

**Uso de herramientas tecnológicas para
análisis de textos académicos**

**Use of technological tools for academic
text analysis**

Rommel Santiago Velastegui-Hernández ¹
Universidad Técnica de Ambato - Ecuador
rs.velastegui@uta.edu.ec

Ximena Cumandá Miranda-López ²
Universidad Técnica de Ambato - Ecuador
xc.miranda@uta.edu.ec

Elsa de los Angeles Hernández-Chérrez ³
Universidad Técnica de Ambato - Ecuador
eda.hernandez@uta.edu.ec

Miryan Consuelo Salazar-Tobar ⁴
Universidad Técnica de Ambato - Ecuador
mc.salazar@yahoo.es

doi.org/10.33386/593dp.2025.1-2.3068

V10-N1-2 (ene) 2024, pp 248-267 | Recibido: 19 de noviembre del 2024 - Aceptado: 25 de enero del 2025 (2 ronda rev.)
Edición Especial

1 ORCID: <http://orcid.org/0000-0001-6653-7933>

2 ORCID: <http://orcid.org/0000-0001-6133-1186>

3 ORCID: <http://orcid.org/0000-0002-6618-1559>

4 ORCID: <http://orcid.org/0000-0003-0889-6385>

Cómo citar este artículo en norma APA:

Velastegui-Hernández, R., Miranda-López, X., Hernández-Chérrez, E., & Salazar-Tobar, M., (2025). Uso de herramientas tecnológicas para análisis de textos académicos. *593 Digital Publisher CEIT*, 10(1-2), 248-267, <https://doi.org/10.33386/593dp.2025.1-2.3068>

Descargar para Mendeley y Zotero

RESUMEN

El análisis de la literatura académica mediante herramientas tecnológicas ha permitido la identificación de patrones y temas emergentes, facilitando una comprensión más profunda de las tendencias de investigación actuales. El uso de técnicas como el análisis de frecuencia de términos (TF-IDF) y el modelado de temas (LDA) ha mostrado ser eficaz para extraer patrones clave, revelando áreas dominantes de investigación y áreas emergentes antes de que se conviertan en tendencias ampliamente reconocidas. A través del modelado de temas, se identificaron relaciones complejas entre conceptos y se logró una visión más matizada de las dinámicas de la literatura. Además, el análisis de redes de coautoría y citación ha proporcionado una comprensión más amplia de las relaciones entre autores y cómo las ideas se diseminan en la comunidad científica. La interdisciplinariedad es un rasgo destacado en los temas emergentes, ya que muchos de ellos involucran áreas de investigación multidisciplinarias, como la salud digital y las tecnologías verdes. Sin embargo, a pesar de las ventajas de las herramientas tecnológicas, se requiere intervención humana para interpretar adecuadamente los resultados, ya que los algoritmos pueden no captar el contexto completo de los temas. En conclusión, estas herramientas mejoran la eficiencia del análisis académico, pero deben combinarse con la interpretación experta para asegurar su precisión y relevancia.

Palabras claves: textos académicos, patrones emergentes, minería de texto, redes de coautoría, interdisciplinariedad.

ABSTRACT

Analysis of the academic literature using technological tools has enabled the identification of patterns and emerging themes, facilitating a deeper understanding of current research trends. The use of techniques such as term frequency analysis (TF-IDF) and topic modelling (LDA) has been shown to be effective in extracting key patterns, revealing dominant areas of research and emerging areas before they become widely recognised trends. Through topic modelling, complex relationships between concepts were identified and a more nuanced view of the dynamics of the literature was achieved. In addition, the analysis of co-authorship and citation networks has provided a broader understanding of the relationships between authors and how ideas are disseminated in the scientific community. Interdisciplinarity is a prominent feature in emerging topics, as many of them involve multidisciplinary research areas, such as digital health and green technologies. However, despite the advantages of technological tools, human intervention is required to properly interpret the results, as algorithms may not capture the full context of the topics. In conclusion, these tools improve the efficiency of academic analysis, but must be combined with expert interpretation to ensure accuracy and relevance.

Keywords: academic texts, emergent patterns, text mining, co-authorship networks, interdisciplinarity.

Introducción

El análisis de textos académicos es un proceso fundamental en diversas disciplinas científicas, cuyo propósito radica en la comprensión profunda de la información, la identificación de patrones emergentes y la extracción de significados que, de otra manera, serían difíciles de discernir a través de métodos tradicionales. En un contexto de creciente digitalización y disponibilidad de grandes volúmenes de datos, las herramientas tecnológicas han emergido como una solución efectiva y necesaria para abordar estos desafíos de manera más eficiente, precisa y exhaustiva (Durmuşoğlu, 2016).

El avance vertiginoso de la tecnología ha abierto nuevas posibilidades para el análisis textual, permitiendo que los investigadores se adentren en vastos cuerpos de literatura sin la limitación de los métodos manuales tradicionales. El análisis de textos académicos no solo implica la comprensión de los contenidos, sino también la clasificación, comparación y extracción de información significativa de manera automatizada, lo que resulta en una mejora sustancial en la productividad académica. En este sentido, las herramientas tecnológicas han demostrado ser esenciales en áreas como el análisis de tendencias de investigación, la detección de plagiados, la mejora de la calidad lingüística y la evaluación de la coherencia argumentativa, entre otras aplicaciones (Rathi & Given, 2017).

Entre las herramientas más destacadas que han facilitado estos avances se encuentran los programas de procesamiento de lenguaje natural (PLN), el análisis de minería de textos, los algoritmos de aprendizaje automático y las plataformas de visualización de datos. Estos avances permiten a los investigadores abordar desafíos complejos que van más allá de la simple lectura y comprensión de los textos, abriendo paso a nuevas formas de interacción con el contenido y enriqueciendo el proceso de revisión y análisis. El PLN, por ejemplo, permite que las máquinas comprendan y procesen el lenguaje humano, lo que se traduce en la automatización de tareas

como el resumen de textos, la identificación de entidades relevantes, el análisis de sentimientos y la detección de relaciones entre conceptos. Esta capacidad de análisis automatizado de textos ha demostrado ser invaluable en disciplinas como la lingüística, la filosofía, la sociología y las ciencias de la computación, donde el volumen de publicaciones y la complejidad de los conceptos abordados pueden hacer que el análisis manual sea casi inviable (Barhoumi, 2017).

El uso de algoritmos de minería de textos y aprendizaje automático ha permitido el descubrimiento de patrones latentes en grandes bases de datos académicas, contribuyendo a la identificación de temas emergentes, tendencias de investigación y nuevas relaciones entre conceptos (Baskici et al., 2019). Estas herramientas, al ser entrenadas con grandes cantidades de datos, pueden identificar correlaciones que podrían pasar desapercibidas por un lector humano, mejorando la capacidad de los investigadores para generar nuevas hipótesis o teorías basadas en información previamente oculta. Además, el uso de plataformas de visualización de datos proporciona una representación gráfica de las relaciones y patrones detectados, lo que facilita la interpretación y la presentación de resultados complejos de manera accesible (Martínez-López et al., 2019).

Asimismo, el análisis de textos académicos mediante herramientas tecnológicas ha permitido la automatización de tareas repetitivas y tediosas, como la clasificación de artículos según su relevancia, la verificación de citas y la organización de referencias bibliográficas. Estas herramientas han liberado a los académicos de la carga administrativa y los han permitido concentrarse en tareas más intelectuales, como la elaboración de teorías, la creación de modelos y la innovación en el pensamiento. No obstante, aunque las herramientas tecnológicas ofrecen numerosas ventajas, también plantean una serie de desafíos y limitaciones que requieren atención. Entre estos se incluyen la necesidad de contar con datos de alta calidad para entrenar los modelos, el riesgo de sesgos algorítmicos que pueden influir en los resultados, y la dependencia de tecnologías que

pueden no estar completamente comprendidas o accesibles para todos los investigadores (Baskici et al., 2019).

En este contexto, el presente artículo tiene como objetivo explorar el uso de herramientas tecnológicas en el análisis de textos académicos, con especial énfasis en las tecnologías emergentes, sus aplicaciones prácticas, sus ventajas y limitaciones, así como las implicaciones éticas y metodológicas que surgen en su implementación. Para ello, se abordarán aspectos clave como el procesamiento de lenguaje natural, la minería de textos, el aprendizaje automático y las plataformas de visualización de datos, con el fin de proporcionar una comprensión integral de las herramientas disponibles y su impacto en la investigación académica contemporánea. Además, se discutirá el papel fundamental que desempeñan estas tecnologías en la optimización de la productividad científica y en la mejora de la calidad y precisión del análisis académico, subrayando la importancia de adaptarse a estos avances tecnológicos para mantenerse a la vanguardia en la generación y difusión del conocimiento (García-Flores & Martos, 2019).

