

Facultad Latinoamericana de Ciencias Sociales, FLACSO Ecuador

Departamento de Desarrollo, Ambiente y Territorio

Convocatoria 2021 - 2023

Tesis para obtener el título de Maestría en Economía del Desarrollo

ANÁLISIS BIBLIOMÉTRICO DEL *BIG DATA* Y EL COMPORTAMIENTO DEL
CONSUMIDOR UTILIZANDO LA BASE DE DATOS DE SCOPUS EN UN PERIODO
DEL 2012 AL 2023

Arévalo Espinoza Oscar Marcelo

Asesor: Martín Mayoral Fernando

Lectores: Albornoz Barriga María Belén, Galvis Correa Andrés Alejandro

Quito, enero de 2025

Dedicatoria

Esta investigación está dedicada principalmente a Dios, quien me ha brindado la fortaleza y sabiduría necesarias para alcanzar este objetivo; a mis padres, por su apoyo incondicional y su confianza constante en mí, siendo un pilar fundamental en mi vida y a mis hermanos, quienes, a pesar de la distancia, siempre han estado presentes, brindándome su apoyo y aliento.

Índice de contenidos

Resumen	8
Agradecimientos	9
Introducción.....	10
Capítulo 1. Marco Teórico.....	15
1.1 Comportamiento del Consumidor	15
1.2 <i>Big Data</i>	19
Capítulo 2. Metodología.....	22
Capítulo 3. Resultados	32
3.1. Discusión de Resultados.....	66
3.1.1. Evolución e Impacto de la Investigación sobre <i>Big Data</i> y Comportamiento del Consumidor	66
3.1.2 Autores, Instituciones y Redes de Investigación Influyentes.....	68
3.1.3 Principales Modelos y Estudios que Conforman la Base del Conocimiento	68
3.1.4 Tendencias Actuales y Emergentes en la Literatura	69
3.1.5 Aplicaciones Prácticas y Relevancia para la Industria.....	70
Conclusiones y Recomendaciones	74
Referencias	77
Anexos	84

Lista de ilustraciones

Gráficos

Gráfico 3. 1. Producción Científica Anual en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	34
Gráfico 3. 2. Citas promedio por año en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	35
Gráfico 3. 3. Distribución y Colaboración Global en Publicaciones en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	40
Gráfico 3. 4. Producción científica por países en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	42
Gráfico 3. 5. Áreas temáticas en investigaciones en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	44
Gráfico 3. 6. Nube de palabras en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	46
Gráfico 3. 7. Red de Co-citación entre autores en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	47
Gráfico 3. 8. Mapa de estructura conceptual en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	49
Gráfico 3. 9. Diagrama de evolución de Sankey en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	51
Gráfico 3. 10. Red de Ocurrencia de Palabras Clave en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	52
Gráfico 3. 11. Mapa Temático en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	54
Gráfico 3. 12.. Relación entre Autores, Palabras Clave y Afiliaciones en la Investigación de Big Data y Comportamiento del Consumidor en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	56

Gráfico 3. 13. Red de Colaboración de autores en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	63
Gráfico 3. 14. Red de Colaboración de afiliaciones en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023.....	64
Gráfico 3. 15. Mapa Mundial de Colaboración entre Países en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	66

Tablas

Tabla 2. 1. Análisis Bibliométricos relacionados al Big Data y el Comportamiento del Consumidor	25
Tabla 3. 1. Información Principal Sobre la Base de Datos de Scopus	33
Tabla 3.2. Autores más relevantes en el campo del Big Data y Comportamiento del Consumidor	36
Tabla 3. 3. Impacto de los autores en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	37
Tabla 3. 4. Afiliaciones relevantes en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	38
Tabla 3. 5. Países más citados en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023.....	41
Tabla 3. 6. Artículos más citados de Scopus en el campo de Big Data y Comportamiento del Consumidor en el periodo 1995-2023	43
Tabla 3. 7. Principales resultados de la Herramienta Tree of Science en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023.....	59
Tabla 3. 8. Síntesis de técnicas y enfoques analíticos en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023	72

Anexos

Anexo 4. 1. Citas promedio por año.....	84
Anexo 4. 2. Distribución y Colaboración Global en Publicaciones.....	84

Anexo 4. 3. Producción científica por país.	85
Anexo 4. 4. Áreas temáticas.....	85
Anexo 4. 5. Nube de palabras	86
Anexo 4. 6. Red de Co-citación entre autores.....	86
Anexo 4. 7. Red de Colaboración entre autores	87
Anexo 4. 8. Red de Colaboración entre Afiliaciones	87
Anexo 4. 9. Red de Colaboración entre países.....	88

Declaración de cesión de derecho de publicación de la tesis

Yo, Oscar Marcelo Arévalo Espinoza, autor de la tesis titulada “Análisis Bibliométrico del *Big Data* y el Comportamiento del Consumidor utilizando la Base de Datos de Scopus en un periodo del 2012 al 2023”, declaro que la obra es de mi exclusiva autoría, que la he elaborado para obtener el título de maestría en Economía del Desarrollo, concedido por la Facultad Latinoamericana de Ciencias Sociales, FLACSO Ecuador.

Cedo a la FLACSO Ecuador los derechos exclusivos de reproducción, comunicación pública, distribución y divulgación, bajo la licencia Creative Commons 3.0 Ecuador (CC BY-NC-ND 3.0 EC), para que esta universidad la publique en su repositorio institucional, siempre y cuando el objetivo no sea obtener un beneficio económico.

Quito, enero de 2025



Firma

Oscar Marcelo Arévalo Espinoza

Resumen

El análisis del comportamiento del consumidor ha experimentado un notable crecimiento en las últimas décadas, impulsado por los avances en *Big Data* y tecnologías relacionadas. Este estudio realiza un análisis bibliométrico descriptivo de la literatura sobre *Big Data* y comportamiento del consumidor, examinando 1.861 documentos publicados hasta el 2023 en la base de datos Scopus, utilizando herramientas como Bibliometrix en RStudio y Tree of Science.

El análisis bibliométrico mostró un crecimiento constante en la producción científica, con una tasa anual del 21,51%. Esto refleja el aumento del interés en el tema a lo largo del tiempo, así como un impacto creciente en la comunidad académica, con un promedio de 16,5 citas por documento. Las principales áreas de enfoque incluyen la aplicación de técnicas de minería de datos, aprendizaje automático e inteligencia artificial para estudiar las preferencias y comportamientos de los consumidores. Además, se identificaron temas emergentes como el análisis de sentimientos, el Internet de las Cosas, y la ética y privacidad de los datos.

Autores influyentes como Wang Y., Li Y. y Zhang Y. han contribuido significativamente al desarrollo del campo. Instituciones líderes como la University of Technology Sydney y Tsinghua University destacan por su alta producción científica. China y Estados Unidos y Singapur emergen como líderes en producción científica, aunque se observa una tendencia hacia colaboraciones nacionales más que internacionales.

Esta investigación subraya que, el *Big Data* aplicado al comportamiento del consumidor se encamina hacia una mayor personalización de experiencias, anticipación de necesidades y un enfoque ético en el manejo de datos. Las investigaciones futuras probablemente integrarán tecnologías avanzadas como el aprendizaje automático, análisis de redes sociales, internet de las cosas e inteligencia artificial, mientras abordan desafíos clave relacionados con la privacidad y la colaboración internacional. Este estudio ofrece una base sólida para entender el estado actual y las direcciones futuras de un campo en rápida evolución, con aplicaciones relevantes tanto en la academia como en la industria.

Agradecimientos

Deseo expresar mi profundo agradecimiento a FLACSO por brindarme la oportunidad de formar parte de este programa de maestría, lo cual ha enriquecido mi desarrollo académico y profesional; a mi tutor, Fernando Martín, por su valiosa orientación y guía a lo largo de este proceso; y, asimismo, a mis lectores, María Belén Albornoz y Andrés Alejandro Galvis Correa, por sus observaciones constructivas, las cuales contribuyeron significativamente a fortalecer este trabajo.

Introducción

En las últimas décadas, el estudio del comportamiento del consumidor ha capturado el interés de académicos y profesionales, impulsado por el avance acelerado de las tecnologías digitales y la proliferación de datos generados por los usuarios en línea. El *Big Data* se ha convertido en una herramienta esencial para comprender las dinámicas de mercado y las preferencias de los consumidores, permitiendo a las empresas y académicos analizar grandes volúmenes de información para tomar decisiones más informadas y desarrollar estrategias comerciales efectivas (Thayyib et al. 2023; Kusnetzky 2010).

El comportamiento del consumidor abarca una serie de procesos psicológicos y sociales que influyen en la selección, adquisición y uso de productos y servicios (Mollá et al. 2012; Raiteri 2016). Las teorías clásicas, como la Teoría de la Utilidad de William Jevons, ofrecen una perspectiva económica sobre cómo los consumidores toman decisiones para maximizar su satisfacción (Gómez y Sequería 2015). Sin embargo, enfoques más recientes en economía conductual y psicología del consumidor han evidenciado que las decisiones de compra están profundamente influenciadas por factores emocionales, sociales y cognitivos limitados (Kahneman y Tversky 1979; Thaler 1985).

La transformación digital ha modificado radicalmente las expectativas y comportamientos de los consumidores. La adopción masiva de dispositivos móviles, redes sociales y plataformas de comercio electrónico ha generado un ecosistema en el que los consumidores están constantemente conectados y generan una cantidad inmensa de datos en cada interacción en línea (Ferrer 2018; Montero y El Kadi 2017). Este fenómeno ha dado lugar al "Consumidor 2.0", caracterizado por su participación activa en entornos digitales y su influencia en las percepciones de otros consumidores a través de contenidos generados por usuarios (Muñoz y Aguilar 2020).

En este contexto, el *Big Data* permitirá a las empresas personalizar ofertas y mejorar la experiencia del cliente, y también proporciona a los investigadores una oportunidad de analizar patrones de comportamiento a gran escala y en tiempo real (Gómez 2021). Las técnicas avanzadas de análisis de datos, como la minería de datos, el aprendizaje automático y la inteligencia artificial, han abierto nuevas vías para explorar y comprender los factores que influyen en las decisiones de los consumidores (Erevelles et al. 2016).

No obstante, el abundante volumen de investigaciones sobre *Big Data* y comportamiento del consumidor ha generado una fragmentación en la literatura que dificulta una comprensión

integral del campo (Chávez et al. 2022). Esta dispersión de estudios impide identificar claramente las tendencias emergentes, las conexiones entre diferentes enfoques y las áreas que requieren mayor atención. Sin una síntesis adecuada, es desafiante avanzar en el conocimiento y aplicar de manera efectiva los hallazgos al ámbito empresarial. Por ello, se hace esencial realizar un análisis bibliométrico exhaustivo que permita mapear la evolución de la investigación, reconocer a los principales actores y tendencias, y vincular los hallazgos empíricos con las teorías fundamentales del comportamiento del consumidor.

El presente estudio utiliza la base de datos Scopus para llevar a cabo un análisis bibliométrico exhaustivo de las publicaciones sobre *Big Data* y comportamiento del consumidor entre 1995 y 2023. Empleando herramientas avanzadas como Bibliometrix en RStudio y Tree of Science, se analizan 1.861 documentos para explorar la producción científica, las redes de colaboración, las temáticas emergentes y la base conceptual del campo (Aria y Corrado 2017; Zuluaga et al. 2022). Este enfoque permite organizar estructuradamente la literatura existente, identificar patrones significativos y áreas potenciales para futuras investigaciones.

La revisión de las principales teorías económicas y psicosociales sobre el comportamiento del consumidor es fundamental para contextualizar los hallazgos del análisis bibliométrico. Teorías como la jerarquía de necesidades de Maslow y el condicionamiento operante de Skinner proporcionan un marco para entender cómo las motivaciones intrínsecas y los estímulos externos moldean las prácticas de consumo (Recalde, Suárez y Páez 2021; Orellana 2022; Skinner 1938). La teoría de la racionalidad limitada de Simon (1956) y la teoría prospectiva de Kahneman y Tversky (1979) destacan que las decisiones de los consumidores no siempre son completamente racionales y están influenciadas por sesgos cognitivos y heurísticas.

Al vincular estas teorías con las técnicas de *Big Data*, se abre una oportunidad para enriquecer la comprensión del comportamiento del consumidor. El análisis de grandes volúmenes de datos permite explorar patrones y relaciones que no son evidentes a través de métodos tradicionales, ofreciendo insights profundos sobre las preferencias y comportamientos de los consumidores en diferentes contextos culturales y de mercado (Solomon 2006; Hofacker Malthouse y Sultan 2016).

La metodología de este estudio es de carácter cuantitativo, descriptivo y bibliométrico. A través de una búsqueda exhaustiva con términos clave específicos configurados en una ecuación de búsqueda en Scopus, se recopilaron los documentos relevantes para el periodo de

estudio. El uso de Bibliometrix y Tree of Science facilita el análisis de indicadores bibliométricos, la visualización de redes de colaboración y la identificación de estructuras temáticas y conceptuales (Aria y Corrado 2017; Cobo et al. 2011).

Este análisis meticuloso evalúa la producción académica y las redes de colaboración, y también identifica brechas de conocimiento y tendencias emergentes. Al vincular los avances teóricos y metodológicos, se ofrece una comprensión profunda de cómo la tecnología de *Big Data* está redefiniendo los paradigmas del comportamiento del consumidor. Con ello, se espera contribuir tanto a la teoría académica como a la práctica profesional, sugiriendo direcciones estratégicas para futuras investigaciones y aplicaciones comerciales.

Además, este estudio aborda las consideraciones éticas y de privacidad inherentes al uso de *Big Data* en el análisis del comportamiento del consumidor. La gestión responsable de los datos personales y el cumplimiento de regulaciones son aspectos fundamentales para mantener la confianza del consumidor y asegurar el uso sostenible y beneficioso de estas tecnologías (Colmenarejo 2018; Aragona 2022).

