 49(2): p. 229-247.
5. Zhang K, Aslan AB. AI technologies for education: Recent research and future directions. Computers and Education: Artificial Intelligence. 2021;2:100025.
6. Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education - where are the educators? International Journal of Educational Technology in Higher Education. 2019 Octubre;16(39).
7. Salas-Pilco SZ, Yang Y. Artificial intelligence applications in Latin American higher education: a systematic review. International Journal of Educational Technology in Higher Education. 2022 Abril;19(21).
8. van de Schoot R, de Bruin J, Schram R, Zahedi P, de Boer J, Weijdema F, et al. An open source machine learning framework for efficient and transparent systematic reviews. Nature Machine Intelligence. 2021 Febrero;3:125-133.
9. Wallace BC, Trikalinos TA, Lau J, Brodley C, Schmid CH. Semi-automated screening of biomedical citations for systematic reviews. BMC Bioinformatics. 2010 Enero;11(55).
10. Larsen KR, Hovorka DS, Dennis AR, West JD. "Understanding the Elephant: The Discourse Approach to Boundary Identification and Corpus Construction for Theory Review Articles". Journal of the Association for Information Systems. 2019;887-927.
11. Zhang X, Yang A, Li S, Wang Y. Machine Reading Comprehension: a Literature Review. arXiv. 2019 Junio.
12. Wagner G, Lukyanenko R, Paré G. Artificial intelligence and the conduct of literature reviews. Journal of Information Technology. 2021 Octubre;37(2):209-226.
13. Okoli C. A Guide to Conducting a Standalone Systematic Literature Review. Communications of the Association for Information Systems. 2015;37.
14. Collins C, Dennehy D, Conboy K, Mikalef P. Artificial intelligence in information systems research: A systematic literature review and research agenda. International Journal of Information Management. 2021 Octubre;60:102383.
15. Van de Schoot R, De Bruin J. Researcher-in-the-loop for systematic reviewing of text databases. SciNLP: Natural Language Processing and Data Mining for Scientific Text. 2020 Junio.
16. Lombaers P, de Bruin J, van de Schoot R. Reproducibility and Data storage Checklist for Active Learning-Aided Systematic Reviews. 2023 Enero.
17. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Systematic Reviews. 2016 Diciembre;5.
18. Ought. Building Elicit, the AI research assistant. [Online]. 2022 [citado 2023 Mayo 6]. Disponible en: <https://ought.org/updates/2022-03-22-elicit>.
19. Zhao WX, Zhou K, Li J, Tang T, Wang X, Hou Y, et al. A Survey of Large Language Models. arXiv. 2023 Marzo.
20. Alkaissi H, McFarlane SI. Artificial Hallucinations in ChatGPT: Implications in Scientific Writing. Cureus. 2023 Febrero;15(2).
21. OpenAI. GPT-4 Technical Report. arXiv. 2023 Marzo.
22. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv. 2019.
23. Beltagy I, Lo K, Cohan A. SciBERT: A Pretrained Language Model for Scientific Text. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Proces-

- sing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP); 2019; Hong Kong, China: Association for Computational Linguistics. p. 3615-3620.
- 24. Nicholson JM, Mordaunt M, Lopez P, Uppala A, Rosati D, Rodrigues NP, et al. scite: A smart citation index that displays the context of citations and classifies their intent using deep learning. *Quantitative Science Studies*. 2021;2(3):882-898.
 - 25. Marshall IJ, Kuiper J, Wallace BC. RobotReviewer: evaluation of a system for automatically assessing bias in clinical trials. *Journal of the American Medical Informatics Association*. 2016 Enero;23(1):193-201.
 - 26. Boell S, Wang B. www.litbaskets.io, an IT Artifact Supporting Exploratory Literature Searches for Information Systems Research. In 30th Australasian Conference on Information Systems; 2019 December.
 - 27. Li J, Larsen K, Abbasi A. TheoryOn: A Design Framework and System for Unlocking Behavioral Knowledge Through Ontology Learning. *MIS Quarterly*. 2020 Diciembre;44:1733-1772.
 - 28. d'Obrenan CvdB, Galliano F, Minton J, Botev V, Wu R. Searching for Carriers of the Diffuse Interstellar Bands Across Disciplines, using Natural Language Processing. *arXiv*. 2022 Noviembre.
 - 29. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. *Nature Machine Intelligence*. 2019 Septiembre; 1:389-399.