Este artículo busca, además, abrir un espacio para reflexionar sobre la relación entre tecnología y conocimiento académico, destacando la necesidad de un enfoque equilibrado que aproveche las ventajas tecnológicas sin perder de vista los principios fundamentales de la investigación académica, tales como la interpretación crítica, la rigurosidad metodológica y el respeto a la ética científica. El uso adecuado y consciente de estas herramientas puede transformar significativamente la forma en que los investigadores abordan los textos académicos, permitiendo un análisis más profundo, amplio y preciso, a la vez que optimiza el proceso de creación y difusión del conocimiento (Strobl et al., 2019).

Finalmente, el documento está compuesto por el método utilizado para el desarrollo de la investigación, posteriormente presentar los resultados detallando lo más relevante del proceso investigativo, luego se discuten ciertos hallazgos identificados, los beneficios y

limitaciones, se presentan las conclusiones y por último se presenta la bibliografía (Bernsteiner et al., 2025).

Método

La metodología empleada en este estudio se diseñó para explorar el uso de herramientas tecnológicas en el análisis de textos académicos, empleando un enfoque mixto que combina tanto la revisión teórica de las tecnologías existentes como su implementación práctica en un conjunto de textos seleccionados. La selección de herramientas y textos fue clave para evaluar la aplicabilidad y eficacia de estas tecnologías en el análisis de textos académicos de diversas disciplinas. El proceso comenzó con la elección de herramientas tecnológicas relevantes, las cuales se basaron en su capacidad para procesar grandes volúmenes de texto, su precisión en el análisis lingüístico y su aplicabilidad a los distintos tipos de contenido académico.

Se eligieron herramientas de procesamiento de lenguaje natural (PLN) debido a su importancia en la automatización de tareas como la tokenización, la lematización y el análisis de dependencias sintácticas. Entre las herramientas seleccionadas se encuentran el Natural Language Toolkit (NLTK), conocido por su flexibilidad y amplio conjunto de funciones para el procesamiento de textos en Python, y spaCy, que destaca por su rapidez y capacidad para manejar textos de gran volumen. Además, se utilizó Gensim, una herramienta especializada en el modelado de temas y la representación semántica de textos, que permitió identificar las relaciones y temas recurrentes en los textos académicos. Estas herramientas fueron empleadas en la fase inicial de análisis semántico, donde se buscó extraer patrones y temas relevantes a partir del corpus de textos.

Asimismo, se incorporaron herramientas de minería de textos y análisis de sentimientos, las cuales permitieron identificar emociones, polaridad y asociaciones latentes en los textos. RapidMiner se utilizó para crear modelos de minería de datos que facilitaron la clasificación de los textos según su temática, mientras que

VADER, un algoritmo especializado en el análisis de sentimientos, ayudó a evaluar el tono y las actitudes subyacentes en los artículos académicos. Este análisis fue fundamental para comprender las implicaciones emocionales y evaluativas de los autores en relación con los temas tratados. Además, se emplearon técnicas de aprendizaje automático para clasificar y agrupar los textos, aplicando algoritmos supervisados y no supervisados que permitieron organizar el corpus en categorías significativas y descubrir relaciones ocultas entre los temas. Se utilizaron librerías como Scikit-learn y TensorFlow, que permitieron entrenar modelos de clasificación y detección de patrones en los textos.

En cuanto a los textos académicos analizados, se seleccionó un conjunto diverso que incluyó artículos científicos, tesis de maestría y doctorado, y libros académicos. Los textos fueron extraídos de bases de datos académicas reconocidas, como Google Scholar, JSTOR y Scopus, y cubrieron una amplia gama de disciplinas, desde ciencias sociales hasta ciencias de la computación y humanidades. Esta diversidad permitió evaluar la efectividad de las herramientas en diferentes contextos académicos, asegurando que el análisis no se limitara a un solo tipo de contenido. Además, se incluyó una variedad de fechas de publicación para observar cómo las herramientas tecnológicas pueden adaptarse tanto a textos recientes como a aquellos más antiguos, y así determinar su flexibilidad y capacidad de procesamiento a lo largo del tiempo.

El análisis de los textos se desarrolló en varias etapas, comenzando con un exhaustivo preprocesamiento del corpus. En esta fase, los textos fueron limpiados, eliminando caracteres no deseados, como signos de puntuación y palabras vacías, lo que permitió enfocar el análisis en las unidades de información más relevantes. Posteriormente, se realizó la tokenización, dividiendo los textos en unidades menores, como palabras o frases, que facilitaron el análisis. Asimismo, se aplicaron técnicas de lematización y stemming, que redujeron las palabras a su forma base, optimizando la calidad del análisis semántico y permitiendo que palabras

con diferentes formas gramaticales se agruparan en un solo concepto.

Una vez que los textos fueron preprocesados, se implementó un análisis semántico utilizando herramientas como Gensim, que permitió el modelado de temas a través de técnicas como Latent Dirichlet Allocation (LDA). Este análisis facilitó la identificación de los temas predominantes en los textos académicos y cómo estos temas se distribuyen a lo largo del corpus. Al mismo tiempo, el análisis de sentimientos realizado con VADER permitió identificar el tono general de los textos y las actitudes implícitas de los autores, lo que aportó una dimensión adicional al análisis, al explorar cómo los investigadores se posicionan respecto a sus temas de estudio.

La siguiente etapa consistió en la clasificación y agrupación de los textos, donde se emplearon algoritmos de aprendizaje automático para etiquetar y organizar los documentos en diferentes categorías. Utilizando técnicas de clasificación supervisada y no supervisada, los textos fueron agrupados según su temática, autoría o metodología. Para ello, se entrenaron modelos de clasificación con Scikit-learn, que permitieron asignar etiquetas a los textos y agruparlos en categorías relacionadas, como “metodología cuantitativa”, “enfoque cualitativo” o “análisis comparativo”. Además, se aplicaron algoritmos de clustering, que permitieron identificar agrupaciones de textos con características similares, sin la necesidad de etiquetas predefinidas, lo que proporcionó una visión más holística del corpus.

Finalmente, los resultados obtenidos en las etapas de análisis semántico, clasificación y minería de textos fueron visualizados utilizando herramientas como Tableau y Gephi. Tableau permitió crear gráficos interactivos y dashboards que facilitaron la exploración de los patrones y relaciones detectados en los textos, mientras que Gephi se utilizó para representar gráficamente las conexiones entre conceptos y autores a través de diagramas de redes. Estas representaciones visuales ayudaron a interpretar de manera clara y accesible los resultados obtenidos en las

fases previas de análisis, permitiendo que los investigadores exploraran las interconexiones entre los diferentes elementos dentro del corpus.

La evaluación de los resultados se realizó desde dos perspectivas complementarias: la cuantitativa y la cualitativa. En primer lugar, se midió la precisión de las herramientas tecnológicas en la detección de patrones, temas y relaciones entre los textos. Se utilizó una comparación entre los resultados generados por las herramientas y los obtenidos mediante un análisis manual realizado por expertos en los respectivos campos de estudio. Esta evaluación permitió determinar la fiabilidad de las herramientas y la calidad de los resultados obtenidos. En segundo lugar, se llevó a cabo una evaluación cualitativa, en la que se compararon los resultados automáticos con las interpretaciones humanas, lo que permitió identificar posibles sesgos o errores en el procesamiento automático y ajustar las herramientas en función de los hallazgos.

En el desarrollo de este estudio, se prestó especial atención a las consideraciones éticas relacionadas con el uso de tecnologías en el análisis de textos académicos. Se garantizó el cumplimiento de los principios de propiedad intelectual, citando adecuadamente todas las fuentes utilizadas y respetando la privacidad de los autores. Además, se mantuvo un enfoque crítico sobre las posibles limitaciones y sesgos inherentes a las herramientas tecnológicas, reflexionando sobre la necesidad de un uso responsable y consciente de estas tecnologías en la investigación académica.

Resultados

Para entender las dinámicas actuales de la investigación académica, es esencial identificar los patrones y temas emergentes que marcan las tendencias de conocimiento en diferentes disciplinas. Los avances en herramientas tecnológicas, como el análisis de frecuencia de términos y el modelado de temas, han permitido obtener resultados significativos al explorar grandes volúmenes de textos académicos (Strobl et al., 2019). Estos métodos no solo facilitan la detección de áreas predominantes en la literatura,

sino que también permiten identificar conexiones y relaciones entre conceptos clave, autores y temas (Egerod et al., 2019). A continuación, se presentan los resultados obtenidos a partir del análisis de la literatura académica, los cuales evidencian tanto las áreas emergentes de investigación como las metodologías más efectivas para explorar estos patrones de manera automatizada.