En resumen, el presente trabajo busca proporcionar una visión integral del estado actual de la investigación sobre *Big Data* y comportamiento del consumidor, identificando los principales actores, tendencias y áreas de oportunidad. Al integrar perspectivas teóricas y prácticas, se pretende fortalecer el vínculo entre la academia y la industria, promoviendo el desarrollo de estrategias basadas en datos que sean éticamente responsables y centradas en el consumidor.

Preguntas de investigación:

Pregunta General

¿Como ha evolucionado y cuáles son las tendencias emergentes en el uso del *Big Data* para comprender y predecir el comportamiento del consumidor, identificando autores, instituciones y revistas más influyentes, así como cuales son las futuras direcciones de investigación en este campo?

Preguntas específicas

1. ¿Cómo ha evolucionado y cuál es el impacto de la investigación sobre *Big Data* y comportamiento del consumidor?
2. ¿Quiénes son los autores, universidades y países más productivos e influyentes, y cuáles son las principales redes de investigación en este campo?

3. ¿Cuáles son los principales modelos y estudios que conforman la base del conocimiento en el campo?

4. ¿Cuáles son las tendencias actuales y emergentes en la literatura relacionada con *Big Data* y comportamiento del consumidor?

Objetivos

Objetivo general

Analizar la evolución y las tendencias emergentes en el uso del *Big Data* para comprender y predecir el comportamiento del consumidor, identificando autores, instituciones y países más influyentes, así como futuras direcciones de investigación en este campo, a través de un análisis bibliométrico sobre la base de datos Scopus para el período 1995 hasta 2023.

Objetivos específicos

- Evaluar la evolución y el impacto de la investigación sobre *Big Data* y comportamiento del consumidor a través de un análisis bibliométrico para identificar tendencias en el crecimiento de publicaciones, citas, y principales áreas temáticas para comprender el desarrollo de este campo y su relevancia en la investigación actual.
- Identificar los autores, universidades y países más productivos e influyentes, así como las principales redes de colaboración en el campo del *Big Data* y comportamiento del consumidor, utilizando herramientas de análisis bibliométrico para mapear el ecosistema académico y profesional predominante en este campo de estudio.
- Analizar los principales modelos y estudios que conforman la base del conocimiento en el campo del *Big Data* y comportamiento del consumidor, aplicando técnicas de co-citación y utilizando la metodología de Tree of Science para determinar las teorías y marcos conceptuales que han sido influyentes en la literatura de este campo de investigación.
- Examinar las tendencias actuales y emergentes en la literatura relacionada con *Big Data* y comportamiento del consumidor, destacando nuevos temas de investigación, metodologías innovadoras y aplicaciones emergentes, con el fin de proporcionar un panorama integral que guíe la dirección de futuras investigaciones.

Hipótesis

El análisis bibliométrico de la literatura sobre *Big Data* y comportamiento del consumidor en la base de datos Scopus desde 1995 hasta 2023 permitirá identificar y describir la evolución científica, tanto en cantidad como en calidad, de las investigaciones en este campo.

Determinará los autores, instituciones y países más influyentes, así como las tendencias temáticas, metodológicas y emergentes que guiarán las futuras direcciones de investigación en este ámbito

Capítulo 1. Marco Teórico

Este apartado proporciona un fundamento teórico, centrado en los conceptos esenciales del comportamiento del consumidor y el *Big Data*. Examinan las teorías y los modelos pertinentes que facilitan la interpretación de grandes volúmenes de datos en relación con el comportamiento del consumidor y el *Big Data*.

1.1 Comportamiento del Consumidor

El estudio del comportamiento del consumidor se ha consolidado como un pilar fundamental para la comprensión de las dinámicas del mercado y la formulación de estrategias comerciales efectivas. Desde la perspectiva de la teoría económica clásica, los consumidores son considerados agentes racionales que toman decisiones de compra con el objetivo de maximizar su utilidad bajo restricciones presupuestarias. Esto implica que los consumidores toman decisiones basadas en la mejor combinación posible de bienes y servicios que maximicen su satisfacción dentro de las limitaciones de su ingreso. En el siglo XIX William Jevons propone la teoría de la utilidad, en donde sostiene que los consumidores asignan valores de utilidad a los bienes y servicios según sus preferencias, y buscan maximizar esta utilidad en sus elecciones de consumo (Varian 2014).

Posteriormente Alfred Marshall y Léon Walras a finales del siglo XIX y principios del XX, amplían la teoría de la utilidad al incorporar conceptos como las curvas de indiferencia y la restricción presupuestaria, enfatizando que los consumidores eligen combinaciones de bienes que les proporcionan la mayor utilidad posible dentro de sus limitaciones económicas. Este enfoque asume que los consumidores tienen información perfecta y pueden evaluar todas las opciones disponibles para maximizar su bienestar (Gómez y Sequería 2015).

Sin embargo, la economía conductual ha revelado limitaciones en esta visión neoclásica del consumidor racional. Herbert Simon introdujo el concepto de “racionalidad limitada”, sugiriendo que las decisiones de los consumidores están restringidas por la información disponible y las capacidades cognitivas, desafiando así la noción de un consumidor plenamente racional (Simon 1956). Esto implica que, en la práctica, los consumidores pueden no maximizar su utilidad debido a sesgos y limitaciones cognitivas.

En este contexto, Kahneman y Tversky (1973, 1979), pioneros en impulsar la Teoría Prospectiva, definen cómo las personas evalúan las pérdidas y ganancias de manera asimétrica, prefiriendo evitar pérdidas en lugar de obtener ganancias equivalentes. Este

fenómeno, conocido como aversión a las pérdidas, conduce a decisiones que desvían de la máxima utilidad esperada (Kahneman y Tversky 1979). La teoría prospectiva resalta que los consumidores no siempre actúan de manera racional, sino que sus decisiones pueden estar influenciadas por percepciones subjetivas del riesgo y la ganancia (Reisch y Zhao 2017).

El premio Nobel de Economía 2017, Richard Thaler (1985), amplió estos hallazgos con la teoría de la Contabilidad Mental, que describe cómo las personas toman decisiones subjetivas e irracionales, asignando sus recursos en diferentes “cuentas mentales” de forma emocional. Esto afecta sus decisiones de consumo y ahorro de maneras que a menudo contradicen la maximización de la utilidad, ya que los consumidores pueden gastar de manera diferente según la fuente del dinero o el propósito asignado mentalmente (Thaler 1985).

En la misma línea, Mick, Broniarczyk y Haidt (2004) sugieren que un exceso de opciones podría reducir el bienestar del consumidor en lugar de aumentarlo, esto se debe a que tener demasiadas opciones puede llevar a la ansiedad, la parálisis en la toma de decisiones y la insatisfacción con las elecciones realizadas, debido a la inseguridad sobre si se tomó la decisión "correcta". Mullainathan y Shafir (2013) discuten cómo las limitaciones en tiempo, conocimiento y poder de negociación pueden complicar la toma de decisiones para los consumidores. Como resultado, los consumidores adoptan reglas generales y atajos mentales para simplificar estas decisiones.

Solomon (2006), Mollá et al. (2012) amplían este enfoque al conceptualizar el comportamiento del consumidor como el conjunto de procesos a través de los cuales los individuos o grupos seleccionan, adquieren, usan o descartan productos y servicios disponibles en el mercado en un esfuerzo por satisfacer sus necesidades y deseos. Raiteri (2016) fortalece esta conceptualización al distinguir entre el "consumidor", aquel que adquiere un producto o servicio, el "cliente", que realiza compras habituales en un comercio específico; y el "usuario", que disfruta continuamente de un servicio o del uso de un producto.

Este enfoque es complementado por la perspectiva de Orellana (2022) y Recalde, Suárez y Páez (2021), quienes subrayan la importancia crítica de la toma de decisiones en la asignación de recursos como el tiempo, el dinero y el esfuerzo en la adquisición de bienes y servicios, destacando el viaje del consumidor desde la identificación de una necesidad hasta las etapas de post-compra.

Desde una perspectiva psicosocial, a principios del siglo XX, se comenzó a analizar cómo el entorno social y cultural moldea las actitudes y comportamientos de consumo, Thorstein Veblen en 1899, introdujo el concepto de consumo ostentoso y la emulación donde los individuos adquieren bienes y servicios no solo por su utilidad, sino para demostrar estatus social y obtener aprobación externa (Coronado 2019).

La teoría de la jerarquía de las necesidades de Maslow (1943), propone que las necesidades y motivaciones humanas estructuran el comportamiento de consumo en diferentes niveles, desde las más básicas, como las fisiológicas y de seguridad, hasta las más elevadas, como la autoestima y la autorrealización (Recalde, Suárez y Páez 2021).

Estos comportamientos en los consumidores pueden conectarse con conceptos económicos como la elasticidad de la demanda y la segmentación de mercados. Por ejemplo, Los bienes de lujo, suelen tener una elasticidad de demanda distinta a la de los bienes de primera necesidad. Comprender estas diferencias permite a las empresas ajustar sus estrategias de precios y marketing para dirigirse eficazmente a diferentes segmentos del mercado.

La teoría conductual del aprendizaje, discutidas por autores como Tapia, Alvarado y Cox (2019) y Callzilla (2023), inicialmente desarrollada por Pavlov en 1903, sugiere que los comportamientos de compra se adquieren y modifican a través de la experiencia. Este enfoque resalta el impacto de los estímulos externos y la publicidad en la formación de hábitos de consumo, lo cual es esencial para comprender cómo las empresas pueden influir en la demanda y el comportamiento del consumidor.

La teoría del condicionamiento operante, descrita por Reina y Rodríguez (2023), fue desarrollada por Edward Thorndike en 1911, B.F. Skinner en 1938, explica cómo las decisiones de compra pueden estar influenciadas por deseos y motivaciones inconscientes. Sostienen que los productos y servicios pueden actuar como símbolos de necesidades psicológicas subyacentes, como el estatus o la seguridad, llevando a los consumidores a tomar decisiones que trascienden la lógica económica racional.

Los elementos psicológicos, como la motivación, percepción, aprendizaje, actitudes, personalidad, valores y creencias influyen profundamente en el comportamiento del consumidor. Según Hervé y Mullet (2009), afectan desde cómo interpretan la información y modifican su comportamiento, hasta cómo se predisponen hacia productos y toman decisiones de compra basadas en sus características individuales y principios fundamentales.

Los factores sociales, incluyendo el impacto de la familia y los grupos de referencia, según Stávková, Stejskal y Toufarová (2008), moldean las preferencias y comportamientos de compra. La cultura, con sus normas y valores, junto con factores personales como la edad, el ingreso y el estilo de vida, influyen significativamente en las decisiones de compra (Muniady et al. 2014 y Rangaiah 2023). Estos factores se complementan con factores económicos, como el ingreso personal y la situación económica general, afectando la capacidad y la voluntad de los consumidores para gastar en productos y servicios.

La era digital ha dado origen al Consumidor 2.0, caracterizado por su interacción activa y constante con el contenido en línea, impulsada por la evolución de Internet hacia la Web 2.0. Como señalan Muñoz y Aguilar (2020), esta transformación ha fomentado un entorno de cooperación e interactividad, donde los consumidores consumen contenido, y también participan activamente en su creación y difusión.

En este entorno, conceptos económicos como la asimetría de información y el costo de búsqueda adquieren una nueva dimensión. Álvarez (2018) expone que, el *Big Data* permite reducir la asimetría de información al proporcionar a las empresas datos detallados sobre las preferencias y comportamientos de los consumidores. Esto facilita la personalización de ofertas y estrategias de marketing más efectivas. Además, los consumidores pueden acceder más fácilmente a información sobre productos y servicios, reduciendo sus costos de búsqueda. Aparecen figuras como el "prosumer" y el "influencer", que juegan roles indispensables en la dinámica de consumo actual, afectando las decisiones de compra de otros consumidores a través de sus recomendaciones y contenido generado (Macías 2016; Sánchez, Zamora y Cuéllar 2015; Ambrosio 2013). La capacidad de estos actores para influir en las percepciones y decisiones de un amplio público destaca la importancia de las estrategias de marketing digital y la gestión de la reputación en línea.

La tecnología digital y la comunicación en línea han rediseñado las costumbres y expectativas de estos consumidores, quienes utilizan las herramientas y recursos en línea para mejorar su experiencia de compra y personalizar su interacción con las marcas (Raiteri 2016; Ferrer 2018). Este cambio ha llevado a las empresas a repensar sus estrategias de marketing y comunicación, adoptando enfoques que fomentan la participación del consumidor y empleando metodologías como el crowdsourcing para adaptar y mejorar sus ofertas en función de las aportaciones de los usuarios (Montero y El Kadi 2017).

Esta "huella digital", al componerse de una gran masa de datos, se convierte en una herramienta estratégica para que las empresas comprendan mejor a su público objetivo y desarrollen ofertas que respondan de manera más eficaz y personalizada a sus necesidades y expectativas (UNESCO 2020). En este contexto, el análisis de *Big Data* adquiere una gran importancia, la interacción constante de los consumidores 2.0 con las plataformas digitales genera enormes volúmenes de datos que, correctamente analizados, proporcionan insights valiosos sobre sus comportamientos, preferencias y tendencias de consumo. Esto puede aumentar el bienestar del consumidor al proporcionar productos y servicios más alineados con sus preferencias reales y la eficiencia de las empresas en el mercado

1.2 *Big Data*

Según Kusnetzky (2010), el término *Big Data* hace referencia a un extenso y complejo conjunto de datos que desafía los métodos tradicionales de procesamiento y análisis, es decir, trasciende la capacidad de los procesos analíticos convencionales debido a su volumen, velocidad y variedad. De manera similar, Dans (2012) enfatiza que el *Big Data* implica el manejo y análisis de repositorios de datos tan vastos que sobrepasan las capacidades de las herramientas analíticas y de bases de datos tradicionales. Camargo et al. (2015), amplían esta definición al considerar que *Big Data* no solo abarca los datos en sí, sino también las herramientas, procesos y procedimientos que permiten a las organizaciones gestionar y manipular estos grandes conjuntos de datos.

