



Big data en *marketing*: tendencias y desafíos desde un enfoque bibliométrico

Big data in marketing: trends and challenges from a bibliometric approach

DOI: 10.22458/rna.v16iE2.6014

Edgar Alfonso Sansores-Guerrero¹

Juana Edith Navarrete-Marneou²

1,2. Universidad Autónoma Metropolitana, Ciudad de México, México
easg@azc.uam.mx; jenm@azc.uam.mx

Recepción: 07 de agosto de 2025

Corrección: 10 de noviembre de 2025

Aceptación: 19 de noviembre de 2025

RESUMEN

Este artículo es parte de un proyecto sobre el uso de *big data* (BDA) en la gestión. El objetivo fue evaluar el impacto de la aplicación de BDA en el *marketing* desde un enfoque bibliométrico; para ello, se empleó un análisis basado en Biblioshiny y VOSviewer. Este enfoque representa una alternativa para el estudio de la literatura especializada en el tema. El trabajo se posiciona entre las primeras investigaciones latinoamericanas que examinan tendencias, desafíos y brechas en la aplicación del BDA en el *marketing*. Los hallazgos revelaron que el estudio en este campo aún se encuentra en una fase inicial.

Palabras clave: Base de datos, mercadotecnia, bibliometría.

ABSTRACT

This article is part of a project focusing on the use of big data (BDA) in management. The objective was to evaluate the impact of applying BDA to marketing using a bibliometric approach. To this end, an analysis based on Biblioshiny and VOSviewer was conducted. This approach is an alternative way to study specialized literature on the subject. This study is among the first Latin American research to examine trends, challenges, and gaps in the application of BDA in marketing. The findings revealed that research in this field is still in its early stages.

Keywords: Database, marketing, bibliometrics.

RÉSUMÉ

Cet article fait partie d'un projet sur l'utilisation de Big Data (BDA) dans la gestion. L'objectif était d'évaluer l'impact de l'application du BDA sur le marketing à partir d'une approche bibliométrique. Pour cela, une analyse basée sur Biblioshiny et VOSviewer a été utilisée. Cette approche représente une alternative pour l'étude de la littérature spécialisée sur le sujet. Ce travail fait partie des premières recherches latino-américaines qui examinent les tendances, les défis et les lacunes dans l'application du BDA au marketing. Les résultats ont révélé que l'étude dans ce domaine est encore à un stade précoce.

Mots-clés: Base de données, marketing, bibliométrie.

Nota: Este artículo forma parte del número especial publicado en el marco del V Congreso Internacional de Marketing, celebrado los días 16 y 17 de junio de 2025 en la Universidad Estatal a Distancia (UNED), Costa Rica.

1. Doctor en Estudios Organizacionales, Universidad Autónoma Metropolitana, Profesor-Investigador de la Universidad Autónoma del Estado de Quintana Roo.

2. Doctora en Dirección y Finanzas, Universidad Popular Autónoma del Estado de Puebla, Profesora-Investigadora de la Universidad Autónoma del Estado de Quintana Roo.

RESUMO

Este artículo faz parte de um projeto sobre o uso de Big Data (BDA) na gestão. O objetivo foi avaliar o impacto da aplicação do BDA no marketing a partir de uma abordagem bibliométrica. Para isso, foi utilizada uma análise baseada no Biblioshiny e no VOSviewer. Essa abordagem representa uma alternativa para o estudo da literatura especializada no tema. Este trabalho está entre as primeiras pesquisas latino-americanas que examinam tendências, desafios e lacunas na aplicação do BDA no marketing. Os resultados revelaram que o estudo neste campo ainda se encontra em uma fase inicial.

Palavras-chave: Banco de dados, marketing, bibliometria.

INTRODUCCIÓN

En el contexto de la cuarta revolución industrial, el término *big data* (BDA) hace referencia a conjuntos de datos de una magnitud y complejidad que no pueden ser procesados mediante herramientas o sistemas de gestión de información convencionales (Lu *et al.*, 2022)³. Estos datos pueden alcanzar volúmenes en el rango de exabytes, lo cual hace que su recopilación, almacenamiento, búsqueda, intercambio, transferencia, análisis, interpretación y visualización requieran de un alto nivel de recursos y una gran interacción entre humanos y máquinas (Zhang *et al.*, 2021)⁴.

En ese sentido, dichos macrodatos poseen cuatro dimensiones: a) volumen: representan una enorme cantidad de información (hasta diez zettabytes) que las organizaciones compilan mediante fuentes tradicionales; b) variedad: se relaciona con los diversos tipos de datos, tanto estructurados como no estructurados, que se obtienen de múltiples fuentes; c) veracidad: se remite a la calidad y confiabilidad de los datos; y d) velocidad: se refiere a la rapidez con la que los datos se generan y procesan. Esta velocidad es asombrosa y requiere que los datos se analicen casi en tiempo real para ser útiles.

Aunque existen distintos enfoques para su aplicación, se observan elementos comunes en los métodos y tecnologías que se emplean con frecuencia. En ese sentido, para procesar esta gran cantidad de información, los sistemas informáticos, generalmente, dividen esta actividad en cuatro pasos: entrada de datos, almacenamiento, procesamiento y análisis, e interpretación de los resultados.

Bajo esta visión sistémica, se puede evidenciar que el valor del BDA no reside en el volumen de datos disponibles, sino en el análisis estratégico que se realiza de los resultados. Por ejemplo, para las tareas de *marketing*, el ingreso de datos se refiere a la recolección de información estructurada (como órdenes de compra) y no estructurada (comentarios de los clientes en redes sociodigitales), mientras que el almacenamiento hace referencia al uso de la infraestructura física y virtual para el acceso a los datos en tiempo real.

En tanto, el análisis de esta información puede implicar el uso de una estadística descriptiva, la aplicación de algoritmos avanzados de inteligencia artificial, redes neuronales artificiales, técnicas para la interpretación del discurso y métricas para estimar las emociones de los consumidores. Finalmente, la presentación gráfica de los resultados facilita la construcción de mapas, figuras, ilustraciones y esquemas para su interpretación.

Ante esto, la gestión y el análisis de grandes volúmenes de información representan un área de oportunidad para las empresas que se encuentran en la constante búsqueda de desarrollar ventajas competitivas (Ningsih *et al.*, 2025)⁵. Sin embargo, la existencia de barreras de entrada (como la inversión y la especialización) ha limitado la adopción del BDA en áreas de comercialización y ventas.