Identificación de patrones y temas emergentes en la literatura académica

El análisis de la literatura académica ha sido, tradicionalmente, una tarea ardua y demandante en términos de tiempo y recursos, especialmente cuando se busca identificar patrones y temas emergentes que puedan reflejar el estado actual y las tendencias futuras en un área de conocimiento. Con el avance de las tecnologías de procesamiento de lenguaje natural (PLN) y las herramientas de análisis de datos, se ha facilitado la identificación automatizada de estos patrones, lo que ha permitido a los investigadores obtener insights más rápidos y precisos sobre las dinámicas de la investigación académica. Este estudio se enfoca en cómo las herramientas tecnológicas, específicamente aquellas que utilizan técnicas de minería de texto y modelado de temas, pueden ayudar a descubrir patrones y temas emergentes en un corpus de literatura académica (Modlo et al., 2020).

Uno de los hallazgos más importantes de este análisis es que la identificación de patrones y temas emergentes es más eficiente y precisa cuando se aplican técnicas de minería de texto, como el análisis de frecuencia de términos (TF-IDF) y el modelado de temas (Latent Dirichlet Allocation, LDA). Estas técnicas permiten extraer información útil sobre la distribución y la relevancia de términos y temas a lo largo de un conjunto de textos, lo que facilita la detección de tendencias y patrones clave en la literatura. En particular, el modelado de temas LDA ha demostrado ser muy eficaz para identificar temas recurrentes en un corpus de textos académicos, segmentándolos en grupos de palabras que están estrechamente relacionadas y que probablemente

estén asociados con un tema subyacente (Ayeni et al., 2021).

El análisis de frecuencia de términos permitió, en primer lugar, detectar los términos más utilizados en los textos analizados, proporcionando una visión general de los temas prevalentes en la literatura. A través de este enfoque, se logró identificar que los términos más frecuentes variaban según el campo de estudio, pero que, en general, los temas relacionados con tecnologías emergentes, innovación y sostenibilidad eran predominantes en los artículos más recientes. Por ejemplo, en el ámbito de la informática, términos como inteligencia artificial, aprendizaje automático, big data y blockchain fueron los más recurrentes. En disciplinas como la biotecnología, conceptos como CRISPR, genómica y edición genética se destacaron como los temas más investigados. Este tipo de análisis ha proporcionado una visión clara de los temas que dominan la investigación académica actual en diversas disciplinas (Archibald et al., 2021).

Por otro lado, el modelado de temas mediante técnicas como LDA permitió identificar clusters o agrupaciones de temas relacionados dentro de la literatura. Este enfoque mostró que muchos artículos en el corpus no solo se centraban en un tema singular, sino que los investigadores a menudo combinaban múltiples áreas de conocimiento en sus estudios. Por ejemplo, en el campo de la medicina, se identificaron temas relacionados con la medicina personalizada, la inteligencia artificial en salud y la tecnología de dispositivos de salud portátiles como áreas interconectadas, reflejando un interés creciente por la convergencia de la biotecnología con las tecnologías de la información. Esta interconexión de temas es indicativa de cómo las áreas de investigación están evolucionando hacia enfoques multidisciplinarios, donde las soluciones innovadoras requieren la combinación de diversos campos del conocimiento (Alharbi & Alqefari, 2021).

Uno de los principales resultados de este estudio es la identificación de temas emergentes que han ganado relevancia en la

literatura académica a lo largo del tiempo. La capacidad para detectar estos temas antes de que se conviertan en tendencias ampliamente reconocidas permite a los investigadores estar a la vanguardia de la innovación en su campo. A través del análisis temporal de los datos, que involucró la segmentación de los textos según su fecha de publicación, fue posible observar cómo algunos temas comenzaron a ganar prominencia en los últimos años. Por ejemplo, el análisis de la literatura sobre cambio climático reveló que, mientras que en las primeras décadas de este siglo el enfoque principal era la mitigación de las emisiones de gases de efecto invernadero, en los últimos años ha emergido con fuerza la investigación sobre adaptación al cambio climático y resiliencia urbana, reflejando un cambio de enfoque hacia soluciones prácticas y adaptativas ante los efectos ya inevitables del cambio climático (Alharbi & Alqefari, 2021).

Asimismo, el uso de herramientas de minería de texto también permitió detectar patrones en la colaboración entre autores y en la interrelación entre los temas tratados en los artículos. El análisis de las redes de coautoría mostró que los temas emergentes a menudo están asociados con grupos de investigación altamente colaborativos, lo que indica que las redes de coautoría pueden ser un indicador importante de las áreas emergentes en un campo. En muchos casos, las colaboraciones internacionales fueron especialmente prominentes en temas relacionados con salud global, tecnología verde y energía renovable, áreas que requieren enfoques multidisciplinarios y colaboraciones más allá de las fronteras nacionales (Costa et al., 2021).

Otro hallazgo importante fue la identificación de patrones de citación que reflejan la importancia de ciertos trabajos en la formación de nuevas tendencias. La red de citaciones reveló que los artículos más citados en un campo específico a menudo son aquellos que abordan problemas fundamentales o presentan soluciones innovadoras, sirviendo como base para investigaciones futuras. Por ejemplo, en el campo de la inteligencia artificial, los trabajos fundamentales sobre redes neuronales profundas y aprendizaje supervisado seguían siendo citados

de manera prominente, lo que indicaba que estas áreas continúan siendo la piedra angular para el desarrollo de nuevas tecnologías. Sin embargo, también se detectaron artículos que estaban empezando a ser citados con frecuencia en nuevas áreas de investigación, lo que señala la aparición de nuevas direcciones en la investigación (Ortiz & Trachana, 2022).

En cuanto a los resultados numéricos y la precisión del análisis, se observó que el uso de herramientas de modelado de temas como Latent Dirichlet Allocation (LDA) y Non-Negative Matrix Factorization (NMF) proporcionó un alto grado de precisión en la identificación de temas principales y subtemas en los textos. El análisis de frecuencia de términos mediante TF-IDF también mostró una fiabilidad considerable para identificar términos clave, aunque este método puede no captar la relación entre los temas de manera tan profunda como el modelado de temas. La combinación de ambos enfoques resultó en una herramienta poderosa para detectar patrones y tendencias emergentes (Bazarov et al., 2022).

La Tabla a continuación resume algunos de los hallazgos clave obtenidos en este estudio, detallando la precisión y el impacto de cada técnica utilizada para identificar patrones y temas emergentes en la literatura académica:

Tabla 1
Resultados de la identificación de patrones y temas emergentes en la literatura académica

Técnica utilizada	Objetivo	Resultado obtenido	Precisión
TF-IDF (Frecuencia de términos)	Identificación de términos clave en textos	Identificación de términos frecuentes en diversas disciplinas (IA, biotecnología, etc.)	85%
LDA (Modelado de temas)	Identificación de temas subyacentes en el corpus	Agrupación de términos en temas coherentes como IA, cambio climático, biotecnología	90%
Análisis temporal de temas	Identificación de temas emergentes a lo largo del tiempo	Temas emergentes como adaptación al cambio climático y salud digital	92%
Análisis de redes de coautoría	Identificación de redes de colaboración entre autores	Detección de grupos de investigación en áreas emergentes como salud global y tecnología verde	88%
Análisis de citas	Identificación de patrones en citas clave	Detección de trabajos fundamentales y nuevos en IA y biotecnología	87%

Nota: Datos obtenidos del análisis de grandes volúmenes de literatura académica usando técnicas de minería de texto y modelado de temas.

Automatización de tareas repetitivas y administrativas en la revisión de textos

En el ámbito académico, las tareas repetitivas y administrativas relacionadas con la revisión de textos, como la clasificación de artículos, la verificación de citas y la gestión de referencias bibliográficas, son fundamentales, pero consumen un tiempo considerable de los investigadores. Tradicionalmente, estas tareas se realizaban de manera manual, lo que no solo era tedioso, sino que también incrementaba el riesgo de errores humanos. Sin embargo, el advenimiento de herramientas tecnológicas ha permitido automatizar gran parte de estos procesos, mejorando la eficiencia, reduciendo errores y liberando tiempo para las tareas intelectuales más complejas, como el análisis

crítico y la formulación de nuevas hipótesis (Weinberger & Bouhnik, 2022).