Al “Modelo de las tres V”, de volumen, velocidad y variedad introducido por Laney (2001) para caracterizar el *Big Data* se han sumado otras dimensiones como la veracidad, que asegura la fiabilidad de los datos para la toma de decisiones, y el valor, que se refiere a la conversión de datos en conocimiento aplicable (Hernandez y Duque 2017). Sun et al. 2018 y Chavez et. al. 2022, introducen conceptos adicionales como la versatilidad, vulnerabilidad y volatilidad, ampliando aún más el espectro de características asociadas al *Big Data*.

Desde una perspectiva económica, el *Big Data* permite a las empresas optimizar recursos y aprovechar las economías de escala en el procesamiento y análisis de grandes volúmenes de datos. Al manejar eficientemente vastas cantidades de información, las organizaciones pueden reducir costos operativos y mejorar la eficiencia en la toma de decisiones (Erevelles, Fukawa y Swayne 2016). Las economías de escala se manifiestan en la capacidad de procesar más datos sin un aumento proporcional en los costos, lo que resulta en una ventaja competitiva en mercados altamente dinámicos (Álvarez 2018).

Desde la perspectiva estructural, el *Big Data* se categoriza en datos estructurados, semi-estructurados y no estructurados, cada uno con sus propias implicaciones para el almacenamiento y análisis (Yu y Yang 2016). El procesamiento de estos datos implica diversas fases, comenzando por la adquisición de información de múltiples fuentes, seguido por el almacenamiento, que desafía las capacidades de los sistemas de bases de datos tradicionales debido a su magnitud, y finalmente la analítica, que busca interpretar y aplicar los datos de manera coherente y fiable (Paredes 2015).

La analítica de *Big Data* va más allá de la recopilación de datos, utilizando análisis matemáticos y estadísticos para identificar patrones no visibles con métodos tradicionales y predecir comportamientos futuros. Este proceso, conocido también como el descubrimiento de conocimiento en datos, implica la utilización de modelos algorítmicos y técnicas de aprendizaje automático para generar información procesable a partir de grandes volúmenes de datos (Páez y Monroy 2020). Según Álvarez (2018), esta capacidad predictiva es esencial, ya que permite a las empresas identificar patrones no visibles con métodos tradicionales, anticipar comportamientos del mercado y de los consumidores, y así personalizar estrategias de manera precisa y eficiente.

En el análisis del comportamiento del consumidor, el *Big Data* facilita la extracción de información sobre las preferencias y comportamientos de los consumidores, permitiendo a las organizaciones adaptar sus estrategias de manera más efectiva, tal como lo destacan Erevelles, Fukawa y Swayne (2016). Hofacker et al. (2016), inspirándose en el modelo de toma de decisiones de Blackwell, Miniard y Engel (2005), adaptan este proceso al consumidor moderno de la era digital, destacando etapas como el reconocimiento del problema, la búsqueda y evaluación de alternativas y el compromiso de compra, reflejando así la evolución hacia el consumidor 2.0. Con el análisis del *Big Data*, las empresas pueden personalizar ofertas y comunicaciones, optimizando la experiencia del consumidor y fomentando la lealtad, lo que se traduce en ventajas competitivas significativas.

La creciente importancia del *Big Data* en la economía y el comportamiento del consumidor plantea varias preocupaciones éticas, principalmente relacionadas con la privacidad y el uso responsable de la información. Colmenarejo (2018) argumenta que el manejo masivo de datos no es solo un fenómeno tecnológico, sino también social, que transforma la comunicación y las relaciones sociales. Esto lleva a conflictos éticos en áreas como la privacidad, la propiedad y la identidad. Por su parte, Aragona (2022) destaca que uno de los aspectos fundamentales de estos problemas éticos es principalmente la relacionada a la calidad de los datos y el acceso a

ellos. Además, menciona la necesidad de desarrollar pautas éticas adaptadas al contexto, ya que no pueden ser universales ni establecidas de una vez por todas. El consentimiento informado y el tratamiento no intrusivo de los datos digitales son pilares centrales en este debate, pues las técnicas de investigación encubiertas y el acceso a datos personales desafían las prácticas éticas tradicionales. Por ello, es fundamental abordar la responsabilidad y las consecuencias de la recopilación y el uso de datos personales en las prácticas comerciales y de marketing (Colmenarejo 2018; Aragona 2022).

Desde entonces, numerosos autores han explorado estos dos temas desde diversas perspectivas, evidenciando la necesidad de emplear metodologías analíticas complementarias para profundizar en la comprensión del comportamiento del consumidor a través del *Big Data*. En este contexto, el análisis bibliométrico se constituye como una herramienta esencial que, al integrarse con técnicas de análisis de *Big Data*, permite identificar y analizar de manera más eficiente las tendencias y patrones en la literatura especializada. El análisis bibliométrico facilita la detección de las principales líneas de investigación en *Big Data* y comportamiento del consumidor, mapeando la evolución de conceptos clave y revelando vacíos en el conocimiento (Zupic y Cater 2015).

La sinergia entre el *Big Data* y el análisis bibliométrico amplía la capacidad para manejar y examinar datos bibliográficos masivos, posibilitando una comprensión más profunda y matizada de las dinámicas del comportamiento del consumidor y las tendencias emergentes en la investigación académica (Cobo et al. 2011).

En el siguiente capítulo se desarrollarán los principales argumentos relacionados con el análisis bibliométrico.

Capítulo 2. Metodología

Este capítulo detalla la metodología adoptada para realizar un análisis bibliométrico sobre impacto del *Big Data* en el comportamiento del consumidor, utilizando la base de datos de Scopus, con el objetivo de identificar, analizar y visualizar las principales tendencias, patrones y redes de colaboración en este campo de estudio a lo largo de una década.

La bibliometría, es definida por Broadus (1987) y Jiménez (1993) como un análisis cuantitativo de la producción científica mediante métodos matemáticos y estadísticos. Incluye el análisis de las contribuciones de autores individuales, instituciones y países a sus áreas de estudio, así como el impacto de sus publicaciones en el avance del conocimiento (Peralta, Frías y Gregorio 2015). En el contexto de esta investigación, la bibliometría permite cuantificar y evaluar cómo ha evolucionado el conocimiento sobre *Big Data* y comportamiento del consumidor, lo cual es esencial para comprender las dinámicas actuales y anticipar futuros direccionamientos investigativos en estos campos.

El análisis bibliométrico ofrece una panorámica integral de la literatura científica, facilitando la comprensión del desarrollo histórico y actual de un campo, así como la identificación de tendencias futuras. Es indispensable seleccionar técnicas bibliométricas alineadas con los objetivos del estudio desde la fase inicial de recopilación de datos con el fin de efectivizar tiempo en el análisis y discusión de resultados. Estas técnicas se dividen en análisis de desempeño, que mide la contribución de entidades como autores, instituciones y países al conocimiento de un campo, y mapeo científico, que explora las relaciones y temas emergentes dentro de una disciplina (Bienert et al. 2015). Ambas técnicas son esenciales para esta investigación, ya que el análisis de desempeño ayuda a identificar las contribuciones más significativas, y el mapeo científico permite entender cómo se relacionan y evolucionan las temáticas clave.

Según Aria y Corrado (2017), las técnicas bibliométricas, como el análisis de citas, co-citas, acoplamiento bibliográfico y análisis de co-palabras, revelan la influencia, estructura intelectual, conexiones y relaciones temáticas dentro de una disciplina, ofreciendo una visión actualizada de los temas y tendencias predominantes. En este estudio, cada técnica bibliométrica se emplea para obtener una comprensión detallada de diferentes aspectos del campo.

El análisis de citas permite discernir la relevancia y el impacto de investigaciones específicas, autores y revistas, identificando aquellos trabajos que han moldeado significativamente el

entendimiento del comportamiento del consumidor en la era del *Big Data* (Garfield 2004). A través de esta técnica, es posible determinar cuáles son los estudios y autores más influyentes en el campo, lo que ayuda a comprender las bases teóricas y prácticas que sustentan el área de estudio.

El análisis de co-citas facilita el desglose de la estructura intelectual del campo, revelando cómo distintas investigaciones se interrelacionan y conforman subdominios específicos. Esto es fundamental para comprender las corrientes de pensamiento predominantes y las conexiones conceptuales entre estudios sobre consumidores y análisis de datos masivos (Donthu et al. 2021; Small 1973). Al identificar grupos de artículos citados conjuntamente, se pueden mapear las redes intelectuales y entender cómo se han desarrollado las líneas de investigación en este ámbito.

El acoplamiento bibliográfico, al enfocarse en las referencias compartidas, destaca las similitudes temáticas y arroja luz sobre las tendencias emergentes y las perspectivas más recientes en la investigación del comportamiento del consumidor en el contexto del *Big Data* (Gan et al. 2022). Esta técnica permite identificar cómo las investigaciones recientes se relacionan entre sí a través de las fuentes que citan, lo cual es útil para detectar nuevas áreas de interés y posibles colaboraciones.

El análisis de co-palabras, al examinar la co-ocurrencia de términos clave en la literatura relevante, ayuda a mapear el terreno conceptual del estudio del comportamiento del consumidor en relación con el *Big Data*, identificando nuevos ejes de investigación y posibilitando la predicción de futuras áreas de interés (Chang, Wang y Hawamdeh 2015). Esto permite entender las tendencias temáticas y cómo evolucionan los tópicos de investigación a lo largo del tiempo.

Estas metodologías bibliométricas, al ser aplicadas conjuntamente, proporcionan insights valiosos para investigadores y profesionales interesados en las intersecciones entre el *Big Data* y el comportamiento del consumidor, facilitando la comprensión de las tendencias actuales y futuras en estos campos interconectados.

Para la evaluación de la producción científica por medio de indicadores bibliométricos, Bordons y Zuleata (1999) enfatizan la importancia de elegir una base de datos con un reconocido prestigio académico y un volumen significativo de publicaciones relevantes. Según Cadavid et al. (2012) las tres bases de datos de publicaciones científicas más importantes son Scopus, Web of Science y Scholar Google.

Se eligió la base de datos de Scopus, ya que es reconocida por ser la base de datos más grande de resúmenes y citas de literatura revisada por pares, lo que incluye una amplia gama de disciplinas científicas (Burnham 2006); Además, cuenta con una mayor cobertura de revistas en ciencias sociales y humanidades en comparación con WoS, lo que es particularmente relevante para estudios sobre comportamiento del consumidor (Mongeon y Paul-Hus 2016); Ofrece ventajas significativas como su interfaz, herramientas de filtrado y análisis, y acceso a documentos revisados por pares, lo que facilita la recopilación y depuración de datos (Granda et al. 2013); Asimismo, incluye indicadores bibliométricos como el índice h, que permiten evaluar el impacto y la calidad de las investigaciones y autores (Falagas et al. 2008).

Es importante considerar las metodologías previamente implementadas en estudios cercanos para esclarecer las tendencias y enfoques predominantes en la investigación sobre *Big Data* y comportamiento del consumidor. La revisión de estas metodologías facilita comprensión de las prácticas actuales, y también la identificación de posibles áreas de mejora y oportunidades para la innovación en la metodología propuesta. En este sentido, se presenta a continuación una síntesis de los estudios bibliométricos más relevantes desde 2020, seleccionados por su relevancia y aplicabilidad al contexto de esta investigación.

Tabla 2. 1. Análisis Bibliométricos relacionados al *Big Data* y el Comportamiento del Consumidor

Título de la Investigación	Autores	Período de Análisis	Documentos y Base de Datos	Palabras Clave	Metodología	Principales Resultados
<i>Artificial Intelligence and Behavioral Economics: A Bibliographic Analysis of Research Field</i>	(Aoujil et al. 2023)	2012-2022	637 documentos (Web of Science)	<i>Behavioral economics, artificial intelligence, machine learning, neural networks, deep learning</i>	Emplearon el paquete Bibliometrix en R y el software VOSviewer para análisis de citas, coautoría, acoplamiento bibliográfico y co-ocurrencia de palabras clave. Identificaron autores, instituciones influyentes y redes de colaboración internacional.	Identificaron un aumento en publicaciones sobre la aplicación de IA en economía del comportamiento. Los autores más citados incluyen a Kahneman y Thaler. Las instituciones líderes son MIT y Harvard. Se destacó el uso de redes neuronales y aprendizaje profundo para modelar decisiones bajo incertidumbre.
<i>Green Consumer Research: Trends and Way Forward Based on Bibliometric Analysis</i>	(Haba, Bredillet y Dastane 2023)	1993-2022	493 documentos (Scopus)	<i>Green Consumer, Green Consumption, Sustainable Consumption</i>	Realizaron análisis de red de palabras clave, coautoría y co-citación utilizando VOSviewer. Aplicaron SciMAT para análisis de mapeo científico longitudinal, identificando la evolución temática y el mapa estratégico del campo.	China, EE.UU. e India lideran en investigaciones sobre consumo verde. Temas emergentes incluyen energías renovables y políticas ambientales. Li Y. es el autor más influyente, mientras que la Florida International University es la institución con más publicaciones. Los productos verdes y el marketing verde son áreas de mayor impacto.
<i>Two Decades of M-Commerce Consumer</i>	(Thangavel y Chandra 2023)	2001-2022	467 documentos (Web of Science)	<i>M-Commerce, Mobile Commerce, Consumer</i>	Utilizaron el software Bibliometrix de RStudio para realizar análisis bibliométricos completos, incluyendo técnicas	Publicaciones sobre <i>m-commerce</i> crecieron significativamente post-2010. Keng-Boon Ooi es el autor más citado. Observan un