Por esta situación, en el desarrollo de la literatura especializada en el tema, se observa un predominio del enfoque micro que se ha centralizado en el análisis y desarrollo de plataformas digitales (Gupta *et al.*, 2021)⁶. No obstante, las aportaciones que ha realizado esta perspectiva al desarrollo de un cuerpo teórico robusto siguen siendo incipientes, por lo que aún persisten brechas significativas en la literatura.

3 Lu, Yusheng y Jiantong Zhang, "Bibliometric analysis and critical review of the research on big data in the construction industry", *Engineering, Construction and Architectural Management* 29, no. 9 (2022): 3574-3592.

4 Zhang, Yucheng, Meng Zhang, Jing Li, Guangjian Liu, Miles M. Yang y Siqi Liu, "A bibliometric review of a decade of research: Big data in business research—Setting a research agenda", *Journal of Business Research* 131 (2021): 374-390.

5 Ningsih, Fitriah, Muhammad Ilham y Murtiadi Awaluddin, "Industrial Competitiveness through Porter's Five Forces Framework", *Al-Kharaj: Journal of Islamic Economic and Business* 7, no. 2 (2025).

6 Gupta, Shivam, Théo Justy, Shampy Kamboj, Ajay Kumar y Eivind Kristoffersen, "Big data and firm marketing performance: Findings from knowledge-based view", *Technological Forecasting and Social Change* 171 (2021): 120986.

Dentro de las principales brechas identificadas destacan la integración del BDA en el análisis de *marketing*, la validación de marcos analíticos, el estudio de los efectos a largo plazo de la personalización basada en BDA, la falta de métricas para clasificar la lealtad del consumidor, y la ausencia de un marco teórico robusto (Tang *et al.*, 2022)⁷. Además, resulta necesario desarrollar nuevos enfoques y marcos analíticos para tratar temas como la transformación digital impulsada por el BDA, la gestión de la privacidad de los datos, el desarrollo de competencias profesionales y la evaluación de diversas herramientas digitales (Fathi *et al.*, 2022)⁸.

Por lo anterior, el propósito de la presente investigación es examinar el impacto de la BDA en el *marketing* a través de un análisis bibliométrico (técnicas de minería de texto y modelado de temas). Si bien existen diversos estudios de este tipo, estos se orientan a medir la cantidad de publicaciones que abordan las diferentes aplicaciones del DBA en las actividades del *marketing*; por lo que, a diferencia de los análisis bibliométricos previos, este estudio utiliza bases de datos de revistas científicas indexadas en el ámbito del impacto del BDA en la gestión comercial y emplea el término “*Big Data Marketing*” como uno de los criterios de búsqueda temática. Bajo este enfoque, se examinan los metadatos bibliográficos relevantes para la construcción e interpretación del conocimiento científico sobre el tema, se determina el estado actual de la literatura y se establece una agenda de futuras líneas de investigación.

Cabe destacar, que el análisis bibliométrico permite identificar patrones, tendencias e impactos dentro de un campo específico del conocimiento. Con ello, se pueden trazar rutas para el desarrollo de nuevas perspectivas, diseñar innovadoras metodologías que faciliten una mejor comprensión del fenómeno y la aplicación de marcos analíticos transdisciplinarios.

Dada sus bondades, esta metodología ha sido la más utilizada en la investigación académica, al consolidarse como una herramienta eficiente y eficaz en el análisis de grandes volúmenes de información científica. En dichas investigaciones se resaltan las tendencias y temas emergentes presentados en revistas científicas, así como el análisis de patrones de colaboración que contribuyen al perfil de los actores clave en el campo de conocimiento analizado; en otras palabras, presentan una cartografía de la estructura intelectual de un campo específico dentro del cuerpo existente de literatura científica.

Así, los datos usados en esta clase de análisis se caracterizan por ser extensos y objetivos (el número de citas y publicaciones, la frecuencia de las palabras clave o los temas abordados). Sin embargo, su interpretación es completa ya que mezcla evaluaciones objetivas (como el análisis de desempeño) con aproximaciones subjetivas (como el análisis temático), sustentadas en metodologías rigurosas y fundamentadas (Passas, 2024)⁹.

Por ello, al sistematizar estos extensos volúmenes de datos no estructurados, la bibliometría (cienciometría) permite dibujar mapas del conocimiento científico acumulado y comprender las dinámicas evolutivas de campos consolidados (Donthu *et al.*, 2021)¹⁰. Por ello, un análisis bibliométrico aplicado adecuadamente puede mapear las diferentes rutas construidas para el acceso a un campo del conocimiento específico.

Aunque el análisis bibliométrico brinda diversas ventajas, su aplicación para generar nuevo conocimiento es relativamente reciente (una década), y aún no se ha podido analizar completamente su potencial como herramienta científica. Por esta situación, los estudios bibliométricos, al sustentarse en grupos limitados de datos y enfoques metodológicos, presentan como principal limitante una visión fragmentada de los campos examinados (Caputo y Mariya, 2022)¹¹.

Dada esta circunstancia, se recomienda complementar este tipo de estudios con cartografía científica (como VOSviewer). Esta técnica tiene por objetivo visualizar la estructura y dinámica de la investigación científica mediante un enfoque multidimensional: revisión de publicaciones citadas

7 Tang, Ling, Jieyi Li, Hongchuan Du, Ling Li, Jun Wu y Shouyang Wang, “Big data in forecasting research: a literature review”, *Big Data Research* 27 (2022): 100289.

8 Fathi, Marzieh, Mostafa Haghi Kashani, Seyed Mahdi Jameii y Ebrahim Mahdipour, “Big data analytics in weather forecasting: A systematic review”, *Archives of Computational Methods in Engineering* 29, no. 2 (2022): 1247-1275.

9 Ioannis Passas, “Bibliometric analysis: the main steps”, *Encyclopedia* 4, no. 2 (2024).

10 Donthu, Naveen, Satish Kumar, Debmalya Mukherjee, Nitesh Pandey y Weng Marc Lim, “How to conduct a bibliometric analysis: An overview and guidelines”, *Journal of business research* 133 (2021): 285-296.

11 Andrea Caputo y Mariya Kargina, “A user-friendly method to merge Scopus and Web of Science data during bibliometric analysis”, *Journal of Marketing Analytics* 10, no. 1 (2022): 82-88.

que ayuda a identificar las publicaciones con mayor influencia; el análisis de cocitación, que muestra las relaciones entre las obras referenciadas, las redes bibliográficas que agrupan las publicaciones a partir de referencias comunes, el análisis de copalabras, que exhibe las conexiones temáticas entre los manuscritos, y el análisis de coautoría que permite comprender las interacciones sociales y colaborativas entre autores (Pessin *et al.*, 2022)¹².