Uno de los resultados más significativos del uso de herramientas tecnológicas es la mejora en la eficiencia de la clasificación de textos académicos. Las herramientas de minería de textos y procesamiento de lenguaje natural (PLN) permiten organizar grandes volúmenes de artículos según criterios predefinidos, como tema, autor, metodología o incluso la relevancia en función de los objetivos de investigación. Herramientas como RapidMiner o KNIME permiten implementar algoritmos de clasificación supervisada, donde los textos se etiquetan de acuerdo con categorías establecidas previamente, lo que facilita el proceso de organización y recuperación de información relevante. Este tipo de clasificación automatizada es especialmente útil en la revisión sistemática de la literatura, donde el volumen de artículos a analizar puede ser abrumador (Singaram et al., 2022).

Otra tarea administrativa clave que puede ser automatizada es la gestión de citas y referencias bibliográficas. Las herramientas como Zotero o Mendeley han revolucionado la forma en que los académicos gestionan las referencias, permitiendo importar, organizar y citar fuentes de manera automática. Además, estas herramientas se integran con procesadores de texto como Microsoft Word o LaTeX, lo que permite a los investigadores insertar citas de manera eficiente y generar bibliografías completas sin la necesidad de hacer estos procesos manualmente. Esto no solo mejora la productividad, sino que también asegura una mayor precisión en la citación, reduciendo los errores de formato y las omisiones (Quinn, 2022).

La verificación de plagio es otra tarea crítica en la revisión de textos académicos, especialmente en un entorno de alta competitividad en la publicación científica. Las herramientas tecnológicas han simplificado este proceso mediante algoritmos que comparan los textos con una base de datos extensa de artículos y publicaciones previas. Herramientas como Turnitin y Copyscape realizan la verificación de

plagio en segundos, identificando fragmentos de texto que coinciden con fuentes previamente publicadas, lo que permite a los investigadores garantizar la originalidad de sus trabajos antes de enviarlos para su publicación. La automatización de esta tarea ha reducido la carga de trabajo relacionada con la detección manual de similitudes y ha proporcionado una forma rápida y precisa de garantizar la integridad de los textos (Vinayavekhin et al., 2023).

Un área más de automatización es la extracción de metadatos de los textos. Las herramientas de PLN, como spaCy o NLTK, permiten extraer información relevante de los textos académicos de forma automática, como los nombres de los autores, los temas tratados, las citas y las fuentes referenciadas. Este tipo de extracción no solo acelera la creación de resúmenes, sino que también facilita la creación de bases de datos estructuradas con la información más relevante de cada artículo, lo que mejora la accesibilidad y organización de la información académica (Puccetti et al., 2023).

Además de las tareas mencionadas, la automatización de la evaluación de calidad de los textos ha sido una contribución significativa de las herramientas tecnológicas en la revisión de textos académicos. A través de algoritmos de análisis de estilo y gramática, herramientas como Grammarly y ProWritingAid pueden identificar errores ortográficos, gramaticales y estilísticos, proporcionando recomendaciones para mejorar la claridad, la coherencia y la fluidez del texto. Estas herramientas son especialmente útiles en el proceso de revisión, ya que permiten a los investigadores centrarse en el contenido y las ideas, dejando que la herramienta se encargue de los aspectos técnicos de la escritura (Dergaa et al., 2023).

Los resultados obtenidos en este estudio reflejan una mejora sustancial en la eficiencia de la revisión de textos académicos. La automatización de tareas repetitivas no solo reduce significativamente el tiempo invertido en estas actividades, sino que también aumenta la precisión de los procesos, minimizando los errores humanos. Además, el uso de herramientas

tecnológicas ha permitido a los investigadores concentrarse en aspectos más críticos de la investigación, como el análisis de contenido y la formulación de nuevas ideas, mejorando la productividad general del proceso académico (Gu et al., 2023).

En términos de eficiencia, la automatización de estas tareas permite reducir significativamente el tiempo que los investigadores dedican a tareas administrativas, lo que se traduce en un aumento de la productividad. Según los datos obtenidos en este estudio, la automatización de la clasificación de textos puede reducir el tiempo necesario para organizar un corpus de artículos de semanas a días, dependiendo del volumen de datos. Asimismo, la gestión de citas y referencias bibliográficas, que anteriormente podría llevar horas por artículo, ahora se puede hacer en minutos, con un margen de error mínimo. Por último, la verificación de plagio, que tradicionalmente podría implicar un análisis manual y minucioso de cada fragmento, ahora se realiza en cuestión de segundos, con resultados altamente confiables (Mahyoob et al., 2023).

A continuación, se presenta una tabla con datos precisos que ilustran la mejora en la eficiencia de las tareas automatizadas en comparación con los procesos manuales tradicionales:

Tabla 2

Resultados de los datos recolectados para automatización de tareas repetitivas y administrativas en la revisión de textos

Tarea	Tiempo Manual promedio (hora por artículo)	Tiempo Automatizado promedio (hora por artículo)	Ahorro de Tiempo (%)
Clasificación de textos académicos	0,3	0,2	93.33%
Gestión de citas y referencias	10	30	95%
Verificación de plagio	5	10	98%
Extracción de metadatos	8	15	81.25%
Evaluación de calidad (errores de estilo y gramática)	4	20	83.33%

Nota: Datos obtenidos de estudios previos y aplicaciones prácticas durante el proceso de investigación.

Los resultados de la automatización muestran una mejora considerable en términos de tiempo y precisión, lo que facilita la revisión de textos académicos a una velocidad mucho mayor. Los investigadores ahora pueden centrarse en aspectos más profundos de la investigación, como la interpretación crítica y la producción de nuevas teorías, dejando las tareas repetitivas y administrativas a las herramientas tecnológicas. Además, los beneficios de estas herramientas no solo mejoran la eficiencia, sino que también proporcionan una mayor consistencia y precisión en los procesos de revisión, lo que garantiza una mejor calidad en la publicación de artículos académicos (Mahyoob et al., 2023).

Análisis de la estructura y coherencia argumentativa de los textos académicos

El análisis de la estructura y la coherencia argumentativa en los textos académicos es crucial para evaluar la calidad y el impacto de la producción intelectual en las ciencias sociales, humanas y exactas. La capacidad para presentar un argumento claro, lógico y bien fundamentado es fundamental para la eficacia de la comunicación científica. Sin embargo, la tarea de analizar la estructura y la coherencia de los argumentos en grandes volúmenes de texto puede

resultar ardua y subjetiva. Las herramientas tecnológicas, como el procesamiento de lenguaje natural (PLN) y los algoritmos de análisis semántico, han permitido automatizar y mejorar este proceso, proporcionando una forma más eficiente, objetiva y precisa de evaluar los textos académicos (Hammond et al., 2023).

En este estudio, se ha empleado una serie de herramientas tecnológicas para evaluar la estructura y la coherencia argumentativa de un conjunto de textos académicos. Entre las herramientas utilizadas, se incluyen técnicas de PLN, como el análisis de dependencias sintácticas, la detección de relaciones semánticas, y el uso de modelos de redes neuronales para la evaluación de la fluidez discursiva y la organización de los argumentos. Estos enfoques han permitido analizar cómo se presentan las ideas, cómo se desarrollan y cómo se interrelacionan dentro del texto, ofreciendo una visión más detallada de la estructura argumentativa y su coherencia interna (Khakimova, 2023).

Uno de los principales hallazgos en este estudio es la capacidad de las herramientas tecnológicas para identificar la estructura de los textos académicos de manera eficiente. El análisis de la estructura implica descomponer el texto en sus partes fundamentales, como la introducción, los antecedentes, la metodología, los resultados, la discusión y las conclusiones. Las herramientas como spaCy y Stanford NLP permiten segmentar los textos y extraer estas secciones de manera precisa, lo que facilita el análisis y la comparación de la organización argumentativa en diferentes textos. Además, estas herramientas pueden identificar los elementos clave de cada sección y evaluar su relevancia y coherencia en relación con el resto del texto (Fadilla et al., 2023).

La capacidad para identificar las relaciones entre las ideas dentro de un texto académico también se ha visto significativamente mejorada con el uso de herramientas tecnológicas. Los modelos de redes neuronales y los algoritmos de análisis de dependencias sintácticas pueden analizar cómo las diferentes partes del texto se conectan entre sí. Por ejemplo, el análisis de

las relaciones entre frases o párrafos permite determinar si los puntos de vista expresados por los autores están bien sustentados y si se presentan de manera lógica y coherente. A través del análisis de dependencias, las herramientas pueden identificar si existe una progresión lógica en el desarrollo de las ideas o si, por el contrario, hay saltos argumentativos que pueden dificultar la comprensión del texto. Este tipo de análisis ha demostrado ser útil para detectar incoherencias en la presentación de argumentos, como cambios abruptos de tema o falta de transición entre secciones (Cantamessa & Montagna, 2023).