<i>Research: A Bibliometric Analysis Using R Biblioshiny</i>				<i>Behavior, Mobile Shopping, Bibliometric Analysis, Literature Review</i>	estadísticas y visualizaciones avanzadas.	cambio de enfoque del Modelo de Aceptación Tecnológica a la Teoría Unificada de Aceptación y Uso de la Tecnología, reflejando la evolución en la confianza y lealtad del usuario en el comercio móvil. COVID-19 aceleró la adopción del <i>m-commerce</i> , destacando el comercio social móvil y <i>fintech</i> como áreas emergentes.
<i>Artificial Intelligence in E-Commerce: A Bibliometric Study and Literature Review</i>	(Bawack et al. 2022)	1975-2020	4335 documentos (Web of Science)	<i>Electronic Commerce, e-business, Internet Commerce, Artificial Intelligence, y varios términos relacionados con técnicas y conceptos de IA</i>	Evaluaron impacto de autores, revistas e instituciones usando R Bibliometrix. Utilizaron análisis de co-citas y espectroscopía de año de publicación de referencias (RPYS) para identificar contribuciones clave.	El análisis de sentimientos, sistemas de recomendación y personalización son los temas más investigados. Las instituciones en China lideran la producción de investigaciones, mientras que EE.UU. domina en términos de citas. Proponen que futuras investigaciones se centren en la integración de IA en sistemas de información para mejorar las experiencias de los consumidores en <i>e-commerce</i> .
<i>Evolutionary Overview of Consumer Health Informatics: A Bibliometric Study on the Web of Science</i>	(Ouyang et al. 2020)	1999-2019	986 documentos (Web of Science core collection database)	<i>Consumer Health Informatics, Consumer Health Information, Consumer Media * Informatics,</i>	Utilizaron SciMAT para análisis de co-ocurrencia y clúster, construyendo mapas de conocimiento y evolución temática. Clasificaron temas en cuadrantes basados en importancia y madurez.	Se identificaron temas clave como la educación del paciente y la búsqueda de información de salud como áreas centrales y bien desarrolladas. El análisis muestra que <i>eHealth</i> y la alfabetización en salud digital son campos en rápido crecimiento. La Universidad de Toronto y la

<i>from 1999 to 2019</i>				<i>Consumer Media * Information</i>		Universidad de California son las instituciones más destacadas en este campo.
<i>The Usage of Large Data Sets in Online Consumer Behaviour: A Bibliometric and Computational Text-Mining–Driven Analysis of Previous Research</i>	(Vanhala et al. 2020)	2000-2018	495 documentos (ISI Web of Science)	<i>Big Data, Clickstream, Text Data, Path Analysis, Customer Analytics, Customer, Consumer, Purchase, Buying, Online Discussion, Information Search, Customer Journey, Online y Mobile.</i>	Utilizaron <i>Network Analysis Interface for Literature Studies</i> (NAILS) para análisis bibliométrico y <i>Structural Topic Modeling</i> (STM) para descubrir temas subyacentes en los documentos.	Identificaron un notable incremento en investigaciones desde 2013 sobre <i>Big Data</i> y comportamiento del consumidor. Los temas clave incluyen experiencia de compra, privacidad en comunidades digitales y retos en mercados minoristas. Los autores más influyentes son Van Doorn y Verhoef. Se sugiere una mayor integración de teorías de comportamiento con análisis de <i>Big Data</i> para entender mejor las motivaciones del consumidor.

Elaborado por el Autor a partir de datos obtenidos de Google Académico (2024).

En el análisis bibliométrico del *Big Data* y el comportamiento del consumidor, diversas metodologías han sido empleadas para entender las tendencias y enfoques predominantes en este campo. Aoujil et al. (2023) utilizaron el paquete Bibliometrix en R y VOSviewer para realizar un análisis de citas, coautoría, acoplamiento bibliográfico y co-ocurrencia de palabras clave, identificando un aumento constante en publicaciones sobre economía conductual y la aplicación de inteligencia artificial en la personalización de servicios. De manera similar, Haba, Bredillet y Dastane (2023) aplicaron SciMAT y VOSviewer para realizar un análisis de red de palabras clave y co-citación, revelando un crecimiento en la investigación sobre consumo verde y la formación de redes académicas en sostenibilidad. Estos enfoques metodológicos permiten mapear el desarrollo del campo y destacar las áreas de colaboración y avance científico.

Otros estudios, como el de Thangavel y Chandra (2023), emplearon el software Bibliometrix de RStudio para analizar el m-commerce, observando un crecimiento significativo en la investigación desde 2001, influenciada recientemente por la pandemia de COVID-19. Bawack et al. (2022) utilizaron R Bibliometrix y técnicas de co-citación para evaluar la investigación en inteligencia artificial en comercio electrónico, identificando sistemas de recomendación y personalización como temas dominantes. Ouyang et al. (2020) aplicaron SciMAT para analizar la evolución de la informática de la salud del consumidor, destacando la importancia de la educación del paciente y la tecnología de salud. Finalmente, Vanhalaa et al. (2020) usaron NAILS y STM para modelar temas en la investigación de comportamiento del consumidor en línea, subrayando la relevancia de la privacidad digital. Estos estudios proporcionan un marco comprensivo de metodologías que informan y justifican las propuestas metodológicas de este análisis bibliométrico, estableciendo una base sólida para la aplicación de técnicas avanzadas en el estudio del *Big Data* y el comportamiento del consumidor.

Siguiendo estos estudios, esta investigación adopta una metodología cuantitativa, descriptiva y bibliométrica. Se utiliza la base de datos Scopus, actualizada hasta el 26 de mayo de 2024, para realizar un análisis exhaustivo del impacto del *Big Data* en el comportamiento del consumidor desde 1995, año de la primera publicación relevante, hasta 2023. Este enfoque se caracteriza por la utilización de herramientas avanzadas de análisis bibliométrico, proporcionadas por Bibliometrix en RStudio y Tree of Science. Estas herramientas facilitan la evaluación de la producción académica y las citas, el análisis de redes de colaboración y

también una organización estructurada de la literatura existente, permitiendo una comprensión profunda y ordenada del campo de estudio.

La recopilación de datos se estructuró siguiendo el protocolo PICOC (*Population, Intervention, Comparison, Outcome, Context*) para garantizar una estrategia de búsqueda integral y sistemática (Khan et al. 2003). La estrategia de búsqueda se diseñó de la siguiente manera:

- Población (P): Estudios relacionados con *Big Data* en el contexto del comportamiento del consumidor.
- Intervención (I): Uso y aplicación de técnicas de *Big Data*, análisis de datos y minería de datos.
- Comparación (C): No se establece una comparación directa, pero se consideran diferentes enfoques y metodologías dentro del campo.
- Resultado (O): Impacto y comprensión mejorada del comportamiento del consumidor a través del uso de *Big Data*.
- Contexto (C): Publicaciones académicas indexadas en Scopus entre 1995 y 2023.

Se incorporaron sinónimos y variantes ortográficas en la ecuación de búsqueda para ampliar la cobertura y asegurar la inclusión de todos los documentos relevantes. Por ejemplo, se utilizaron términos como "*Big Data*", "*Data Analytics*" y "*Data Mining*" para abarcar diferentes enfoques relacionados con el análisis de grandes volúmenes de datos (Chen et al. 2014; Witten et al. 2016). Asimismo, se incluyeron variantes de "*Consumer Behavior*" y "*Customer Behavior*", considerando tanto la ortografía americana como británica, para abarcar publicaciones internacionales.

La ecuación de búsqueda aplicada en Scopus fue:

```
TITLE-ABS-KEY ( ( "Big Data" OR "Data Analytics" OR "Data Mining" ) AND ( "Consumer Behavior" OR "Consumer Behaviour" OR "Customer Behavior" OR "Customer Behaviour" ) ) AND ( EXCLUDE ( PUBYEAR , 2025 ) OR EXCLUDE ( PUBYEAR , 2024 ) ) AND ( LIMIT-TO ( LANGUAGE , "English" ) )
```

Se establecieron criterios de inclusión para garantizar la relevancia y la calidad de los documentos seleccionados. Uno de estos criterios fue limitar la búsqueda a documentos en inglés, dado que este es el idioma universal de la escritura académica, lo que facilita la difusión y comprensión de la investigación a nivel internacional, asegurando la recopilación de datos pertinentes para una evaluación detallada.

Es importante reconocer que los estudios bibliométricos pueden estar sujetos a diversos sesgos que podrían afectar la validez y generalización de los resultados obtenidos. Uno de los principales sesgos proviene de la selección de la base de datos; al utilizar exclusivamente Scopus, se limita la cobertura de la literatura relevante, ya que ninguna base de datos es completamente exhaustiva (Mongeon y Paul-Hus 2016). Este enfoque puede excluir publicaciones importantes presentes en otras bases de datos como Web of Science o Google Scholar, introduciendo un sesgo de selección que puede influir en los hallazgos y conclusiones del estudio (Falagas et al. 2008). Además, al restringir la búsqueda a documentos en inglés, se introduce un sesgo lingüístico que excluye investigaciones significativas publicadas en otros idiomas, afectando la representatividad global del análisis y favoreciendo una perspectiva anglocéntrica de la literatura científica (van Leeuwen et al. 2001).

Otro aspecto para considerar es el sesgo de citación, donde ciertos artículos o autores reciben una cantidad desproporcionada de citas debido a su visibilidad, el prestigio de la revista o prácticas de autocitación, lo que puede influir en la percepción de su influencia real dentro del campo (MacRoberts y MacRoberts 2018). Asimismo, existe el sesgo institucional o de país, ya que instituciones o países con mayores recursos tienden a producir más investigaciones que son más citadas, creando una visión desequilibrada del desarrollo científico en el área (Aksnes, Piro y Rorstad 2019). Reconocer y abordar estos sesgos es fundamental para garantizar que el análisis bibliométrico proporcione una visión precisa y equilibrada del estado de la investigación en *Big Data* y comportamiento del consumidor.

El análisis descriptivo de los datos proporcionó una visión global de la producción científica, destacando el número total de publicaciones, su distribución anual y las principales revistas y países contribuyentes. Esta etapa inicial fue indispensable para comprender la evolución del campo de estudio. Posteriormente, se realizó un análisis bibliométrico detallado de 1861 documentos seleccionados, utilizando el paquete estadístico Bibliometrix en R Studio. Este software facilitó una gestión eficiente y meticulosa de los datos bibliométricos, esencial para la interpretación y el manejo estadístico de la información. Este análisis se alineó con las metodologías propuestas por Malpica et al. (2021) y Saleem et al. (2024), quienes destacan la utilidad de Bibliometrix en R para una exploración exhaustiva de las tendencias bibliográficas y los patrones de publicación, utilizando metadatos bibliométricos completos y comparables, con el objetivo de asegurar la máxima significancia y precisión en los resultados.

Adicionalmente, la herramienta Tree of Science (TOS) añadió un valor significativo al análisis al proporcionar una visualización estructurada de la jerarquía de las investigaciones. Su algoritmo permite mapear y visualizar las conexiones entre estudios clave, identificando los trabajos fundamentales que forman la base y los nodos más influyentes dentro del campo de estudio. Esta capacidad de discernir la estructura jerárquica y la interconexión de las publicaciones facilita una comprensión más profunda de las tendencias y los avances dentro del ámbito investigado (Valencia et al. 2020).

Esta metodología ofrece una perspectiva clara de cómo los estudios fundamentales han configurado el campo, y cómo los trabajos subsiguientes han expandido y diversificado el área de estudio. Este enfoque destaca los documentos pioneros que han establecido las bases teóricas y metodológicas, que introducen innovaciones y exploran nuevas direcciones, subrayando así las tendencias emergentes y delineando áreas potenciales para futuras investigaciones (Zuluaga et al. 2022).

De acuerdo con Core of Science (2020), TOS utiliza un algoritmo basado en métricas de teoría de grafos para analizar y organizar la evolución del conocimiento científico en forma de árbol en 4 categorías: raíces (ROOT) que incluye los 10 trabajos seminales más citados en un campo determinado, y que en el algoritmo se identifican como vértices en el grafo con grado exterior cero. Esto implica que son los puntos de origen de otros trabajos, pero no hacen referencia a ningún otro trabajo fundacional. El tronco (TRUNK) formado por artículos que se derivan de las raíces y sirven de paso hacia nuevas vías de investigación. En el algoritmo se determinan a partir de los caminos más cortos entre hojas y raíces buscando los vértices que se encuentran en el camino. Las ramas (BRANCHES) que agrupa los artículos en distintos subcampos o áreas especializadas de investigación. Finalmente, las hojas (LEAVES) recogen las investigaciones de los últimos 5 años y de mayor impacto en dicho campo de estudio. El algoritmo ordena las hojas como vértices que no han sido citados por otros artículos en función de su antigüedad (Valencia et al. 2020; Zuluaga et al. 2022)

Capítulo 3. Resultados

Al evaluar la completitud de los metadatos de los 1.861 documentos extraídos de Scopus, la herramienta Bibliometrix reveló una variabilidad significativa en la calidad de los distintos tipos de metadatos, los cuales se clasifican en categorías que van desde "Excelente" hasta "Completamente Ausente". Según esta herramienta, los metadatos asociados con "Abstract", "Document Type", "Journal", "Language", "Publication Year", "Title", y "Total Citation" demostraron una completitud total, siendo categorizados como "Excelente". Además, con menos del 5% de datos perdidos, los metadatos de "Author", "Affiliation", y "Cited References" fueron clasificados como "Buenos", y los de "Keywords" como "Aceptables".

Por otro lado, metadatos como "Keywords Plus" y "Corresponding Author" fueron categorizados como "Pobres" debido a altas tasas de ausencia del 26,8% y 39,1%, respectivamente. Estas deficiencias indican áreas donde la información es insuficiente, lo que podría limitar la profundidad de los análisis temáticos y de liderazgo dentro del campo.