Asimismo, el análisis de citas estudia cómo se relacionan las publicaciones a través de las referencias bibliográficas con el fin de identificar trabajos clave y tendencias dentro de un área del conocimiento. Por su parte, la cocitación analiza la frecuencia con la que dos documentos son citados en conjunto, lo que permite detectar vínculos entre ellos y reconocer temas de investigación relevantes.

El acoplamiento bibliográfico opera de forma inversa: conecta documentos recientes que comparten las mismas referencias, lo cual sugiere afinidades temáticas. Por otro lado, el análisis de copalabras se enfoca en la coincidencia de términos clave dentro de los registros, y esto facilita la identificación de relaciones entre diferentes áreas de investigación. Finalmente, el análisis de coautoría examina las redes de colaboración entre investigadores, ya sea a nivel individual o institucional, basándose en su producción científica conjunta.

Sin embargo, dado que los estudios bibliométricos pueden ofrecer perspectivas limitadas si se utilizan de manera aislada, es crucial complementar esta técnica con herramientas que amplíen y enriquezcan el análisis. En este contexto, la cartografía científica se presenta como una estrategia poderosa y visualmente atractiva que permite representar de manera gráfica las relaciones, estructuras y dinámicas del conocimiento científico.

Frente a esta limitación, la cartografía científica se presenta como una herramienta metodológica clave. No solo ayuda a visualizar de manera gráfica las estructuras, relaciones y dinámicas del conocimiento científico, sino que también enriquece nuestra comprensión de esos patrones invisibles que organizan la producción académica. Gracias a su habilidad para integrar y representar visualmente redes de coautoría, patrones de citación, conexiones temáticas y clústeres conceptuales, la cartografía científica convierte los datos bibliométricos en representaciones cognitivas que son más accesibles, fáciles de interpretar y estratégicas.

Además, este enfoque visual hace que sea más fácil identificar comunidades científicas, áreas emergentes, nodos de innovación e incluso vacíos en la investigación. Esto es crucial tanto para las personas investigadoras como para los tomadores de decisiones en ciencia, tecnología e innovación. En resumen, la cartografía científica no solo complementa a la bibliometría, sino que la enriquece y la potencia, al añadir una dimensión relacional y espacial que permite una comprensión más profunda y completa del conocimiento científico en constante evolución.

METODOLOGÍA

En este estudio, se empleó el análisis bibliométrico como metodología de investigación. Este enfoque utiliza técnicas matemáticas y estadísticas para examinar las bases de datos de publicaciones científicas. Es importante destacar que esta metodología es muy útil para investigar comunidades científicas, la estructura del conocimiento y los temas emergentes relacionados con el impacto del BDA en el *marketing*. Por lo tanto, la revisión de datos bibliométricos, como citas, palabras clave, referencias, colaboraciones, redes y tendencias, proporcionará información valiosa para diseñar una agenda de investigación futura.

Se utilizaron las bases de datos Scopus y Web of Science (WoS) como fuentes de información, reconocidas a nivel internacional por su rigor editorial, amplia variedad temática y cobertura multidisciplinaria. La elección de estas plataformas se basó en criterios de calidad, confiabilidad y relevancia para los objetivos de este estudio. Ambas bases de datos proporcionan registros bibliográficos revisados por pares, lo cual asegura un alto nivel de solidez académica, además de contar con funciones avanzadas para la exportación y manejo de metadatos, que son esenciales para los análisis bibliométricos y de redes científicas abordadas en esta investigación.

12

Pessin, Vilker Zucolotto, Luciana Harue Yamane y Renato Ribeiro Siman, "Smart bibliometrics: an integrated method of science mapping and bibliometric analysis", *Scientometrics* 127, no. 6 (2022): 3695-3718.

La estrategia de búsqueda se diseñó siguiendo los principios de precisión y exhaustividad, con el objetivo de reunir la producción científica más representativa y relevante sobre el tema. En esta etapa, se propuso una fórmula de búsqueda que incluía palabras clave cuidadosamente seleccionadas, operadores booleanos y filtros por idioma, tipo de documento, periodo de publicación e indexación. La selección de estos elementos se basó en una revisión preliminar del estado del arte y en las metodologías observadas en estudios bibliométricos anteriores.

Con el fin de validar esta estrategia y garantizar que el estudio mantuviera una metodología consistente, se realizó una prueba piloto como parte de la fase exploratoria. Este proceso previo fue clave para ajustar los criterios de búsqueda, evaluar la relevancia y coherencia de los resultados obtenidos, así como verificar la efectividad de las herramientas de análisis a utilizar más adelante.

La prueba piloto también ayudó a detectar posibles sesgos, omisiones o redundancias en la recuperación de información. Además, permitió establecer los parámetros necesarios para asegurar que el estudio fuera replicable y confiable en su totalidad. Solo después de confirmar que los procedimientos eran sólidos y apropiados, se avanzó con la ejecución completa del análisis.

Como fuentes de información se emplearon las bases de datos Scopus y Web of Science (WoS), debido a su amplia cobertura, calidad de los registros y reconocimiento en la comunidad científica. La estrategia de búsqueda se diseñó cuidadosamente para capturar literatura relevante en torno al uso de BDA en el ámbito del *marketing*. Por esta razón, como primer paso se realizó una prueba piloto, cuyo principal propósito fue asegurarse de que el diseño metodológico, los criterios de búsqueda y los procedimientos de análisis fueran válidos y replicables antes de realizar el estudio completo.

Como resultado, se llevó a cabo la validación de la estrategia de búsqueda. Esto significó asegurarse de que las palabras clave elegidas, los operadores booleanos utilizados y los filtros aplicados (como el rango de tiempo, el idioma, el tipo de documento, entre otros) permitieran obtener información relevante, alineada con los objetivos del estudio. Además, se comprobó que las publicaciones obtenidas estuvieran claramente relacionadas con el tema de investigación, asegurando así su relevancia documental.

También se revisó la funcionalidad del software seleccionado para el análisis bibliométrico, y asegurar que fuera compatible con los datos y que pudiera procesar la información de manera efectiva. A partir de esta etapa, se pudieron ajustar y perfeccionar los criterios de búsqueda iniciales, lo cual ayudó a mejorar la precisión de los resultados. Al final, esta prueba piloto permitió evaluar la viabilidad tanto metodológica como técnica del estudio en su conjunto, esto facilitó una planificación más eficiente para el análisis final.