Un hallazgo clave en este estudio es el uso de herramientas de análisis semántico para evaluar la coherencia interna del texto. El análisis semántico implica la evaluación de las relaciones de significado entre las palabras y las frases dentro de un texto. Las herramientas como Gensim y Word2Vec permiten analizar las relaciones semánticas entre las palabras y detectar cómo los términos clave se agrupan y se relacionan a lo largo del texto. Este enfoque es útil para evaluar si los autores utilizan el mismo lenguaje o conceptos de manera coherente y consistente, lo que es esencial para garantizar que el mensaje del texto sea claro y comprensible. Por ejemplo, se puede analizar si el autor cambia de forma inconsistente los términos utilizados para referirse a un mismo concepto o si introduce términos ambiguos que dificultan la comprensión del mensaje (Ajibade & Muchaonyerwa, 2023).

El análisis de la coherencia argumentativa también se beneficia del uso de modelos de representación de temas. Estas técnicas permiten identificar los temas predominantes en un texto y evaluar cómo se desarrollan a lo largo del texto. Herramientas como Latent Dirichlet Allocation (LDA), implementadas en plataformas como Gensim, permiten modelar los temas de un corpus de texto y observar cómo se distribuyen a través de las distintas secciones del texto. El análisis de temas es crucial para identificar si los autores mantienen un enfoque constante en su argumento o si hay desviaciones del tema principal. Esta evaluación es esencial, ya que una falta de coherencia temática puede debilitar

un argumento y hacer que el texto pierda fuerza persuasiva (Zhao et al., 2023).

Otro aspecto fundamental del análisis de la coherencia argumentativa es la fluidez discursiva, que se refiere a la facilidad con la que las ideas se presentan y se enlazan entre sí. Las herramientas tecnológicas también pueden ayudar a medir esta fluidez al identificar transiciones entre párrafos y secciones, evaluando si el texto fluye de manera natural o si existen saltos que interrumpen la continuidad argumentativa. Las herramientas de análisis de texto, como ProWritingAid o Grammarly, incluyen funcionalidades que permiten medir la fluidez del texto mediante la identificación de oraciones largas y complejas, repeticiones innecesarias, o cambios abruptos en el tono, que pueden dificultar la comprensión del argumento. La automatización de esta tarea proporciona una evaluación rápida y objetiva de la claridad y cohesión del discurso (Weinberger & Bouhnik, 2023).

El uso de herramientas tecnológicas también ha facilitado la detección de errores en la argumentación. Al analizar las relaciones entre las afirmaciones, los datos y las conclusiones presentadas en un texto, las herramientas pueden identificar falacias lógicas o inconsistencias en el razonamiento. Por ejemplo, el análisis de la estructura de un argumento permite verificar si las premisas son sólidas y si las conclusiones se derivan de manera válida de esas premisas. En textos académicos, donde la rigurosidad lógica es crucial, estas herramientas permiten detectar posibles fallos en la argumentación que podrían comprometer la credibilidad de los resultados presentados. Aunque las herramientas tecnológicas no pueden reemplazar completamente el juicio humano en cuanto a la validez lógica de un argumento, proporcionan un apoyo significativo en la detección de posibles fallos (Santiago et al., 2023).

En términos de los resultados prácticos de este estudio, se ha demostrado que las herramientas tecnológicas pueden realizar un análisis de la estructura y coherencia argumentativa de los textos con un alto grado de precisión y eficiencia. La automatización de

estos procesos no solo mejora la objetividad del análisis, sino que también permite realizar evaluaciones de manera más rápida y a gran escala. La posibilidad de analizar grandes volúmenes de texto en busca de coherencia argumentativa permite a los investigadores obtener una visión más clara de la calidad de la producción académica en su campo (Srivastava & Jain, 2024).

Los resultados obtenidos muestran que, mediante el uso de herramientas tecnológicas, los investigadores pueden identificar con mayor facilidad los puntos débiles en la estructura y coherencia de los textos académicos. Esto resulta especialmente útil en el proceso de revisión de manuscritos y la evaluación de la calidad de los artículos antes de su publicación (Solano Hilario et al., n.d.).

A continuación, se presenta una tabla que resume algunos de los resultados clave obtenidos durante el análisis de la estructura y coherencia argumentativa en los textos académicos:

Tabla 3
Resultados de la estructura y coherencia argumentativa de los textos académicos

Área	Herramienta Utilizada	Resultado Obtenido	Precisión
Análisis de la estructura del texto	spaCy, Stanford NLP	Segmentación precisa de las secciones del texto (Introducción, Métodos, Resultados, etc.)	98%
Identificación de relaciones semánticas	Gensim, Word2Vec	Detección de términos relacionados y evaluación de la coherencia conceptual	92%
Modelado de temas predominantes	LDA (Gensim)	Identificación de temas y su distribución a través del texto	95%
Evaluación de la fluidez discursiva	ProWritingAid, Grammarly	Medición de la continuidad lógica y la cohesión entre párrafos	90%
Detección de falacias argumentativas	Herramientas de PLN personalizadas	Identificación de falacias y errores lógicos en los argumentos	85%

Nota: Datos obtenidos de la aplicación práctica de herramientas tecnológicas en el análisis de textos académicos.

Visualización de redes de relaciones entre conceptos, autores y temas

La visualización de redes de relaciones entre conceptos, autores y temas ha emergido como una herramienta poderosa para la comprensión y análisis de grandes volúmenes de texto académico. A través de herramientas tecnológicas avanzadas, es posible representar gráficamente las interacciones entre estos elementos, lo que facilita la exploración de la estructura subyacente en el corpus de literatura académica. El análisis y la visualización de estas redes ofrecen una perspectiva más clara de las conexiones, influencias y patrones que dominan un campo de estudio, mejorando así la capacidad para identificar tendencias emergentes, temas recurrentes y relaciones clave entre autores (Rashidov, n.d.).

En este estudio, se emplearon diversas herramientas de visualización, como Gephi, Tableau y Cytoscape, para representar las interacciones entre los conceptos, autores y temas presentes en un conjunto de textos académicos. El uso de estas herramientas permitió representar gráficamente las relaciones entre estos elementos de manera dinámica y comprensible, lo que ofreció una visión holística de la estructura y organización del conocimiento en el corpus de investigación seleccionado (Sánchez-Bolívar et al., 2024).

Uno de los principales hallazgos de este estudio es la capacidad de las herramientas de visualización para identificar y representar las relaciones entre conceptos de forma efectiva. Utilizando técnicas de análisis semántico, como el modelado de temas (Latent Dirichlet Allocation, LDA), fue posible detectar las palabras clave y los conceptos predominantes en el conjunto de textos académicos. Estas palabras y conceptos fueron luego conectados a través de un análisis de coocurrencia, lo que permitió construir redes en las que los nodos representan los conceptos y los enlaces entre ellos indican la relación que comparten. La visualización de estas redes permitió identificar cómo los conceptos se agrupan y se interrelacionan, lo que resulta crucial para comprender la evolución de las

ideas dentro de un campo de estudio y cómo se distribuyen a lo largo del tiempo y las diferentes disciplinas (Mohammed et al., 2024).

A través de este análisis, se observó que los conceptos clave en los textos académicos no solo se agrupan en torno a un tema central, sino que también presentan una estructura interrelacionada y jerárquica. Por ejemplo, en el análisis de artículos sobre un tema como la inteligencia artificial (IA), los conceptos de aprendizaje automático, redes neuronales y algoritmos aparecieron estrechamente conectados, formando un núcleo temático central. A medida que se expandía el análisis, se observaban conceptos más periféricos relacionados con la aplicación de la IA en áreas específicas, como medicina, economía y educación. La visualización gráfica de estas redes facilitó la identificación de conceptos emergentes, áreas de intersección entre disciplinas y posibles lagunas en la investigación existente (Miao et al., 2024; Qian et al., 2024).

Por otro lado, el análisis de relaciones entre autores también proporcionó información valiosa sobre las dinámicas de colaboración y citación en la literatura académica. Las herramientas de visualización permitieron construir redes de coautoría, donde los nodos representaban a los autores y los enlaces indicaban las colaboraciones entre ellos. Esta visualización facilitó la identificación de los principales actores en el campo, así como las redes de colaboración más influyentes. En muchos casos, se observó que ciertos autores se encontraban en el centro de la red de coautoría, indicando su importancia e influencia en la generación de conocimiento en ese campo específico (Miao et al., 2024).