Para asegurar la mayor significancia y precisión en el análisis bibliométrico, y dado el objetivo de identificar las tendencias y el impacto del *Big Data* en el comportamiento del consumidor, se ha decidido centrar el estudio en aquellos metadatos clasificados como "Excelente", "Buenos" y "Aceptables". La adopción de esta estrategia metodológica permite garantizar que los análisis realizados sean fundamentados, representativos y precisos, basándose en la información más completa y confiable disponible en la base de datos de Scopus.

La caracterización de la base de datos Scopus abarca el periodo desde 1995 hasta 2023, durante el cual se recopilaron 1861 documentos provenientes de 1.109 fuentes, incluyendo revistas, libros, capítulos de libros y actas de conferencias. La tasa de crecimiento anual del 21,51% refleja un aumento constante en la producción de investigaciones, lo que indica un interés continuo y sostenido en el estudio de la aplicación del *Big Data* en el comportamiento del consumidor. El promedio de 16,5 citas por documento sugiere que los estudios en este campo tienen un impacto considerable en la comunidad científica, subrayando la relevancia del tema en diversas disciplinas.

Un total de 4458 autores han contribuido a los 1861 documentos, con 237 de estos siendo de autoría única. Esto indica que, aunque una parte significativa de los trabajos son colaborativos, existe una presencia importante de trabajos individuales. El 18,38% de los documentos resultan de colaboraciones internacionales, lo que destaca la cooperación entre

diferentes países en el estudio del *Big Data* y el comportamiento del consumidor, aunque con una ligera disminución respecto a periodos anteriores.

Los principales tipos de documentos incluyen artículos de revistas (754) y actas de conferencias (897), confirmando que estos formatos son los canales más utilizados para la difusión de hallazgos en este campo. También se registraron 33 libros y 120 capítulos de libros, lo que sugiere que el conocimiento en esta área está siendo abordado en textos más extensos que permiten una profundización en el análisis y la comprensión del tema (ver Tabla 3.1).

Tabla 3. 1. Información Principal Sobre la Base de Datos de Scopus

	Description	Results
Main Information About Data	<i>Timespan</i>	1995:2023
	<i>Sources (Journals, Books, etc)</i>	1109
	<i>Documents</i>	1861
	<i>Annual Growth Rate %</i>	21.51
	<i>Document Average Age</i>	6.55
	<i>Average citations per doc</i>	16.5
	<i>References</i>	54643
Document Contents	<i>Keywords Plus (ID)</i>	7647
	<i>Author's Keywords (DE)</i>	4193
Authors	<i>Authors</i>	4458
	<i>Authors of single-authored docs</i>	237
Authors Collaboration	<i>Single-authored docs</i>	250
	<i>Co-Authors per Doc</i>	3.01
	<i>International co-authorships %</i>	18.38
Document Types	<i>article</i>	754
	<i>book</i>	33
	<i>book chapter</i>	120
	<i>conference paper</i>	897
	<i>conference review</i>	29
	<i>data paper</i>	1
	<i>retracted</i>	4
	<i>review</i>	22
	<i>short survey</i>	1

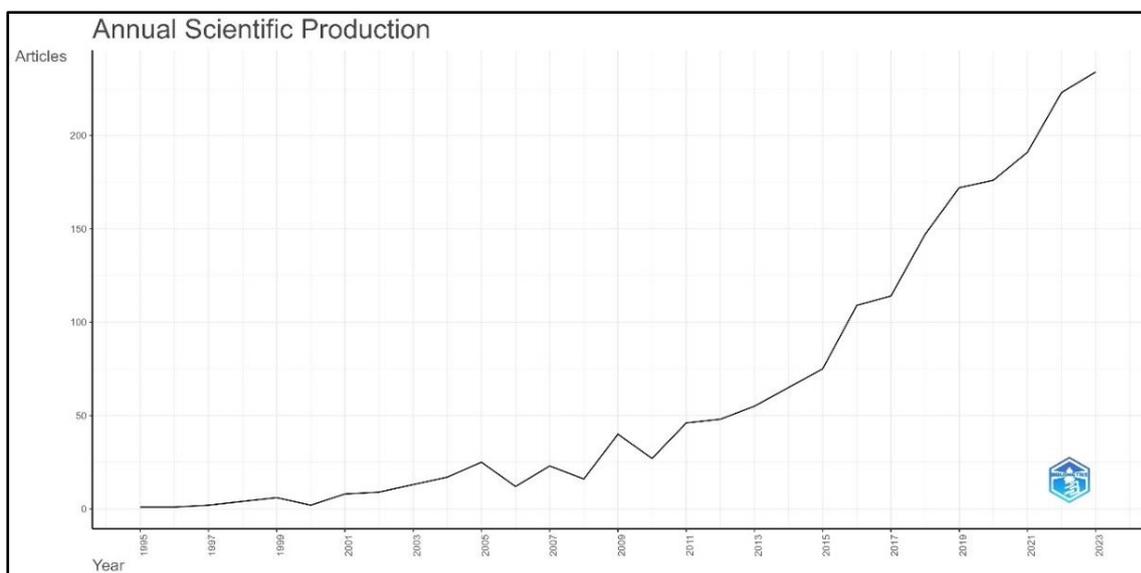
Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

En el Gráfico 3.1, referente a la producción científica anual, se observa la evolución de las publicaciones sobre *Big Data* y comportamiento del consumidor desde 1995 hasta 2023. En los primeros años, el número de publicaciones fue mínimo, con una tendencia estable hasta 2008. A partir de 2009, la producción comenzó a incrementarse de manera constante, alcanzando picos significativos en los últimos años.

El crecimiento notable se evidencia entre 2012 y 2022, con un claro aumento en la cantidad de artículos publicados. En 2022, se alcanzó un máximo de más de 200 artículos. Aunque los datos de 2023 aún no están completos, el gráfico muestra una tendencia de continuidad en la producción científica. Esta consistencia refleja la consolidación del área de investigación en *Big Data* y comportamiento del consumidor como un tema de creciente interés académico.

La producción anual sostenida y en aumento, especialmente en la última década, refuerza la idea de que esta área de investigación ha madurado y continúa en expansión dentro de la comunidad científica, con un crecimiento en la publicación de artículos cada vez más relevante en el análisis bibliométrico del campo.

Gráfico 3. 1. Producción Científica Anual en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

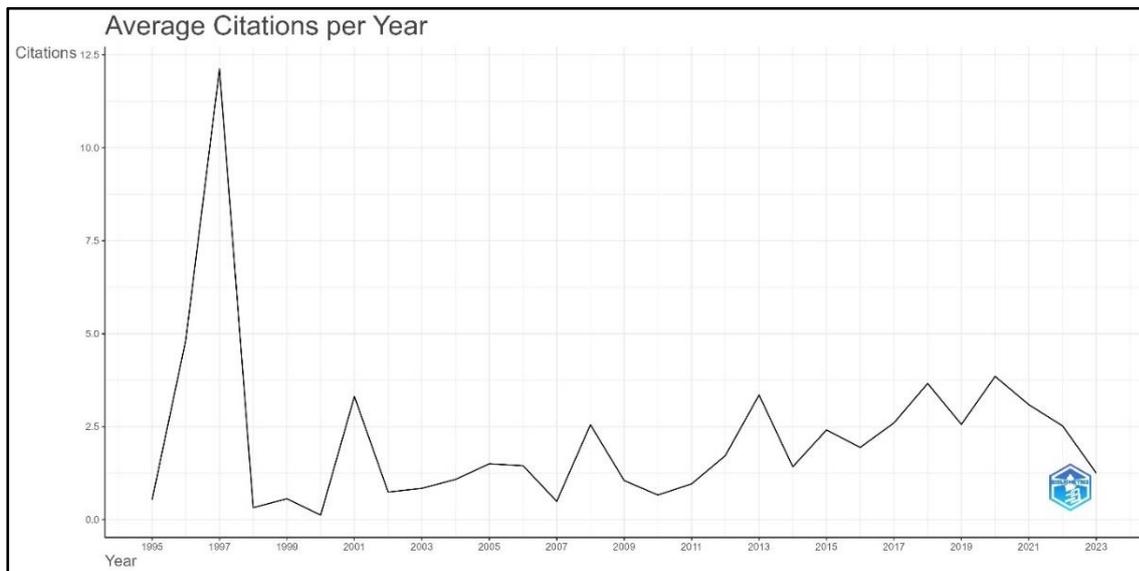
En el Gráfico 3.2, la evolución de las citas promedio por año ilustra cómo ha variado el impacto de los estudios sobre *Big Data* y comportamiento del consumidor a lo largo del

tiempo, desde 1995 hasta 2023. Al inicio del periodo, se observa un pico notable en 1997, donde las citas promedio por artículo alcanzaron más de 12, indicando que los pocos estudios publicados en esos años tuvieron un impacto significativo en la comunidad académica.

Después de este pico inicial, las citas promedio experimentaron una disminución, estabilizándose entre 1 y 3 citas por artículo durante la primera década de los 2000. Esta tendencia refleja una normalización en el impacto a medida que el volumen de publicaciones aumentaba y el campo se expandía.

Desde 2010 hasta 2020, las citas promedio muestran fluctuaciones menores, manteniéndose generalmente por debajo de las 2,5 citas por artículo. El gráfico indica un ligero incremento hacia 2018, seguido de un descenso gradual hacia 2023. La reducción reciente puede atribuirse al corto periodo de tiempo disponible para que los artículos más recientes acumulen citas.

Gráfico 3. 2. Citas promedio por año en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024)

Nota: La tabla de datos de este gráfico se adjunta en el Anexo No. 4.1.

La Tabla 3.2 presenta a los autores más relevantes en la investigación de *Big Data* y comportamiento del consumidor. Encabezando la lista se encuentra Wang Y., con 21 publicaciones y un impacto fraccionado de 7,16, lo que indica una alta participación en investigaciones colaborativas de gran relevancia. Le siguen Li Y. y Zhang Y., ambos con 16

publicaciones y contribuciones fraccionadas de 4,96 y 4,95, respectivamente, reflejando su influencia significativa en el campo.

Liu Y. y Li J. también destacan con 14 y 12 publicaciones, y fracciones de 4,47 y 4,76, subrayando su participación activa en investigaciones relevantes. Autores como Li X., Wang X. y Wang Z., con 11 publicaciones cada uno, presentan impactos fraccionados entre 3,25 y 2,90, demostrando su importancia en proyectos colaborativos. Yada K. y Chen J., con 10 y 9 publicaciones, tienen impactos fraccionados de 3,20 y 2,42, respectivamente, reforzando su rol en el campo.

Finalmente, Li H., con 9 publicaciones y un impacto fraccionado de 3,15, también aparece como uno de los autores destacados. Esta tabla refleja la contribución significativa de estos investigadores a la literatura sobre *Big Data* y comportamiento del consumidor, evidenciando altos niveles de colaboración que se manifiestan en el impacto fraccionado de sus publicaciones. La influencia de estos autores es clave para comprender la evolución y desarrollo de este campo de estudio.

Tabla 3.2. Autores más relevantes em el campo del *Big Data* y Comportamiento del Consumidor

Autores	No. Artículos	Artículos Fraccionados
Wang Y	21	7.16
Li Y	16	4.96
Zhang Y	16	4.95
Liu Y	14	4.47
Li J	12	4.76
Li X	11	3.47
Wang X	11	3.25
Wang Z	11	2.90
Yada K	10	3.20
Chen J	9	2.42

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

La Tabla 3.3 muestra el impacto local de los autores más destacados en el campo, evaluado a través de los índices h, g y m. El índice h mide la productividad y el impacto de las citas de los autores; el índice g otorga más peso a las publicaciones con muchas citas; y el índice m ajusta el índice h por el tiempo transcurrido desde la primera publicación del autor, permitiendo analizar la productividad en función del tiempo.

Wang Y. encabeza la lista con un índice h de 8, un índice g de 21 y un índice m de 0,471, indicando una alta productividad sostenida en el tiempo y una fuerte influencia en el campo. Le siguen Li J. y Liu Y., ambos con un índice h de 7 y con índices g de 12 y 14, respectivamente, mostrando una contribución destacada en cuanto al número de citas y calidad de las publicaciones.

Zhang Y. y Li Y. también sobresalen con índices h de 7 y 6, respectivamente, y con índices g de 16 y 9. Sus índices m, de 0,333 y 0,353, reflejan un impacto positivo pero ligeramente menor en cuanto a productividad ajustada al tiempo en comparación con los autores más productivos. Otros autores como Tuzhilin A., Wang J. y Zhang C., con índices h de 6 e índices m que varían entre 0,25 y 0,667, muestran que, aunque sus contribuciones son más recientes, ya han tenido un impacto considerable.

Chen J. y Chen Q., con índices h de 5, también se encuentran entre los autores destacados, aunque con índices m ligeramente menores, reflejando un impacto más moderado en el tiempo. En conjunto, esta tabla proporciona una imagen detallada de la influencia local de los principales autores en la investigación sobre *Big Data* y comportamiento del consumidor, permitiendo comprender mejor cómo su productividad y el impacto de sus trabajos han evolucionado en este campo.

Tabla 3. 3. Impacto de los autores en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023

Autores	h_index	g_index	m_index
Wang Y	8	21	0.471
Li J	7	12	0.5
Liu Y	7	14	0.438
Zhang Y	7	16	0.333
Li Y	6	9	0.353
Tuzhilin A	6	8	0.25
Wang J	6	9	0.667
Zhang C	6	7	0.667
Zhang M	6	8	0.75
Chen J	5	9	0.227
Chen Q	5	6	0.238

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

En la Tabla 3.4, se observa que la categoría "Not Reported" lidera con 29 publicaciones, lo que puede indicar una amplia gama de contribuciones individuales sin vínculo directo con una institución particular o la falta de información detallada en los metadatos. Entre las instituciones reportadas, la University of Technology Sydney (Australia) se destaca con 23 publicaciones, seguida por Tamkang University (Taiwán) con 22 y Tsinghua University (China) con 20, reflejando la importancia de estas universidades en la producción de conocimiento sobre *Big Data* y comportamiento del consumidor.