Resultado de la prueba piloto, los términos de búsqueda quedaron definidos como: TITLE ((Big Data OR Data Analytics AND Market* AND Manage*) OR (Intelligence OR Big Data * AND "Digital" AND Busin*) OR (Big Data OR Market * AND "New Business" AND Data*) OR (Big Data OR Intelligence AND "Enterp" AND Market*) OR (Data Analytics OR Intelligence AND "Business" AND Startup*)) OR KEYWORDS AUTHOR ((Big Data OR Data Analytics AND Market* AND Manage*) OR (Intelligence OR Big Data * AND "Digital" AND Busin*) OR (Big Data OR Market * AND "New Business" AND Data*) OR (Big Data OR Intelligence AND "Enterp" AND Market*) OR (Data Analytics OR Intelligence AND "Business" AND Startup*)).

La información fue recopilada de las bases de datos utilizadas en las pruebas piloto. Como se mencionó anteriormente, la selección de estas fuentes se basó en su enfoque multidisciplinario y en el amplio reconocimiento que tienen dentro de la comunidad científica para este tipo de estudios. El análisis cubrió el período 2016-2025, utilizando como término de búsqueda la fórmula mencionada antes, la cual se comparó con las que se usaron en varios artículos identificados durante la revisión bibliográfica.

Los criterios establecidos para la selección fueron: a) artículos científicos, b) documentos escritos en inglés, francés o español; c) publicaciones indexadas en Web of Science (WoS) o Scopus, d) manuscritos evaluados a través del sistema de revisión por pares de doble ciego, y e) el uso de combinaciones de palabras clave de acuerdo con la fórmula de búsqueda definida. Los metadatos se procesaron utilizando las herramientas Biblioshiny (de RStudio) y VOSviewer.

Es fundamental destacar que el uso de Biblioshiny y VOSviewer, como herramientas de análisis, se justifica por sus funciones complementarias en estudios bibliométricos. Biblioshiny, que actúa como una interfaz gráfica del paquete Bibliometrix en R, permite llevar a cabo análisis descriptivos y de rendimiento científico de manera eficiente y reproducible, lo cual facilita la exploración de grandes volúmenes de datos (Aria y Cuccurullo, 2017)¹³.

VOSviewer es muy conocido por su habilidad para crear y visualizar mapas de redes de coocurrencia, coautoría y cocitación. Lo anterior permitió descubrir estructuras, conexiones y tendencias en la producción científica (Van Eck y Waltman, 2010)¹⁴. Al combinar ambas herramientas se obtiene una visión más completa y gráficamente atractiva del área de investigación. En la búsqueda inicial, se encontraron 315 artículos científicos; luego, se aplicaron los filtros disponibles en la plataforma de WoS y Scopus, siguiendo los criterios de inclusión para garantizar la relevancia e impacto de las publicaciones. De todos los documentos encontrados, se eligieron 266 registros que cumplieran con los criterios establecidos.

Los metadatos completos de cada publicación se exportaron en formatos BibTeX y TXT, lo que permitió su análisis computacional posterior. Estos archivos fueron procesados utilizando software de código abierto, especialmente Bibliometrix y su interfaz gráfica Biblioshiny, que se desarrollaron en el entorno estadístico R. Bibliometrix permitió llevar a cabo un análisis bibliométrico detallado, que incluye indicadores de productividad científica, el impacto de las publicaciones, la colaboración entre autores, instituciones y países, así como un análisis temático de palabras clave y cocitación. Por otro lado, Biblioshiny logró que la interacción con los datos fuera más sencilla a través de una plataforma amigable e intuitiva, permitiendo crear visualizaciones dinámicas que enriquecen la interpretación de los hallazgos.

Además, se utilizó VOSviewer, una herramienta especializada en la construcción y visualización de mapas bibliométricos, para complementar la representación gráfica de redes de coautoría, cocitación y coocurrencia de términos. Esta integración de herramientas permitió generar mapas y gráficos interactivos que representan la evolución temática, las principales líneas de investigación, los nodos de mayor influencia en la producción científica y los vínculos entre diferentes actores e instituciones a lo largo del periodo analizado.

El enfoque metodológico implementado en esta investigación permitió una visión holística, crítica y evolutiva de estudio del impacto del BDA en el *marketing*. Con la aplicación del análisis bibliométrico y de técnicas de mapeo científico fue posible medir la productividad académica (en términos de volumen de publicaciones, autores más influyentes, fuentes recurrentes y países líderes en la generación de conocimiento) e identificar patrones de colaboración, núcleos temáticos consolidados y áreas de investigación en consolidación o declive.

El principal aporte de esta perspectiva fue la capacidad para detectar tendencias emergentes; es decir, conceptos, metodologías o aplicaciones que comienzan a ganar tracción en la literatura científica, pero que aún no han sido plenamente explorados. Este tipo de hallazgos es relevante en un campo dinámico y en constante evolución como lo es el marketing digital impulsado por tecnologías de análisis de datos. Asimismo, el estudio permitió detectar brechas en la literatura especializada en el tema, lo cual constituye una oportunidad valiosa para orientar futuras investigaciones hacia temas poco abordados, y con ello contribuir al enriquecimiento del conocimiento académico.

En síntesis, la metodología aplicada proporcionó una radiografía del ecosistema científico en torno a BDA, lo cual permite visualizar no solo la cantidad, sino la calidad, la estructura y la direccionalidad del conocimiento producido. Por medio de la construcción de mapas de coocurrencia de palabras clave, redes de coautoría y análisis de citaciones, se evidenciaron las dinámicas que conducen el desarrollo del campo de estudio, como los vínculos interinstitucionales, las líneas de investigación más influyentes y la evolución de los enfoques conceptuales a lo largo del tiempo. Esta perspectiva integral no solo documenta el pasado y describe el presente, sino que también orienta estratégicamente hacia el futuro, al delimitar con claridad los temas que están emergiendo, los cuales presentan saturación teórica y los que aún requieren mayor exploración.

13 Massimo Aria y Corrado Cuccurullo, "bibliometrix: An R-tool for comprehensive science mapping analysis", *Journal of informetrics* 11, 4 (2017): 59-975.

14 Nes Van Eck y Ludo Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping", *Scientometrics* 84, 2 (2010): 523-538

RESULTADOS/DISCUSIÓN

De acuerdo con los resultados obtenidos, el número anual de publicaciones relacionadas con el uso de *big data* en *marketing* presenta una tendencia creciente sostenida a partir del año 2018, lo cual marca un punto de inflexión en la evolución del campo. Este comportamiento se intensificó significativamente en el 2022, cuando se registra el mayor incremento interanual, con un crecimiento del 100 % respecto al año anterior, alcanzando un total de 60 artículos publicados. Este repunte contrasta con los niveles previos al 2018, periodo durante el cual el número de publicaciones anuales no superaba las 15, lo cual evidencia una etapa incipiente del desarrollo del área temática.