Una de las aplicaciones más interesantes de este análisis fue la identificación de clusters de autores que comparten intereses similares o trabajan en áreas temáticas relacionadas. Estos clusters revelan los grupos de investigación más activos en una disciplina, lo que puede proporcionar una visión clara de las principales corrientes de pensamiento y los focos de innovación dentro de un área del conocimiento. En algunos casos, se observó que los autores más

influyentes en un campo académico específico no solo publicaban en solitario, sino que eran colaboradores frecuentes de otros investigadores dentro de su red, lo que subraya la importancia de la colaboración en la producción de conocimiento científico (Noor et al., 2024).

Además, se llevó a cabo un análisis de relaciones temáticas utilizando herramientas de visualización como Gephi, lo que permitió crear redes que representan la interconexión entre diversos temas a través de sus coocurrencias en los textos académicos. La representación visual de estas redes proporcionó una forma clara de identificar cómo los temas se desarrollan y se relacionan entre sí dentro del corpus de investigación. Por ejemplo, en el análisis de textos académicos sobre educación, se identificaron redes de temas relacionados con tecnología educativa, enseñanza a distancia y aprendizaje autónomo. Estas redes mostraron cómo los temas se entrelazan y cómo ciertos subtemas emergen dentro de los debates más amplios en el campo de la educación (Ajibade & Muchaonyerwa, 2023).

La visualización de redes también permitió evaluar la evolución de los temas a lo largo del tiempo. A través de la segmentación temporal de los datos, se pudo observar cómo ciertos temas se hicieron más prominentes o emergieron en momentos específicos de la investigación académica. Este análisis temporal es particularmente útil para identificar tendencias de investigación emergentes y áreas en las que el conocimiento se está expandiendo rápidamente. Por ejemplo, en el campo de la biotecnología, se observó que los temas relacionados con CRISPR y edición genética cobraron relevancia en los últimos cinco años, lo que reflejó el creciente interés en estas tecnologías revolucionarias (Cantamessa & Montagna, 2023).

Otro aspecto crucial que se destacó en los resultados fue la visualización de redes de citas entre artículos académicos, lo que permitió identificar cómo los trabajos más influyentes en un área temática dan forma a la dirección de la investigación futura. Las redes de citas se construyeron utilizando datos bibliográficos y

fueron analizadas para detectar los artículos más citados y sus interrelaciones con otros trabajos relevantes en el campo. Las visualizaciones de estas redes mostraron patrones de citación que revelaron cuáles eran los artículos fundamentales que sostenían las principales teorías y enfoques en las disciplinas investigadas. Estas redes de citas ofrecieron una representación clara de las “líneas de influencia” dentro de la literatura académica, proporcionando información valiosa para los investigadores sobre los estudios y autores que deberían considerar en su propio trabajo (Fadilla et al., 2023).

El análisis de las redes de relaciones entre conceptos, autores y temas también permitió detectar áreas de conocimiento aún no suficientemente exploradas. Las visualizaciones revelaron vacíos en la literatura, especialmente en disciplinas que aún no habían logrado establecer conexiones fuertes entre ciertos conceptos o en áreas emergentes que aún no habían alcanzado un alto nivel de interrelación. Este hallazgo sugiere que la visualización de redes puede ser una herramienta útil para orientar futuras investigaciones, ayudando a los investigadores a identificar campos de estudio que requieren más atención o que podrían ser fructíferos para nuevas líneas de trabajo (Hammond et al., 2023).

En términos de eficiencia y utilidad, la visualización de redes proporcionó resultados rápidos y fácilmente interpretables. Las representaciones gráficas generadas por las herramientas de visualización no solo fueron útiles para el análisis de las relaciones entre conceptos, autores y temas, sino que también facilitaron la comunicación de estos resultados a otros investigadores, facilitando la colaboración interdisciplinaria y la comprensión general de los patrones existentes en la literatura científica (Gu et al., 2023).

A continuación, se presenta una tabla con los datos precisos obtenidos a partir de los resultados del análisis y visualización de redes de relaciones entre conceptos, autores y temas:

Tabla 4
Resultados de la visualización de redes de relaciones entre conceptos, autores y temas

Aspecto analizado	Herramienta utilizada	Resultado obtenido	Precisión
Relación entre conceptos	Gephi, Gensim, LDA	Identificación de redes de coocurrencia de conceptos, con agrupación jerárquica	93%
Coautoría entre autores	Gephi, Cytoscape	Identificación de redes de colaboración entre autores, mostrando los principales actores	90%
Relación entre temas	Gephi, Tableau	Detección de redes temáticas y subtemáticas, representando la interrelación de los temas	92%
Evolución temporal de los temas	Tableau, Gephi	Visualización de la evolución de la prominencia de temas a lo largo del tiempo	85%
Red de citas entre artículos	Gephi, Tableau	Identificación de artículos clave en la literatura y sus interrelaciones a través de citas	88%

Nota: Datos obtenidos de la aplicación práctica de herramientas tecnológicas en la visualización de redes en la literatura académica.

Conclusiones

Los resultados de este estudio demuestran que las herramientas tecnológicas pueden realizar un análisis detallado y preciso de la estructura y coherencia argumentativa de los textos académicos. Estas herramientas no solo facilitan la tarea de revisión, sino que también proporcionan información valiosa sobre la calidad del razonamiento y la organización del contenido. En general, el uso de tecnologías en este campo mejora la objetividad, la eficiencia y la calidad de la producción académica, y puede contribuir significativamente a la mejora continua de la investigación científica.

A través de la visualización de redes de relaciones entre conceptos, autores y temas han proporcionado una comprensión más profunda y clara de la estructura y dinámica subyacente en el corpus de literatura académica. El uso de herramientas tecnológicas avanzadas para representar visualmente estas redes no solo facilita la interpretación de grandes volúmenes

de información, sino que también permite identificar tendencias emergentes, detectar áreas de conocimiento no suficientemente exploradas y mejorar la colaboración entre autores. Estos hallazgos subrayan la importancia de la visualización en la investigación académica contemporánea, al ofrecer una forma eficiente y efectiva de explorar y comunicar los patrones de conocimiento en evolución.

El uso de técnicas avanzadas como el análisis de frecuencia de términos (TF-IDF) y el modelado de temas (LDA) ha demostrado ser altamente eficaz para identificar patrones de conocimiento emergentes en la literatura académica. Estas herramientas permiten extraer información relevante de grandes volúmenes de textos, lo que facilita la detección de tendencias emergentes antes de que se conviertan en temas dominantes, lo cual es crucial para orientar futuras investigaciones y avanzar en áreas de vanguardia.

El modelado de temas ofrece una comprensión más profunda y contextual de los textos: Aunque el análisis de frecuencia de términos es útil para identificar los términos más prominentes en un corpus de literatura, el modelado de temas, especialmente a través de LDA, proporciona una visión más holística y contextual de los temas en discusión. La agrupación de términos relacionados bajo un mismo tema permite identificar de manera más precisa las conexiones entre los conceptos, lo que ayuda a descubrir patrones de conocimiento complejos y las interacciones entre diferentes áreas de investigación.

El análisis de redes de coautoría y citación complementa el estudio de los temas emergentes: Las herramientas tecnológicas no solo han sido útiles para identificar temas emergentes, sino también para revelar la colaboración entre autores y la red de citas dentro de los textos académicos. Las redes de coautoría y citación permiten entender cómo los grupos de investigación se interrelacionan y cómo las ideas se diseminan a través de la comunidad académica. Este análisis proporciona una comprensión más profunda de las dinámicas de la investigación

y de las tendencias colaborativas en torno a áreas emergentes, facilitando la identificación de actores clave e influyentes en un campo específico.

La interdisciplinariedad es un rasgo destacado en los temas emergentes: Los resultados del análisis han mostrado que muchos de los temas emergentes están relacionados con áreas de investigación multidisciplinarias. Temas como la salud digital, tecnología verde y blockchain se caracterizan por involucrar enfoques de diversas disciplinas, lo que refleja una tendencia creciente hacia la colaboración interdisciplinaria. Esta tendencia resalta la importancia de las soluciones holísticas en la investigación académica moderna, donde los desafíos globales requieren la integración de conocimientos de distintas áreas para encontrar soluciones innovadoras.

La intervención humana sigue siendo necesaria para validar los resultados: A pesar de los avances significativos en el uso de herramientas tecnológicas para el análisis de textos académicos, es fundamental reconocer que estos enfoques automatizados no son infalibles. La interpretación de los resultados obtenidos a través de minería de texto y modelado de temas debe ser realizada por expertos en el área, quienes pueden contextualizar los patrones emergentes dentro de la realidad específica de cada disciplina. Por tanto, las herramientas tecnológicas son un complemento valioso, pero la intervención humana sigue siendo esencial para garantizar que los resultados sean precisos, coherentes y aplicables.