Asimismo, otras instituciones como Islamic Azad University (Irán) con 18 publicaciones, Bina Nusantara University (Indonesia) y Northwestern Polytechnical University (China) con 16 publicaciones cada una, resaltan el rol de universidades de diversas regiones en esta área de estudio. En el ámbito europeo, la Agricultural University of Athens (Grecia), con 15 publicaciones, muestra su participación en esta línea de investigación, acompañada por Sejong University (Corea del Sur) y University of Queensland (Australia), ambas con 14 publicaciones.

La distribución de las publicaciones entre estas universidades destaca una fuerte presencia de Asia, particularmente de China, Taiwán y Corea del Sur, subrayando la relevancia de la región en la investigación sobre *Big Data* y comportamiento del consumidor. También se refleja la participación de otras regiones como Europa y Oceanía, mostrando el carácter internacional y multidisciplinario de este campo en expansión.

Tabla 3. 4. Afiliaciones relevantes en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023

Afiliación	Artículos
Not reported	29
University of Technology Sydney	23
Tamkang University	22
Tsinghua University	20
Islamic Azad University	18
Bina Nusantara University	16
Northwestern Polytechnical University	16
Agricultural University Of Athens	15
Sejong University	15
University of Queensland	14

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

El Gráfico 3.3 muestra la distribución y colaboración global en publicaciones sobre *Big Data* y comportamiento del consumidor (1995-2023). Las barras de los países están compuestas por SCP (*Single Country Publications*), que representan publicaciones de autores de un solo país (barras azules), mientras que los MCP (*Multiple Country Publications*) indican colaboración internacional (barras naranjas).

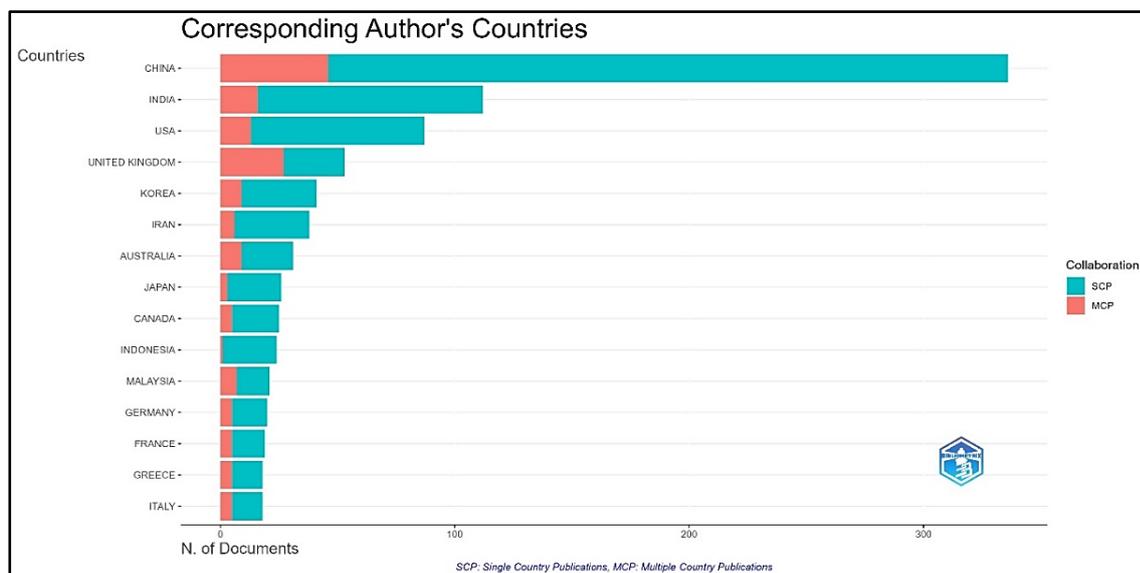
China lidera la producción con 336 publicaciones, de las cuales 290 son de autores de un solo país, subrayando su fuerte capacidad investigativa interna. Solo 46 de sus publicaciones han sido realizadas en colaboración internacional, representando un 13,7% de su producción total. Este patrón destaca la predominancia de la investigación autónoma dentro del país.

India, con 112 publicaciones, también tiene una mayoría de SCP (96), aunque presenta una mayor proporción de colaboraciones internacionales (16 artículos) en comparación con China. Estados Unidos cuenta con 87 publicaciones, de las cuales 13 fueron colaboraciones internacionales. Aunque su número de MCP es relativamente bajo, muestra una dinámica de colaboración mayor que la de China.

El Reino Unido es particularmente notable por su alto nivel de colaboración internacional. De sus 53 publicaciones, 27 fueron resultado de colaboraciones internacionales, representando más de la mitad de su producción (50,9%), lo que evidencia su enfoque hacia una investigación más globalizada y cooperativa. De manera similar, Australia y Corea del Sur también presentan una proporción significativa de MCP, con 9 colaboraciones internacionales cada uno en comparación con su producción total.

En contraste, países como Indonesia y Japón tienen una alta proporción de publicaciones nacionales con mínimas colaboraciones internacionales, lo que puede reflejar una mayor autonomía en sus proyectos de investigación. Este análisis ilustra la tendencia global de la investigación en *Big Data* y comportamiento del consumidor, destacando cómo algunos países optan por un enfoque más autónomo, mientras que otros, como el Reino Unido y Australia, se enfocan en la colaboración internacional, fomentando una mayor diversidad metodológica y enriqueciendo los resultados de la investigación.

Gráfico 3.3. Distribución y Colaboración Global en Publicaciones en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024)

Nota: La tabla de datos de este gráfico se adjunta en el Anexo No. 4.2.

La Tabla 3.5 presenta el análisis de los países más citados en la investigación sobre *Big Data* y comportamiento del consumidor. China lidera en términos de citas totales (TC) con 4.423, pero presenta un promedio de 13,20 citas por artículo, lo que podría sugerir un volumen elevado de publicaciones con un impacto moderado por documento.

Singapur se destaca notablemente con 2.368 citas totales y un promedio excepcional de 263,10 citas por artículo, indicando que sus publicaciones, aunque menos numerosas, tienen un impacto significativamente alto en la comunidad académica internacional. Esto refleja una alta calidad y relevancia de sus investigaciones en el campo.

Estados Unidos sigue a China en el total de citas (2.722), pero con un promedio mucho mayor de 31,30 citas por artículo, sugiriendo que las publicaciones estadounidenses tienen un impacto más profundo y sostenido en la literatura. Australia también se destaca con 1.473 citas totales y un promedio de 47,50 citas por artículo, consolidando su posición entre los países con investigaciones de alto impacto.

Por otro lado, países como Rumania y Portugal, aunque tienen un número total de citas menor (617 y 615, respectivamente), presentan promedios elevados de 68,60 y 43,90 citas por artículo, indicando que las publicaciones de estos países son altamente valoradas dentro de la literatura especializada. Este análisis subraya la importancia tanto de la cantidad de publicaciones como

de la calidad e impacto de las mismas en la investigación global sobre *Big Data* y comportamiento del consumidor.

Tabla 3. 5. Países más citados em el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023

País	Total de citas	Promedio de artículos citados
China	4423	13.20
Usa	2722	31.30
Singapore	2368	263.10
India	1475	13.20
Australia	1473	47.50
United Kingdom	1148	21.70
Korea	892	21.80
Canada	753	30.10
Romania	617	68.60
Portugal	615	43.90

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

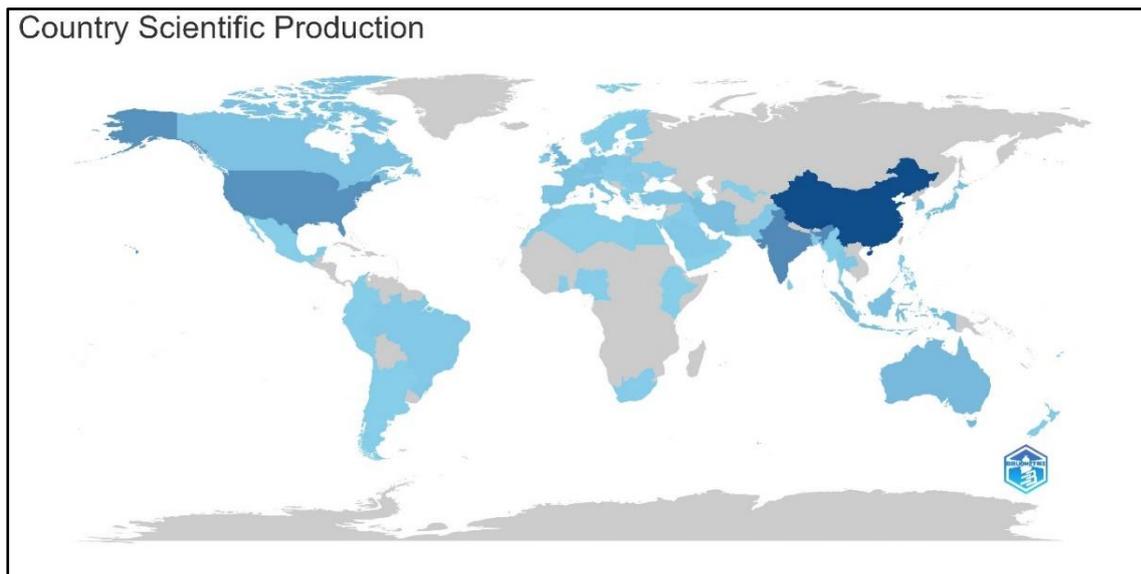
El Gráfico 3.4 muestra la distribución de la producción científica a nivel mundial en el campo del *Big Data* y comportamiento del consumidor durante el periodo 1995-2023. China encabeza esta lista con 1.299 publicaciones, subrayando su preeminencia en la investigación de estos temas, lo cual puede estar vinculado a su fuerte inversión en tecnología y ciencia de datos, así como a su enfoque en la digitalización y la innovación tecnológica.

Estados Unidos e India siguen a China con 580 y 658 publicaciones, respectivamente, manteniendo una sólida presencia en el campo, aunque con un volumen menor en comparación con China. Este liderazgo en producción científica refleja sus avanzadas capacidades tecnológicas y académicas, fortaleciendo su rol en la investigación global en *Big Data* y comportamiento del consumidor.

Otros países como el Reino Unido, Australia y Corea del Sur también demuestran una actividad investigativa considerable, con 209, 188 y 143 publicaciones, respectivamente. Estos números evidencian una participación significativa y diversificada en la investigación global de estas áreas.

Esta distribución resalta el interés global y las contribuciones variadas hacia el avance del conocimiento en estos campos influyentes, ofreciendo una visión panorámica de cómo diferentes regiones están explorando y expandiendo el entendimiento del *Big Data* y el comportamiento del consumidor.

Gráfico 3. 4. Producción científica por países en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024)

Nota: La tabla de datos de este gráfico se adjunta en el Anexo No. 4.3.

La Tabla 3.6 presenta los cinco artículos más citados en la base de datos de Scopus sobre *Big Data* y comportamiento del consumidor, los cuales destacan por su impacto en áreas clave como la automatización, el análisis de datos en redes sociales y la seguridad cibernética. El más citado, "*Brave new world: service robots in the frontline*" (Wirtz et al. 2018), con 1.213 citas, explora cómo los robots de servicio y la inteligencia artificial están transformando los sectores de servicios, abordando tanto su aceptación por los consumidores como las implicaciones éticas a nivel social. Este artículo ha sido relevante por su enfoque en la automatización y las implicaciones futuras del uso de robots en la interacción con los clientes.

Otro estudio destacado es "*Social media brand community and consumer behavior*" (Goh, Heng y Lin 2013), con 1.070 citas, que examina cómo el contenido generado por usuarios y por profesionales del marketing en comunidades de marcas influye en el comportamiento de compra de los consumidores. Sus hallazgos muestran que el contenido generado por usuarios

tiene un impacto mayor en las decisiones de compra, lo que subraya la importancia de las redes sociales en la creación de estrategias de marketing.

Asimismo, "*Adaptive fraud detection*" (Fawcett y Provost, 1997), con 669 citas, presenta un enfoque basado en minería de datos para detectar fraude, lo que ha influido significativamente en los sistemas de seguridad financiera. "*What can big data and text analytics tell us about hotel guest experience and satisfaction?*" (Xiang et al., 2015), con 658 citas, explora cómo el análisis de texto y *Big Data* ayudan a mejorar la experiencia de los huéspedes en la industria hotelera.

Finalmente, "*More than words: Social networks' text mining for consumer brand sentiments*" (Mostafa, 2013), con 460 citas, analiza el uso de la minería de texto en redes sociales para evaluar los sentimientos de los consumidores hacia marcas globales. Estos estudios demuestran la relevancia del *Big Data* y el análisis de texto en diversas industrias, desde la hospitalidad hasta el marketing, y reflejan el creciente interés académico en cómo estas tecnologías influyen en el comportamiento del consumidor.