El año 2019 también se destaca por un aumento notable en la producción científica, con un crecimiento del 50 % (16 artículos), consolidando así una fase de expansión progresiva del interés académico en la intersección entre BDA y *marketing*. Esta evolución cuantitativa sugiere que, a partir del 2018, la comunidad científica comenzó a percibir el potencial transformador de las tecnologías de datos masivos en los procesos de toma de decisiones, segmentación de mercados, análisis del comportamiento del consumidor y optimización de estrategias de valor.

Este crecimiento muestra una maduración temática, que refleja el aumento en la disponibilidad de datos, el desarrollo de nuevas herramientas analíticas y la necesidad de respuestas más precisas y personalizadas en los entornos comerciales digitales. El renovado interés de las personas investigadoras también puede atribuirse al creciente protagonismo del *marketing* basado en datos en las agendas de innovación empresarial, así como a la consolidación de enfoques metodológicos interdisciplinarios que integran estadística, inteligencia artificial, economía del comportamiento y ciencia de datos.

Por consiguiente, la evolución en la cantidad de publicaciones no solo da cuenta del dinamismo del área, sino que también visibiliza un proceso de legitimación académica, donde el uso de *big data* en *marketing* deja de ser emergente para convertirse en un objeto de estudio consolidado, con líneas de investigación cada vez más definidas, rigurosas y diversificadas.

En ese sentido, la evolución del estudio del BDA en el *marketing* puede entenderse como un proceso de maduración progresiva que ha dado lugar a cuatro grandes vertientes temáticas, cada una con características, niveles de consolidación y enfoques metodológicos distintos (Figura 1). Estas vertientes no solo permiten estructurar el conocimiento acumulado, sino que también ilustran la forma en que el campo ha transitado desde preocupaciones iniciales hacia enfoques más complejos y sofisticados:

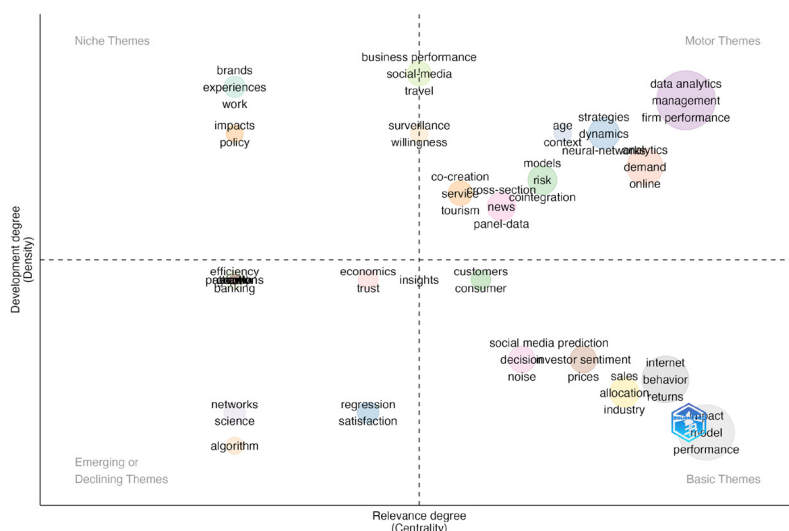
- a.) Temas básicos: esta primera categoría agrupa aquellos estudios orientados a la exploración de fenómenos fundamentales en la intersección entre BDA y *marketing*. Se enfocan principalmente en el análisis de datos provenientes de redes sociales digitales, la identificación de emociones del consumidor mediante técnicas de procesamiento de lenguaje natural y minería de opiniones, así como en el estudio de decisiones de compra vinculadas a factores como el precio, las promociones o la reputación en línea. Estos trabajos cumplen una función introductoria en el campo, al demostrar el potencial de las herramientas de BDA para captar y comprender comportamientos básicos del consumidor en entornos digitales.
- b.) Temas motores: constituyen el núcleo dinámico del campo, aquellos tópicos que no solo generan una gran cantidad de publicaciones, sino que también ejercen una influencia estructurante sobre las demás áreas. Aquí se encuentran investigaciones sobre la gestión estratégica de la analítica de datos dentro de las organizaciones, el uso de redes neuronales y modelos predictivos avanzados para anticipar comportamientos de compra, así como la creación de perfiles de consumidores altamente segmentados. Estos estudios tienen un carácter metodológicamente más robusto y se apoyan en técnicas como *machine learning*, modelado estadístico y sistemas de recomendación, contribuyendo significativamente a la profesionalización del *marketing* orientado por datos.
- c.) Temas nicho: esta categoría incluye líneas de investigación más especializadas y de menor volumen, pero no por ello menos relevantes. Se centran en aspectos como la experiencia del usuario con las marcas, la medición del impacto emocional en interacciones digitales, y la evaluación de la lealtad del cliente a través de indicadores conductuales y métricas longitudinales. Estos estudios abordan dimensiones más subjetivas y cualitativas del *marketing*, al integrar metodologías mixtas que combinan análisis cuantitativo con aproximaciones etnográficas o psicográficas. Si bien su nivel

de citación suele ser menor, su aportación es crucial para la comprensión profunda del vínculo entre consumidor y marca.

- d.) Temas emergentes: esta última vertiente está compuesta por áreas de investigación en rápida expansión, caracterizadas por un alto grado de innovación. En este grupo, se encuentran desarrollos recientes centrados en la construcción de nuevos algoritmos para la automatización del análisis de datos, modelos para medir con precisión el nivel de satisfacción del cliente en tiempo real, sistemas para evaluar la eficiencia de campañas de *marketing* digital y herramientas para interpretar preconcepciones o sesgos cognitivos del consumidor. Estos temas se ubican en la frontera del conocimiento y apuntan a transformar no solo el *marketing* como disciplina, sino también las prácticas empresariales en general.

En conjunto, estas cuatro vertientes conforman un ecosistema dinámico y evolutivo, en el que los temas pueden transitar de una categoría a otra conforme maduran, obtienen visibilidad o se integran en agendas de investigación más amplias. Esta clasificación, por tanto, permite visualizar de manera estructurada el panorama actual del BDA en *marketing*, al tiempo que ofrece una guía útil para identificar oportunidades futuras de investigación, colaboración interdisciplinaria e innovación aplicada.