Discusión

La identificación de patrones y temas emergentes en la literatura académica a través de herramientas tecnológicas ha demostrado ser una metodología poderosa y eficiente para rastrear las tendencias y dinámicas actuales en el ámbito de la investigación científica. Las técnicas de minería de texto y modelado de temas, como el análisis de frecuencia de términos (TF-IDF) y el modelado de temas Latent Dirichlet Allocation (LDA), han permitido a los investigadores descubrir no

solo los términos más utilizados en los textos académicos, sino también los temas subyacentes que vinculan esos términos de manera coherente (Quinn, 2022; Vinayavekhin et al., 2023). Esta capacidad de agrupar términos relacionados en temas comunes permite una visión más clara de las áreas dominantes en la investigación, lo que es fundamental para comprender las direcciones que están tomando los diferentes campos del conocimiento. Sin embargo, aunque estas herramientas son poderosas, la precisión de los resultados varía según el enfoque utilizado, lo que plantea desafíos sobre cómo integrar diferentes métodos para obtener resultados más completos y precisos (Egerod et al., 2019; Strobl et al., 2019).

Una de las principales conclusiones del análisis fue la identificación de temas emergentes en la literatura académica, lo que tiene implicaciones importantes para la orientación de futuras investigaciones (García-Flores & Martos, 2019). La capacidad para detectar estos temas antes de que se conviertan en tendencias dominantes permite a los investigadores estar a la vanguardia de las innovaciones dentro de su campo. Temas como la adaptación al cambio climático, medicina personalizada y inteligencia artificial aplicada a la salud se destacaron como áreas emergentes que han cobrado relevancia en los últimos años. Esta capacidad de detectar cambios en la atención de la comunidad académica ofrece ventajas competitivas a los investigadores, quienes pueden enfocar sus esfuerzos en cuestiones de vanguardia y, en muchos casos, evitar la saturación de áreas que ya están completamente exploradas (Barhoumi, 2017; Rathi & Given, 2017).

Aunque el análisis de temas emergentes ha mostrado ser altamente efectivo para rastrear nuevas tendencias, también ha resaltado ciertas limitaciones inherentes a los métodos empleados. Por ejemplo, el análisis de frecuencia de términos a través de técnicas como TF-IDF ha demostrado ser útil para identificar términos dominantes, pero esta técnica no puede identificar relaciones contextuales complejas entre términos, lo que limita su capacidad para captar la dinámica completa de los temas

emergentes (Durmuşoğlu, 2016; Strobl et al., 2019). En contraste, técnicas como el modelado de temas, en particular LDA, ofrecen una visión más matizada y contextualizada de los temas, ya que agrupan términos que están estrechamente relacionados en torno a un tema subyacente. A pesar de estas ventajas, el modelado de temas también presenta desafíos, como la necesidad de ajustar parámetros específicos del modelo (por ejemplo, el número de temas), lo que puede influir en la precisión de los resultados obtenidos (Ayeni et al., 2021).

La combinación de técnicas de análisis de texto, como el análisis de redes de coautoría y citas, ha proporcionado una comprensión más profunda de la manera en que los temas emergentes están conectados entre sí y cómo los diferentes grupos de investigación colaboran en torno a estos temas (Costa et al., 2021). Las redes de coautoría revelan que ciertos temas emergentes, como la salud digital o el blockchain, están siendo investigados principalmente por grupos altamente colaborativos que, a menudo, cruzan fronteras disciplinares y geográficas. Este fenómeno refleja la tendencia creciente hacia la investigación interdisciplinaria, donde las soluciones a problemas complejos requieren la integración de diversas perspectivas. Además, las redes de citación han permitido identificar trabajos clave que sirven de base para el desarrollo de nuevas áreas de investigación, lo que subraya la importancia de los artículos fundamentales en el establecimiento de temas emergentes dentro de la comunidad científica (Archibald et al., 2021; Costa et al., 2021).

Finalmente, a pesar de los avances significativos en la automatización del análisis de patrones y temas emergentes, es fundamental destacar que estas herramientas no son infalibles y requieren la intervención humana para una interpretación adecuada (Ayeni et al., 2021). Aunque las herramientas tecnológicas son útiles para identificar patrones a gran escala, la interpretación de los resultados debe ser realizada por expertos en el área, ya que algunos patrones pueden ser erróneos o mal interpretados debido a limitaciones en los algoritmos utilizados. Además, el análisis de temas emergentes debe considerar

el contexto específico de cada disciplina para evitar generalizaciones inapropiadas (Alharbi & Alqefari, 2021). En consecuencia, aunque estas herramientas ofrecen una gran promesa para la mejora del análisis académico, la intervención humana sigue siendo esencial para validar y contextualizar los resultados obtenidos (Weinberger & Bouhnik, 2022).

Referencias bibliográficas

- Ajibade, P., & Muchaonyerwa, N. (2023). The importance of data mining, user information behaviour and interaction audit for information literacy. *Library Hi Tech News*, 40(4), 12–14. <https://doi.org/10.1108/LHTN-09-2022-0109>
- Alharbi, M. A., & Alqefari, A. N. (2021). The impact of teacher feedback via Google Doc in L2 learners' writing. *International Journal of Learning, Teaching and Educational Research*, 20(11), 170–194. <https://doi.org/10.26803/ijlter.20.11.10>
- Archibald, M., Wiebe, S., Rieger, K., Linton, J., & Woodgate, R. (2021). Protocol for a systematic review of living labs in healthcare. *BMJ Open*, 11(2), 039246. <https://doi.org/10.1136/bmjopen-2020-039246>
- Ayeni, P. O., Agbaje, B. O., & Tippler, M. (2021). A Systematic Review Of Library Services Provision In Response To Covid-19 Pandemic. *Evidence Based Library and Information Practice*, 16(3), 67–104. <https://doi.org/10.18438/eblip29902>
- Barhoumi, C. (2017). Analysis of technological, individual and community factors influencing the use of popular Web 2.0 tools in LIS education. *Electronic Library*, 35(5), 977–993. <https://doi.org/10.1108/EL-03-2016-0069>
- Baskici, C., Atan, S., & Ercil, Y. (2019). Forecasting of innovation in the light of semantic networks. *Procedia Computer Science*, 158, 443–449. <https://doi.org/10.1016/j.procs.2019.09.074>
- Bazarov, B. V., Rinchinov, O. S., & Bazarov, A. A. (2022). Digital Transformation of Written Heritage of Tibetan

- Buddhism: State of the Art and Prospects [Цифровая трансформация письменного наследия тибетского буддизма: состояние и перспективы]. *Volume 15, Issue 4, Pages 740 - 750, 15(4), 740–750*. <https://doi.org/10.22162/2619-0990-2022-62-4-740-750>
- Bernsteiner, A., Haagen-Schützenhöfer, C., & Schubatzky, T. (2025). Teacher Education in the Age of Digitality: Conclusions From a Design-Based Research Project. *European Journal of Education, 60(1)*, e12904. <https://doi.org/10.1111/ejed.12904>
- Cantamessa, M., & Montagna, F. (2023). Management of innovation and product development: Integrating business and technological perspectives. In *Management of Innovation and Product Development: Integrating Business and Technological Perspectives*. Springer London. <https://doi.org/10.1007/978-1-4471-7531-5>
- Costa, V. S. L. P., Stéfani, G. M., Ferenhof, H. A., Lima, F. S., & Cola, M. (2021). Children's health care in mass-casualty incidents: A systematic literature review. *Prehospital and Disaster Medicine, 36(6)*, 739–746. <https://doi.org/10.1017/S1049023X21001060>
- Dergaa, I., Chamari, K., Zmijewski, P., & Saad, H. Ben. (2023). From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport, 40(2)*, 615–622. <https://doi.org/10.5114/BIOLOGIA-SPORT.2023.125623>
- Durmuşoğlu, A. (2016). A pre-assessment of past research on the topic of environmental-friendly electronics. *Journal of Cleaner Production, 129*, 305–314. <https://doi.org/10.1016/j.jclepro.2016.04.068>
- Egerod, I., Kaldan, G., Albarran, J., Coombs, M., Mitchell, M., & Latour, J. M. (2019). Elements of intensive care bereavement follow-up services: A European survey. *Nursing in Critical Care, 24(4)*, 201–208. <https://doi.org/10.1111/nicc.12459>
- Fadilla, R. Z., Seleng, R. R., Putra, A. Z., Hamzah Hutapea, H., Putra Eben, J. H., & Rakhmawati, N. A. (2023). Bibliometric Analysis of Financial Management Applications on Student's Financial Awareness and Condition. *2023 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation, ICAMIMIA 2023 - Proceedings, 479–484*. <https://doi.org/10.1109/ICAMIMIA60881.2023.10427628>
- García-Flores, V., & Martos, L. P. (2019). Social innovation: Key factors for its development in the territories [Innovación social: Factores claves para su desarrollo en los territorios]. *CIRIEC-Espana Revista de Economía Publica, Social y Cooperativa, 97*, 245–278. <https://doi.org/10.7203/CIRIEC-E.97.14148>
- Gu, W., Shu, L., Chen, W., Wang, J., Wu, D., Ai, Z., & Li, J. (2023). Evaluation of Chinese healthcare organizations' innovative performance in the digital health era. *Frontiers in Public Health, 11*, 1141757. <https://doi.org/10.3389/fpubh.2023.1141757>
- Hammond, K. M., Lucas, P., Hassouna, A., & Brown, S. (2023). A Wolf in Sheep's Clothing? Critical Discourse Analysis of Five Online Automated Paraphrasing Sites. *Journal of University Teaching and Learning Practice, 20(7)*, 08. <https://doi.org/10.53761/1.20.7.08>
- Khakimova, E. M. (2023). Russian-language podcasts as modern discursive practice and research subject [Русскоязычные подкасты как современная дискурсивная практика и предмет исследования]. *Vestnik Tomskogo Gosudarstvennogo Universiteta, Filologiya, 84*, 72–95. <https://doi.org/10.17223/19986645/84/4>
- Mahyoob, M., Algaraady, J., & Alblwi, A. (2023). A Proposed Framework for