Tabla 3. 6. Artículos más citados de Scopus en el campo de *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023

Título	Revista	Autores	País	Citas
<i>Brave new world: service robots in the frontline</i>	Journal of Service Management	(Wirtz et al. 2018)	Singapore; Australia; United States; United Kingdom; Germany;	1213
<i>Social media brand community and consumer behavior: Quantifying the relative impact of user- and marketer-generated content</i>	Information Systems Research	(Goh, Heng y Lin 2013)	Singapore; China	1070
<i>Adaptive fraud detection</i>	Data Mining and Knowledge Discovery	(Fawcett y Provost 1997),	United States	669
<i>What can big data and text analytics tell us about hotel guest experience and satisfaction?</i>	International Journal of Hospitality Management	(Xiang et al. 2015)	United States	658
<i>More than words: Social networks' text mining for consumer brand sentiments</i>	Expert Systems with Applications	(Mostafa, 2013)	Portugal	460

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

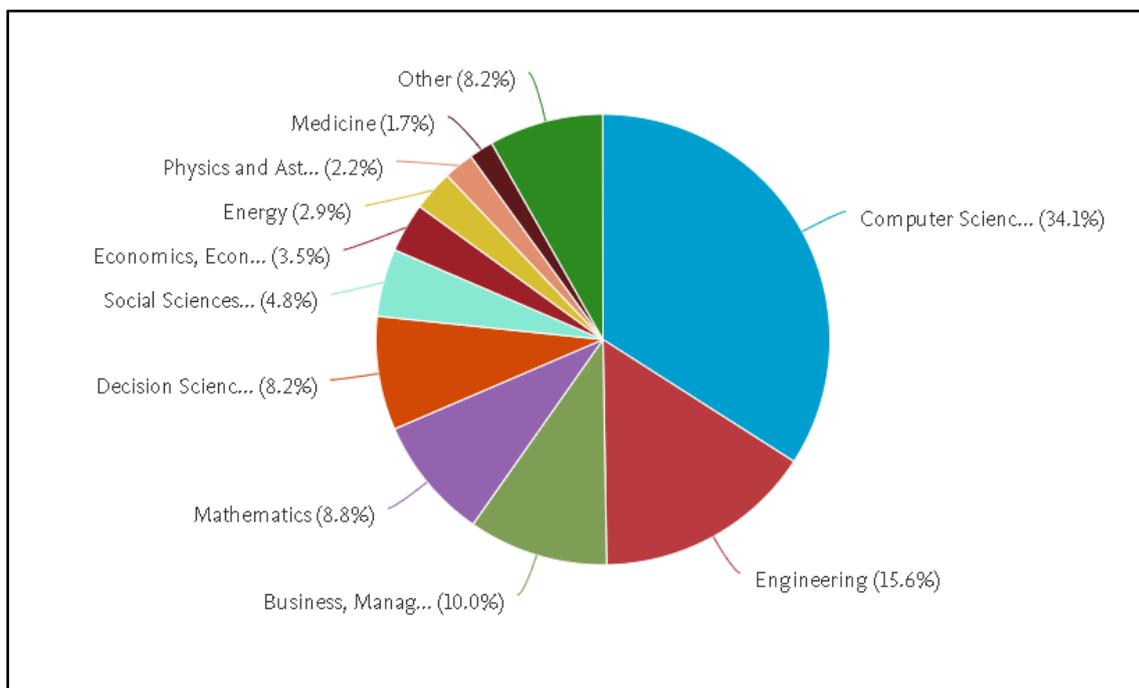
En el Gráfico 3.5, se visualiza la distribución de áreas temáticas entre los 1.861 documentos que abordan *Big Data* y comportamiento del consumidor, evidenciando una amplia

integración interdisciplinaria. Las ciencias de la computación dominan con un 34,1%, subrayando la centralidad de la informática en el procesamiento y análisis de grandes volúmenes de datos, importantes para comprender los patrones y tendencias en el comportamiento del consumidor.

El campo de la ingeniería, representando el 15,6%, junto con los estudios de negocios, gestión y administración (10,0%), resalta la aplicación práctica del *Big Data* para mejorar la eficiencia en los procesos y optimizar la toma de decisiones estratégicas dentro del ámbito empresarial. Estas disciplinas se benefician del análisis de grandes conjuntos de datos para generar soluciones innovadoras y efectivas.

Matemáticas y ciencias de la decisión acumulan el 8,8% y 8,2%, respectivamente, indicando el uso de herramientas analíticas y modelos predictivos avanzados para perfeccionar las estrategias de mercado. Además, disciplinas como las ciencias sociales (4,8%) y la economía (3,5%) están integrando el *Big Data* para mejorar el análisis de fenómenos sociales y económicos, permitiendo una mayor precisión en el desarrollo de políticas y en el estudio de mercados emergentes.

Gráfico 3. 5. Áreas temáticas en investigaciones en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024)

Nota: La tabla de datos de este gráfico se adjunta en el Anexo No. 4.4.

En el Gráfico 3.6, se presenta la nube de palabras generada a partir de las palabras clave más recurrentes en la literatura sobre *Big Data* y comportamiento del consumidor. Los términos "*Data Mining*" y "*Consumer Behavior*", con 860 y 638 menciones respectivamente, destacan como los conceptos predominantes, evidenciando la relevancia del análisis de grandes volúmenes de datos para comprender y prever las acciones del consumidor.

Otros términos clave como "*Sales*" (596), "*Big Data*" (406) y "*Customer Behavior*" (266) resaltan la importancia de entender las dinámicas comerciales y el comportamiento del cliente a partir de la recopilación y análisis de datos. Además, tecnologías y enfoques avanzados como "*Machine Learning*" (99), "*Artificial Intelligence*" (108) y "*Data Analytics*" (172) indican la adopción de herramientas computacionales para gestionar y extraer valor de la información disponible, evidenciando una tendencia hacia la automatización y la personalización en la investigación del comportamiento del consumidor.

Conceptos como "*E-commerce*" (242), "*Customer Satisfaction*" (171) y "*Social Networking - Online*" (110) sugieren que una parte significativa de los estudios se enfoca en plataformas digitales y en cómo optimizar las interacciones y experiencias de los consumidores en estos entornos. Estos términos reflejan el uso creciente de tecnologías disruptivas en el análisis del comportamiento de los consumidores, apoyando la toma de decisiones estratégicas en tiempo real y permitiendo una adaptación ágil a las preferencias y demandas del mercado digital.

Gráfico 3. 6. Nube de palabras en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024)

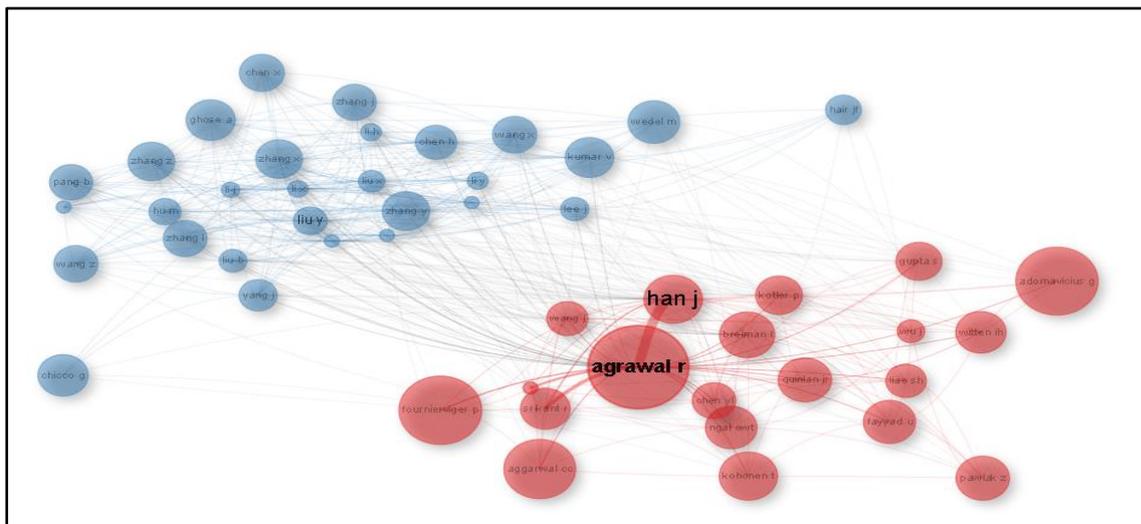
Nota: La tabla de datos de este gráfico se adjunta en el Anexo No. 4.5.

En el Gráfico 3.7, la red de co-citación refleja la estructura de colaboración e influencia en la investigación de *Big Data* y comportamiento del consumidor. El análisis muestra dos clústeres principales:

- Primer clúster, destacan autores como Agrawal R. y Han J., con altos valores de centralidad y PageRank, indicando su influencia en las metodologías de minería de datos y análisis de grandes volúmenes de información. Este clúster se enfoca en aplicaciones tecnológicas y análisis avanzado de datos, con otros autores como Kotler P. y Gupta S. contribuyendo en la intersección entre marketing y *Big Data*.
- Segundo clúster, resalta la presencia de Wang Y. y Liu Y., fuertemente conectados con investigadores en inteligencia artificial y aprendizaje automático aplicado al comportamiento del consumidor. Este grupo refleja un enfoque más técnico y analítico, con autores como Zhang Y. y Chen Y. sobresaliendo por sus contribuciones en áreas tecnológicas. La concentración de estos autores indica un interés creciente en implementar herramientas avanzadas para mejorar la toma de decisiones y el análisis predictivo en mercados dinámicos.

Esta red de co-citación evidencia cómo los autores colaboran y conectan disciplinas clave, permitiendo avances en las metodologías utilizadas para estudiar y comprender el comportamiento del consumidor en la era del *Big Data*. La diversidad de autores y conexiones destaca la multidisciplinariedad del campo, donde convergen ciencias de la computación, ciencias sociales y marketing para generar nuevos conocimientos y enfoques.

Gráfico 3. 7. Red de Co-citación entre autores en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024)

Nota: La tabla de datos de este gráfico se adjunta en el Anexo No. 4.6.

En el Gráfico 3.8, se presenta un mapa de estructura conceptual generado mediante el método de Análisis de Correspondencias Múltiples (MCA). En el Gráfico 3.8, se presenta un mapa de estructura conceptual generado mediante el método de Análisis de Correspondencias Múltiples (MCA), basado en las palabras clave de los autores. Este gráfico visualiza las relaciones entre diferentes términos y cómo se agrupan conceptualmente. Los puntos más cercanos representan términos que suelen aparecer juntos en los estudios, indicando una relación temática.

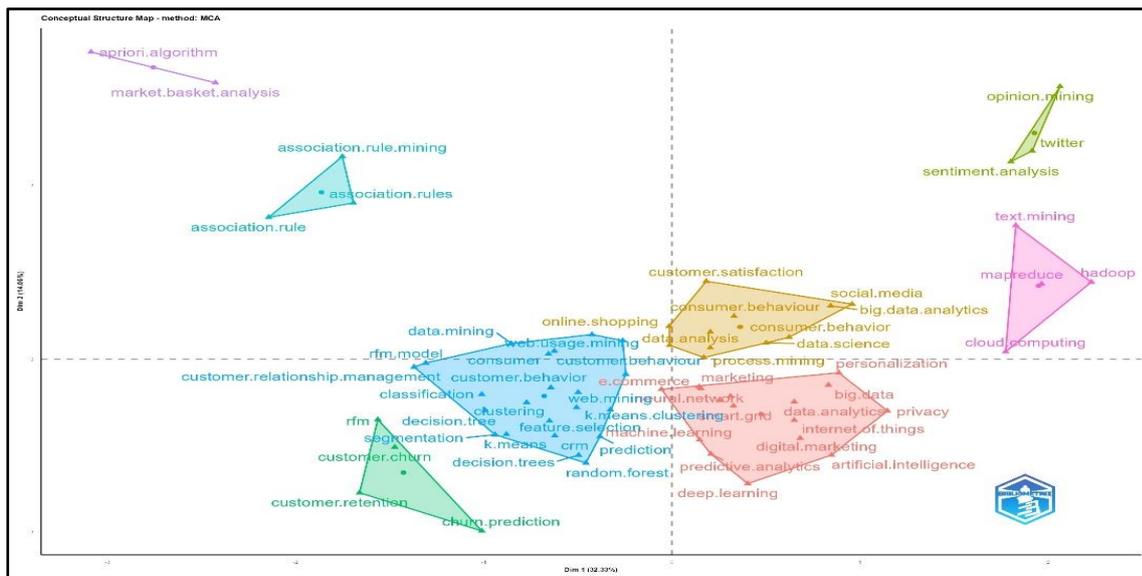
El mapa identifica varios clústeres:

- Clúster morado (superior izquierdo), incluye términos como *apriori algorithm*, *market basket analysis*, *association rule mining* y *association rules*, relacionados con métodos de análisis de patrones de compra en entornos comerciales. Estos algoritmos permiten identificar relaciones y asociaciones frecuentes entre productos comprados, esenciales en aplicaciones de recomendación de productos y minería de datos transaccionales.

- Clúster azul (inferior izquierdo), términos como *data mining*, *customer relationship management* (CRM), *classification*, *decision tree*, *segmentation* y *clustering*. Este clúster está orientado hacia la aplicación de técnicas de minería de datos y análisis predictivo en la gestión de clientes y segmentación de mercados. La integración de técnicas como *decision tree* y *k-means clustering* demuestra la relevancia de los métodos de clasificación y agrupamiento en la segmentación de consumidores basada en grandes conjuntos de datos.
- Clúster verde (cerca del azul), incluye *customer retention*, *churn prediction* y *customer churn*, enfocándose en el análisis y predicción del comportamiento de retención o abandono de clientes. Este grupo es esencial en industrias como telecomunicaciones, banca y servicios de suscripción, donde se utilizan técnicas de predicción para identificar clientes con mayor probabilidad de cancelar un servicio.
- Clúster amarillo (centro-derecha), agrupa *consumer behaviour*, *customer satisfaction*, *online shopping*, *social media* y *e-commerce*, destacando la interrelación entre el comportamiento del consumidor y las plataformas de comercio electrónico. Los estudios en esta área se enfocan en comprender cómo los consumidores interactúan con las plataformas en línea y cómo sus decisiones de compra están influenciadas por la información en redes sociales.
- Clúster rosa (derecha), contiene términos como *mapreduce*, *hadoop*, *cloud computing* y *text mining*, representando tecnologías de procesamiento de grandes volúmenes de datos fundamentales en el análisis de *Big Data*. *Hadoop* y *MapReduce* son tecnologías clave para manejar conjuntos de datos masivos, permitiendo procesar grandes cantidades de información de manera eficiente.
- Clúster verde claro (superior derecha), incluye *opinion mining*, *Twitter* y *sentiment analysis*, enfocándose en el análisis de opiniones y sentimientos en redes sociales. La minería de sentimientos es esencial para comprender las emociones y opiniones expresadas por los consumidores en plataformas como *Twitter*, informando estrategias de marketing y desarrollo de productos.
- Clúster rojo (inferior central), contiene *artificial intelligence*, *deep learning*, *predictive analytics*, *internet of things* y *neural networks*, relacionado con el uso de inteligencia artificial y aprendizaje automático en el análisis y predicción del comportamiento del consumidor. La inclusión de *internet of things* sugiere que el análisis de datos generado por dispositivos conectados está ganando relevancia en este campo.