Figura 1. Mapeo de temáticas

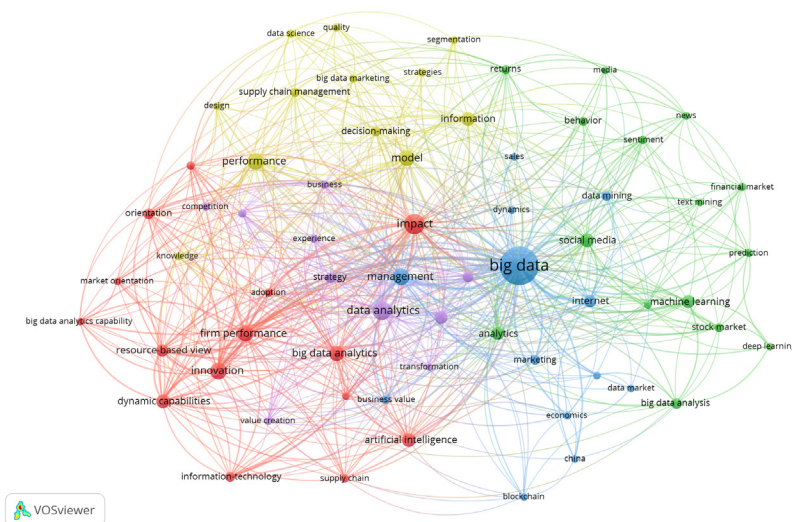


Fuente: Scopus/WoS/Biblioshiny (2024).

Las coocurrencias de términos permiten visualizar una red semántica que revela cómo se interrelacionan los conceptos vinculados al BDA en el contexto del *marketing*, como se muestra en la Figura 2. En esta red, el término “big data” destaca como el nodo central, representado por un círculo azul de mayor tamaño, lo cual indica su papel protagonista en el conjunto de publicaciones analizadas. Esta centralidad sugiere que *big data* actúa como eje articulador de múltiples enfoques y aplicaciones dentro del campo.

Alrededor de este nodo principal emergen otros conceptos clave como “*data analytics*”, “*machine learning*”, “*artificial intelligence*”, “*firm performance*” e “*impact*”, los cuales presentan una alta frecuencia de coocurrencia con BDA. Esta proximidad semántica indica que dichos términos no solo están estrechamente vinculados al uso de BDA, sino que también reflejan las principales líneas de investigación y aplicación en el ámbito del *marketing*. En conjunto, la red sugiere una convergencia temática hacia el aprovechamiento de tecnologías avanzadas para mejorar el rendimiento empresarial y generar valor estratégico.

Figura 2. Coocurrencia de términos



Fuente: Scopus/WoS/VOSviewer (2024).

En el análisis de mapeo temático se identifican cuatro clústeres principales, cada uno representando un eje conceptual dentro del estudio del BDA en el ámbito del *marketing*:

- 1) Tecnología y análisis de datos (azul): este clúster agrupa términos como “internet”, “social media”, “analytics”, “marketing”, “blockchain” y “big data analysis”. Su composición refleja el papel de BDA como herramienta clave en entornos digitales, donde la interconexión entre plataformas tecnológicas y estrategias de *marketing* permite una gestión más eficiente de los datos y una mejor comprensión del comportamiento del consumidor. Además, la presencia de “blockchain” sugiere una creciente preocupación por la seguridad y trazabilidad de los datos en estos ecosistemas.
- 2) Aprendizaje automático y minería de datos (verde): este grupo incluye términos como “machine learning”, “data mining”, “text mining”, “prediction”, “deep learning” y “financial market”. La concentración de estos conceptos evidencia la estrecha relación entre BDA y las técnicas de inteligencia artificial, especialmente en tareas de predicción y análisis de grandes volúmenes de datos no estructurados. Asimismo, la mención del “mercado financiero” indica aplicaciones concretas en sectores donde la anticipación de tendencias y comportamientos resulta crítica.
- 3) Rendimiento empresarial e innovación (rojo): este clúster está conformado por términos como “firm performance”, “innovation”, “dynamic capabilities” y “value creation”. Su presencia destaca el impacto estratégico de BDA en la mejora del desempeño organizacional, la generación de capacidades dinámicas y la creación de valor. En este sentido, BDA no solo se percibe como una herramienta operativa, sino como un catalizador de innovación y ventaja competitiva en entornos empresariales cada vez más dinámicos.
- 4) Gestión y modelos de negocio (amarillo): finalmente, este grupo incluye conceptos como “supply chain management”, “decision-making”, “business”, “quality” y “model”. Aquí se pone de manifiesto el rol de BDA en la optimización de procesos estratégicos, especialmente en la toma de decisiones informadas, la mejora de la calidad operativa y el diseño de modelos de negocio más adaptativos y eficientes.

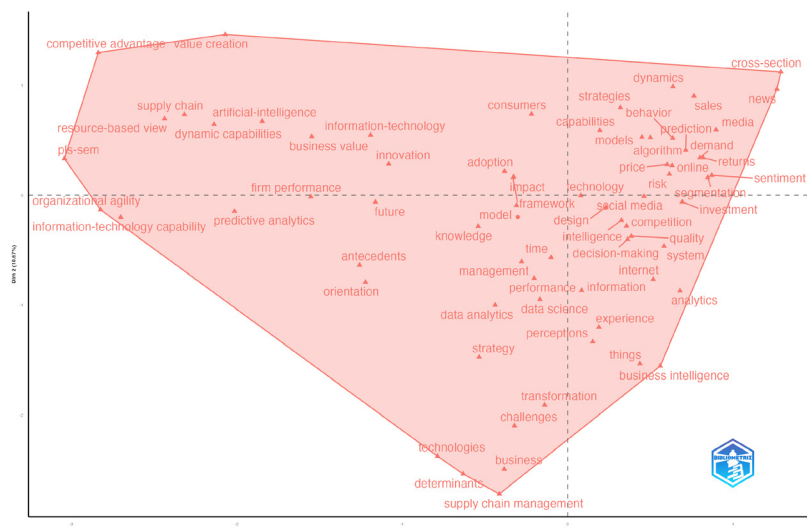
Como complemento, el Mapa de Estructura Conceptual (Petrovich, 2021)¹⁵ permitió identificar y representar gráficamente las relaciones semánticas entre los principales términos extraídos del corpus analizado (Grandi *et al.*, 2024)¹⁶. Este se construyó a partir de la co-ocurrencia de palabras clave, revelando agrupamientos temáticos que reflejan la estructura intelectual del campo de estudio sobre BDA y *marketing* en el periodo 2016-2025 (Figura 3).