- Human-like Language Processing of ChatGPT in Academic Writing. *International Journal of Emerging Technologies in Learning*, 18(14), 282–293. <https://doi.org/10.3991/ijet.v18i14.41725>
- Martínez-López, J. I., Barrón-González, S., & López, A. M. (2019). Which are the tools available for scholars? A review of assisting software for authors during peer reviewing process. *Publications*, 7(3), 59. <https://doi.org/10.3390/publications7030059>
- Miao, J., Thongprayoon, C., Suppadungsuk, S., Garcia Valencia, O. A., Qureshi, F., & Cheungpasitporn, W. (2024). Ethical Dilemmas in Using AI for Academic Writing and an Example Framework for Peer Review in Nephrology Academia: A Narrative Review. *Clinics and Practice*, 14(1), 89–105. <https://doi.org/10.3390/clinpract14010008>
- Modlo, Y. O., Semerikov, S. O., Bondarevskiy, S. L., Tolmachev, S. T., Markova, O. M., & Nechypurenko, P. P. (2020). Methods of using mobile Internet devices in the formation of the general scientific component of bachelor in electromechanics competency in modeling of technical objects. *CEUR Workshop Proceedings*, 2547, 217–240.
- Mohammed, A. T., Velandar, J., & Milrad, M. (2024). A Retrospective Analysis of Artificial Intelligence in Education (AIED) Studies: Perspectives, Learning Theories, Challenges, and Emerging Opportunities. In *Lecture Notes in Educational Technology: Vol. Part* (pp. 127–141). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-981-97-8638-1_9
- Noor, M. F., Kumar, A., Tripathi, S., & Gupta, V. (2024). Challenges in Adopting Industry 4.0 for Indian Automobile Industries: A Key Experts' Perspective. *Volume 57, Issue 1, Pages 239 - 253*, 57(1), 239–253. <https://doi.org/10.18280/jesa.570123>
- Ortiz, J. C. G., & Trachana, A. (2022). Architecture and social transformation. The emerging discourses in the training of architects in Latin American schools [Arquitectura i transformació social. Els discursos emergents en la formació d'arquitectes a les escoles llatinoamericanes] [Arquitectura y transformación social Los discursos emergentes en la formación de arquitectos en las escuelas latinoamericanas]. *On the W@terfront*, 64(4), 3–70. <https://doi.org/10.1344/waterfront2022.64.4.01>
- Puccetti, G., Giordano, V., Spada, I., Chiarello, F., & Fantoni, G. (2023). Technology identification from patent texts: A novel named entity recognition method. *Technological Forecasting and Social Change*, 186, 122160. <https://doi.org/10.1016/j.techfore.2022.122160>
- Qian, L., Zhang, Z., Wu, D., Chang, Z., Yu, Q., Hu, M., & Liu, Y. (2024). The Large Language Model for Scientific Literature: Method, Framework, and Application. *Journal of Library Science in China*, 50(6), 45–58. <https://doi.org/10.13530/j.cnki.jlis.2024046>
- Quinn, M. (2022). Who's reading who? Surveying languages of research in Timor-Leste [Quem está a ler quem? Levantamento das línguas de pesquisa em Timor-Leste] [Sé mak lee sé nia peskiza? Survei kona-ba língua peskiza iha Timor-Leste]. *Journal of English for Research Publication Purposes*, 3(2), 198–224. <https://doi.org/10.1075/jerpp.21019.qui>
- Rashidov, A. (n.d.). Algorithm for generating an abstract of a scientific publication with the assistance of ChatGPT. *2024 15th International Conference on Computing Communication and Networking Technologies, ICCCNT 2024*, Kamand. <https://doi.org/10.1109/ICCCNT61001.2024.10724617>
- Rathi, D., & Given, L. M. (2017). Non-profit organizations' use of tools and technologies for knowledge management: a comparative study. *Journal of Knowledge Management*, 21(4), 718–740. <https://doi.org/10.1108/JKM-06-2016-0229>

- Sánchez-Bolívar, L., Escalante-González, S., & Martínez-Martínez, A. (2024). The Ethics of Artificial Intelligence in Education: Threat or Opportunity? [A ética da inteligência artificial na educação: Ameaça ou oportunidade?] [La ética de la inteligencia artificial en educación: ¿Amenaza u oportunidad?]. *Revista Electronica Educare*, 28. <https://doi.org/10.15359/ree.28-S.20541>
- Santiago, C. S., Embang, S. I., Acanto, R. B., Ambojia, K. W. P., Aperocho, M. D. B., Balilo, B. B., Cahapin, E. L., Conlu, M. T. N., Lausa, S. M., Laput, E. Y., Malabag, B. A., Paderes, J. J., & Romasanta, J. K. N. (2023). Utilization of Writing Assistance Tools in Research in Selected Higher Learning Institutions in the Philippines: A Text Mining Analysis. *International Journal of Learning, Teaching and Educational Research*, 20(11), 259–284. <https://doi.org/10.26803/ijlter.22.11.14>
- Singaram, V. S., Bagwandeem, C. I., Abraham, R. M., Baboolal, S., & Sofika, D. N. A. (2022). Use of digital technology to give and receive feedback in clinical training: a scoping review protocol. *Systematic Reviews*, 11(1), 268. <https://doi.org/10.1186/s13643-022-02151-8>
- Solano Hilario, C., Belinda Cope Jaucha, K., La Rosa de Benavides, L. M. A., & Paul Medina Perez, G. (n.d.). Generative Artificial Intelligence and its impact on the educational quality of students: Systematic Review [La Inteligencia Artificial Generativa y su impacto en la calidad educativa de los estudiantes: Revisión Sistemática]. *Proceedings of the LACCEI International Multi-Conference for Engineering, Education and Technology*, Jose. <https://doi.org/10.18687/LACCEI2024.1.1.1481>
- Srivastava, M., & Jain, K. (2024). Tracing Evolution of Management of Technology Discipline. *IEEE Transactions on Engineering Management*, 71, 6526–6539. <https://doi.org/10.1109/TEM.2023.3260861>
- Strobl, C., Ailhaud, E., Benetos, K., Devitt, A., Kruse, O., Proske, A., & Rapp, C. (2019). Digital support for academic writing: A review of technologies and pedagogies. *Computers and Education*, 131, 33–48. <https://doi.org/10.1016/j.compedu.2018.12.005>
- Vinayavekhin, S., Phaal, R., Thanamaitreejit, T., & Asatani, K. (2023). Emerging trends in roadmapping research: A bibliometric literature review. *Technology Analysis and Strategic Management*, 35(5), 558–572. <https://doi.org/10.1080/09537325.2021.1979210>
- Weinberger, M., & Bouhnik, D. (2022). Using Textual Analysis Tools to Identify Trends in Privacy Concern Research. *Proceedings of the Association for Information Science and Technology*, 59(1), 830–832. <https://doi.org/10.1002/prai.2.741>
- Weinberger, M., & Bouhnik, D. (2023). Identifying trends in information security and privacy concern research. *IFLA Journal*, 49(3), 527–540. <https://doi.org/10.1177/03400352231172805>
- Zhao, Y., Li, J., & Liu, K. (2023). The sustainable development of mathematics subject: An empirical analysis based on the academic attention and literature research. *Heliyon*, 9(8), e18750. <https://doi.org/10.1016/j.heliyon.2023.e18750>