15 Eugenio Petrovich, “Science mapping and science maps”, *Knowledge Organization* 48, no. 7-8 (2021): 535-562.

16 Roges Grandi, Alba Valeria de Sant’Anna de Freitas Loliola, Leandro Krug Wive y Raquel Salcedo Gomes, “A systematic review of the literature on semantic networks of concept maps: study supported by a specialized bibliometric process”, *RDBC: Revista Digital de Biblioteconomia e Ciência da Informação* 22 (2024): e024025.

En ese sentido, el gráfico está compuesto por dos dimensiones principales: la dimensión 1 (eje horizontal), que explica el 25,3 % de la variación total y la dimensión 2 (eje vertical), que representa el 10,9 %. La combinación de ambas dimensiones permite visualizar la distribución y proximidad de los conceptos, al facilitar la identificación de subtemas o líneas de investigación consolidadas.

Figura 3. Mapa de estructura conceptual



Fuente: Scopus/WoS/Biblioshiny (2024).

De acuerdo con el mapa de estructura conceptual, el campo de estudio del *big data* aplicado a la gestión empresarial y el *marketing* se organiza en seis grupos temáticos principales, que delimitan los focos de atención y las líneas de investigación más consolidadas. Estos clústeres no solo reflejan la diversidad del campo, sino que también muestran cómo se interrelacionan distintos enfoques estratégicos, tecnológicos y operativos:

- Enfoque estratégico y de creación de valor: ubicado en la parte superior izquierda del mapa, este clúster agrupa conceptos fundamentales como “competitive advantage”, “value creation”, “resource-based view” y “dynamic capabilities”. Estos términos revelan una perspectiva que asocia el uso del *big data* con la formulación estratégica y la construcción de ventajas competitivas sostenibles. En este grupo, el análisis de datos no es visto solo como una herramienta operativa, sino como un activo clave que se integra en la estrategia corporativa para generar valor a largo plazo, aprovechando recursos internos y capacidades dinámicas.
- Capacidades tecnológicas y analítica predictiva: localizado también en el cuadrante izquierdo, este grupo incluye nociones como “organizational agility”, “information-technology capability”, “predictive analytics” e “innovation”. Representa la dimensión tecnológica del campo, donde se estudia cómo las organizaciones desarrollan capacidades digitales y tecnológicas para adaptarse ágilmente a entornos cambiantes. La analítica predictiva, en particular, aparece como una capacidad avanzada que permite anticipar comportamientos, optimizar procesos y fomentar la innovación en productos y servicios.
- Comportamiento del consumidor y marketing digital: este clúster, situado en el cuadrante derecho del mapa, gira en torno a términos como “online”, “segmentation”, “social media”, “sentiment”, “price” y “returns”. Se trata de una línea de investigación centrada en entender al consumidor en entornos digitales, mediante técnicas de análisis de datos aplicadas al comportamiento online, la segmentación de mercados, la detección de emociones en redes sociales y la valoración de la respuesta del cliente frente a precios o campañas. Aquí, *big data* se convierte en una herramienta esencial para el *marketing* orientado al cliente y la personalización de experiencias.

- d.) Infraestructura analítica y toma de decisiones: en la parte inferior derecha del mapa, se agrupan palabras clave como “business intelligence”, “analytics”, “decision-making”, “system” e “intelligence”. Este grupo refleja el interés por el desarrollo de sistemas y plataformas que facilitan la recolección, procesamiento y análisis de grandes volúmenes de datos. La finalidad es apoyar la toma de decisiones informadas, automatizar procesos analíticos y dotar a las organizaciones de inteligencia operativa. Se trata de una dimensión más técnica y centrada en la arquitectura de datos y la eficiencia en la gestión de la información.
- e.) Transformación digital y gestión operativa: este clúster, que aparece en la parte inferior del mapa, está compuesto por términos como “supply chain management”, “technologies”, “challenges” y “transformation”. Señala la creciente aplicación del *big data* en áreas operativas como la cadena de suministro, la logística y la producción. También destaca los desafíos que enfrentan las organizaciones al incorporar tecnologías emergentes y rediseñar sus procesos. Aquí, el *big data* es interpretado como un catalizador para la transformación digital, al mejorar la eficiencia, la trazabilidad y la capacidad de respuesta de los sistemas operacionales.
- f.) Ejes transversales: en el centro del mapa se ubican conceptos como “data analytics”, “strategy”, “impact”, “framework” y “knowledge”, que actúan como elementos articuladores entre los distintos clústeres. Estos términos representan constructos clave que sustentan el desarrollo teórico del campo y permiten la integración de enfoques diversos. Funcionan como puentes que vinculan la estrategia, la tecnología, el *marketing* y la operación, lo cual contribuye a una comprensión holística del impacto del BDA en la gestión empresarial.

En conjunto, el mapa de estructura conceptual reafirma el carácter multidimensional y dinámico del estudio del BDA en el ámbito del *marketing*, mostrando cómo este campo ha evolucionado desde enfoques fragmentados hacia una visión integrada y transversal, donde confluyen diversas dimensiones como la tecnología, la estrategia empresarial, el comportamiento del consumidor y la transformación organizacional. Esta convergencia no solo refleja la madurez teórica del campo, sino que también evidencia su complejidad práctica, al incorporar herramientas analíticas avanzadas en múltiples niveles de toma de decisiones.

La visualización también destaca una densa interconectividad entre los conceptos clave, lo cual revela que el uso de BDA en *marketing* no se limita a un área funcional específica, sino que opera como un eje transversal que permea desde las operaciones tecnológicas hasta la formulación estratégica. En particular, la notable asociación entre BDA y *machine learning* subraya el papel central que juega la automatización, el modelado predictivo y el aprendizaje computacional como motores de innovación y diferenciación competitiva en el análisis de datos.

En este contexto, el mapa no solo facilita una comprensión integral del estado actual del conocimiento, sino que también permite vislumbrar líneas de investigación emergentes, posibles vacíos temáticos y nuevas oportunidades de exploración interdisciplinaria. Esta visión estructurada y conectada del campo constituye una base sólida para la reflexión crítica sobre sus tendencias, limitaciones y potencial futuro. A partir de este análisis, se derivan las siguientes conclusiones, que sintetizan los hallazgos más relevantes y plantean orientaciones clave para el avance teórico y práctico del uso del BDA en el *marketing*.

CONCLUSIONES

Los hallazgos de esta investigación revelan una brecha significativa en la integración de BDA en el análisis y la estrategia de *marketing*. Las empresas aún enfrentan desafíos sustanciales para aprovechar de manera efectiva el BDA en la planificación, implementación y gestión de relaciones con los clientes. En particular, la conexión entre las técnicas avanzadas de *big data* y los beneficios concretos para el *marketing* no está claramente definida, lo cual resalta la necesidad de enfoques más estructurados y estratégicos que permitan cerrar esta brecha.

Además, la investigación sobre el uso del BDA en *marketing* sigue en una fase incipiente, con una notoria escasez de estudios que aborden de manera integral su aplicación e impacto en el sector. La falta de validación empírica de los modelos teóricos y marcos conceptuales propuestos en la literatura limita el desarrollo de conocimientos aplicables. Asimismo, la ausencia de estudios longitudinales

dificulta la comprensión de los efectos a largo plazo de la personalización basada en BDA sobre la confianza y la lealtad del consumidor, lo que impide evaluar con precisión su efectividad en la construcción de relaciones duraderas con los clientes.

Desde una perspectiva organizacional, el aprovechamiento del BDA implica requerimientos considerables en infraestructura tecnológica, talento humano especializado y capacidades analíticas al interior de los departamentos de *marketing*. Sin embargo, muchas empresas (especialmente las pequeñas y medianas) enfrentan restricciones en la adquisición y gestión de estos recursos. Las carencias en formación y competencias digitales son una barrera crítica para la apropiación efectiva de estas tecnologías.

Además, a medida que el uso de BDA en *marketing* se expande, emergen con fuerza preocupaciones éticas y sociales relacionadas con la privacidad de los consumidores y la transparencia en el uso de sus datos. Estos desafíos exigen marcos regulatorios sólidos, mecanismos de autorregulación y el desarrollo de una cultura organizacional basada en principios de ética digital.

En este sentido, futuras investigaciones deben enfocarse en desarrollar metodologías que faciliten la integración del BDA en las estrategias de *marketing*, abordando tanto los desafíos técnicos como organizacionales. Asimismo, se recomienda profundizar en la formación de capacidades dentro de las empresas, al fomentar el desarrollo de habilidades analíticas y el uso eficiente de herramientas de BDA.

La combinación de estudios teóricos y empíricos permitirá generar un conocimiento más sólido y aplicable, lo cual contribuye con el fortalecimiento del *marketing* basado en datos y a la optimización de las decisiones estratégicas en el entorno empresarial actual.

Finalmente, conforme la aplicación de BDA al *marketing* se incrementa, resultará fundamental abordar las preocupaciones sobre la privacidad de los datos y las cuestiones éticas asociadas. En consecuencia, los análisis deberán enfocarse en identificar estrategias que permitan equilibrar los beneficios de BDA con la necesidad de proteger la privacidad de las personas consumidoras y garantizar un uso ético de la información.

Esto implicará el diseño de marcos regulatorios sólidos, la implementación de prácticas de transparencia en la recopilación y procesamiento de datos, así como la promoción de estándares éticos en el uso de tecnologías avanzadas de análisis de datos.

REFERENCIAS

- Aria, Massimo. y Cuccurullo, Corrado. "bibliometrix: An R-tool for comprehensive science mapping analysis". *Journal of informetrics* 11, 4 (2017): 959-975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Caputo, Andrea y Mariya Kargina. "A user-friendly method to merge Scopus and Web of Science data during bibliometric analysis". *Journal of Marketing Analytics* 10, no. 1 (2022): 82-88. <https://doi.org/10.1057/s41270-021-00142-7>
- Donthu, Naveen, Satish Kumar, Debmalya Mukherjee, Nitesh Pandey y Weng Marc Lim. "How to conduct a bibliometric analysis: An overview and guidelines". *Journal of business research* 133 (2021): 285-296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
- Fathi, Marzieh, Mostafa Haghi Kashani, Seyed Mahdi Jameii y Ebrahim Mahdipour. "Big data analytics in weather forecasting: A systematic review". *Archives of Computational Methods in Engineering* 29, no. 2 (2022): 1247-1275. <https://doi.org/10.1007/s11831-021-09616-4>
- Grandi, Roges, Alba Valeria de Sant'Anna de Freitas Loiola, Leandro Krug Wive y Raquel Salcedo Gomes. "A systematic review of the literature on semantic networks of concept maps: study supported by a specialized bibliometric process". *RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação* 22 (2024): e024025. <https://doi.org/10.20396/rdbci.v21i00.8674477/en>
- Gupta, Shivam, Théo Justy, Shampy Kamboj, Ajay Kumar y Eivind Kristoffersen. "Big data and firm marketing performance: Findings from knowledge-based view". *Technological Forecasting and Social Change* 171 (2021): 120986. <https://doi.org/10.20396/rdbci.v21i00.8674477/en>
- Lu, Yusheng y Jiantong Zhang. "Bibliometric analysis and critical review of the research on big data in the construction industry". *Engineering, Construction and Architectural Management* 29, no. 9 (2022): 3574-3592. <https://doi.org/10.1108/ECAM-01-2021-0005>
- Ningsih, Fitriah, Muhammad Ilham y Murtiadi Awaluddin. "Industrial Competitiveness through Porter's Five Forces Framework". *Al-Kharaj: Journal of Islamic Economic and Business* 7, no. 2 (2025). <https://doi.org/10.24256/kharaj.v7i2.7370>
- Passas, Ioannis. "Bibliometric analysis: the main steps". *Encyclopedia* 4, no. 2 (2024). <https://doi.org/10.3390/encyclopedia4020065>
- Pessin, Vilker Zucolotto, Luciana Harue Yamane y Renato Ribeiro Siman. "Smart bibliometrics: an integrated method of science mapping and bibliometric analysis". *Scientometrics* 127, no. 6 (2022): 3695-3718. <https://doi.org/10.1007/s11192-022-04406-6>
- Petrovich, Eugenio. "Science mapping and science maps". *Knowledge Organization* 48, no. 7-8 (2021): 535-562. <https://hdl.handle.net/11365/1210644>
- Tang, Ling, Jieyi Li, Hongchuan Du, Ling Li, Jun Wu y Shouyang Wang. "Big data in forecasting research: a literature review". *Big Data Research* 27 (2022): 100289. <https://doi.org/10.1016/j.bdr.2021.100289>
- Van Eck, Nes y Waltman, Ludo Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 84, 2 (2010): 523-538. <https://doi.org/10.1007/s11192-009-0146-3>
- Zhang, Yucheng, Meng Zhang, Jing Li, Guangjian Liu, Miles M. Yang y Siqi Liu. "A bibliometric review of a decade of research: Big data in business research-Setting a research agenda". *Journal of Business Research* 131 (2021): 374-390. <https://doi.org/10.1016/j.jbusres.2020.11.004>