Este mapa conceptual revela una estructura compleja de interrelaciones entre tecnologías avanzadas, métodos de análisis y aplicaciones específicas, destacando el impacto significativo del *Big Data* en áreas como la gestión de clientes, la minería de opiniones en redes sociales y el marketing digital.

Gráfico 3. 8. Mapa de estructura conceptual en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024)

En el Gráfico 3.9, se presenta un diagrama de evolución temática en el campo del *Big Data* y el Comportamiento del Consumidor en dos periodos clave: 1995-2012 y 2013-2023. Este diagrama visualiza cómo las palabras clave y temas centrales en la investigación se han transformado y diversificado a lo largo del tiempo, reflejando cambios en las áreas de enfoque y las aplicaciones prácticas de las tecnologías relacionadas con *Big Data*.

La elección de 2013 como un punto de corte en este análisis responde a varios factores críticos señalados en la investigación de Chávez et al. (2022). Antes de esta fecha, la mayor parte de los estudios sobre *Big Data* y su relación con el comportamiento del consumidor estaban restringidos a las ciencias de la computación y la ingeniería, con un fuerte enfoque en el desarrollo de técnicas computacionales para manejar grandes volúmenes de datos. Los términos clave como *web mining*, *data mining*, *fraud detection*, y *business intelligence* dominaron este periodo, reflejando un interés en las metodologías para extraer valor de los datos.

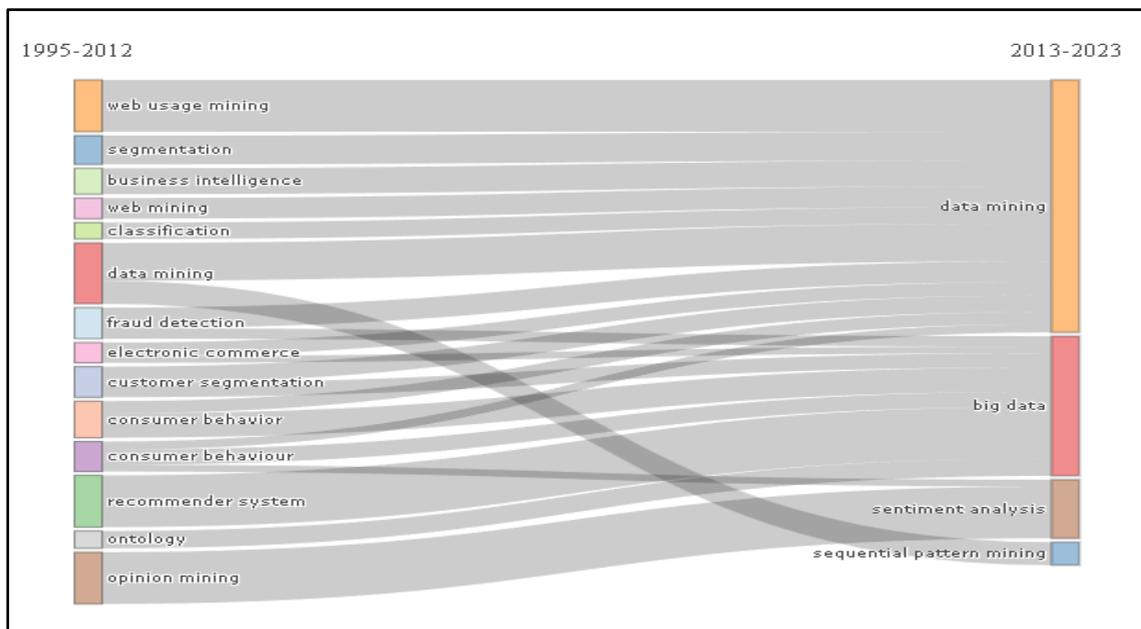
Sin embargo, a partir de 2013, se observó una diversificación notable en el campo, donde las aplicaciones de *Big Data* comenzaron a expandirse más allá de las ciencias técnicas hacia áreas como el marketing, el comportamiento del consumidor, las ciencias sociales y los negocios. Esto se debió en gran parte al crecimiento exponencial de los datos disponibles y a la necesidad de abordar problemas complejos de una manera más interdisciplinaria. La investigación comenzó a incorporar no solo los métodos computacionales para procesar datos, sino también enfoques más orientados hacia la comprensión del impacto de *Big Data* en diversas industrias y comportamientos humanos.

Durante el periodo 1995-2012, los temas predominantes fueron *data mining*, *web mining*, *fraud detection* y *business intelligence*, reflejando un énfasis en las herramientas y tecnologías para analizar grandes volúmenes de datos. La minería de datos y la inteligencia de negocios jugaron un papel central en la extracción de información útil del *Big Data*.

A partir de 2013, se observa una diversificación notable en el campo, donde las aplicaciones de *Big Data* comenzaron a expandirse más allá de las ciencias técnicas hacia áreas como el marketing, el comportamiento del consumidor, las ciencias sociales y los negocios. Temas como *sentiment analysis*, *sequential pattern mining* y *Big Data analytics* indican una maduración en el campo y un mayor enfoque en la aplicación de estos conocimientos a áreas como el marketing digital, la personalización y la toma de decisiones basada en datos.

El surgimiento de *sentiment analysis* y *sequential pattern mining* señala una transición hacia un enfoque más detallado y orientado al consumidor. Estos avances subrayan una mayor sofisticación en el uso del *Big Data* para influir en la estrategia empresarial y el comportamiento del consumidor, permitiendo anticipar tendencias y personalizar ofertas según las necesidades específicas de los clientes.

Gráfico 3. 9. Diagrama de evolución de Sankey en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

En el Gráfico 3.10, se presenta la red de co-ocurrencia de palabras clave basada en los documentos obtenidos de la base de datos de Scopus, reflejando cómo distintas palabras clave interactúan dentro del ámbito de la investigación. Este gráfico es una herramienta visual poderosa para entender las conexiones entre temas de investigación, mostrando la frecuencia y la relación de co-ocurrencia de términos específicos dentro de la literatura científica.

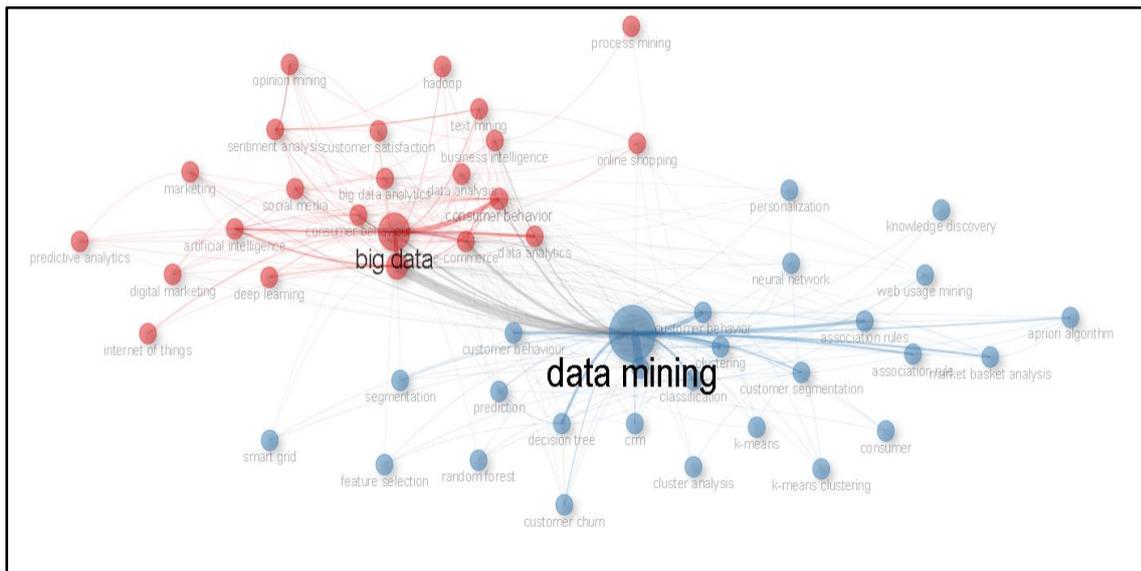
- Centralidad de *Big Data* y *Data Mining*, en el centro del gráfico, estos términos aparecen como nodos predominantes, reflejando su papel central en la investigación actual. Actúan como *hubs* que conectan diversas áreas de estudio, subrayando su relevancia transversal en múltiples disciplinas y aplicaciones.
- Clúster tecnológico (rojo), agrupa términos como *machine learning*, *artificial intelligence* y *deep learning*, ilustrando el enfoque intensivo en tecnología dentro de los estudios de *Big Data*. La densidad de conexiones señala una integración fuerte de técnicas avanzadas de análisis en la exploración del comportamiento del consumidor y otras aplicaciones de negocio.
- Clúster de aplicaciones comerciales y sociales (azul), con términos como marketing, *social media* y *consumer behavior*, este clúster destaca la aplicación del *Big Data* en entender y predecir las dinámicas de mercado y el comportamiento del consumidor. Las conexiones

fuertes con *data analysis* y *predictive analytics* refuerzan la idea de que el análisis de datos está profundamente incorporado en las estrategias comerciales modernas.

El Clúster de innovación y nuevos campos, incluye términos emergentes como *internet of things* y *smart grid*, que vinculan *Big Data* con innovaciones en infraestructura y tecnología. Estas áreas representan campos de creciente interés donde *Big Data* puede tener un impacto significativo en el desarrollo de soluciones sostenibles y avanzadas.

La prominencia de conexiones entre *Big Data* con *artificial intelligence* y *machine learning* indica una tendencia hacia la automatización y la mejora de la eficiencia en el análisis de grandes volúmenes de datos. La intersección entre estas áreas sugiere una evolución hacia sistemas más autónomos y adaptativos que pueden aprender y ajustarse a patrones complejos de datos para ofrecer *insights* más precisos y accionables.

Gráfico 3. 10. Red de Ocurrencia de Palabras Clave en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

En el Gráfico 3.11, se presenta un Mapa Temático generado por la herramienta Bibliometrix de R Studio. Este gráfico organiza las palabras clave en un espacio de dos dimensiones basado en dos criterios: el grado de relevancia (*Centrality*) y el grado de desarrollo (*Density*). El mapa está dividido en cuatro cuadrantes, cada uno representando diferentes niveles de relevancia y desarrollo de los temas en el campo de estudio.

En el cuadrante de Temas Emergentes o en Declive, este cuadrante, ubicado en la parte inferior izquierda del gráfico, agrupa palabras clave que, aunque relevantes en algún

momento, ahora están menos desarrolladas o perdiendo relevancia. En este caso, términos como *"customer behavior"*, *"personalization"*, *"hadoop"* y *"data science"* se encuentran en este espacio. Estos temas, a pesar de ser importantes, parecen estar en una fase de transición, posiblemente debido a la saturación de investigaciones previas o a la evolución de nuevas tecnologías y enfoques que los han desplazado. El declive de estos temas sugiere que las investigaciones podrían estar migrando hacia aplicaciones más avanzadas o enfoques más específicos.

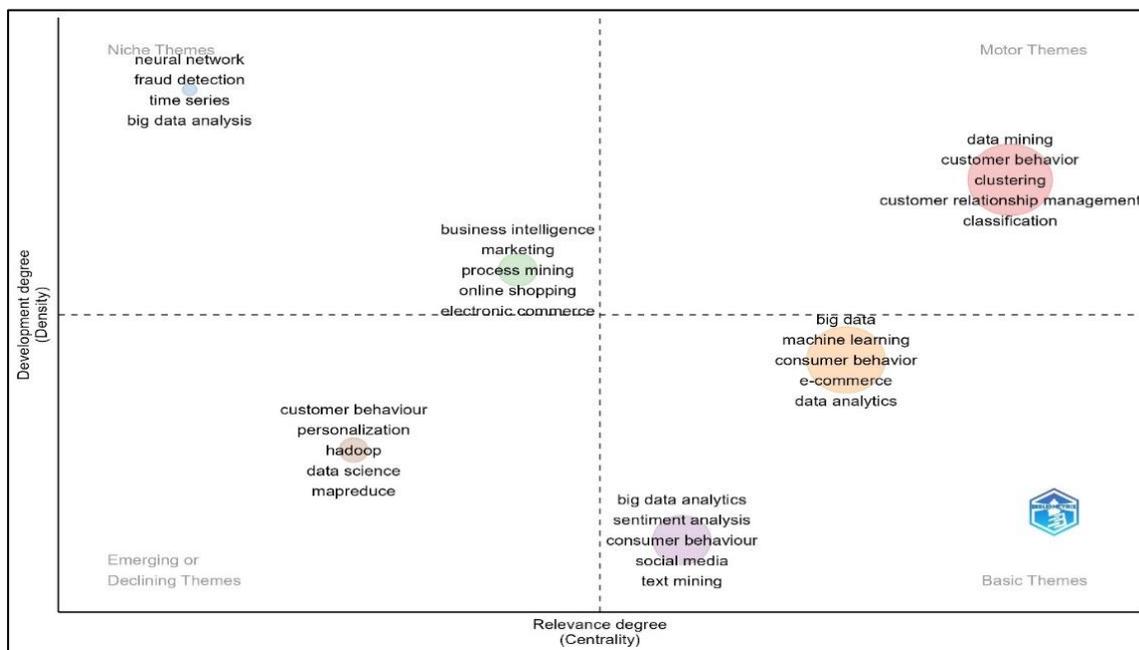
En la parte inferior derecha se encuentran los Temas Básicos, aquellos que, aunque tienen una alta centralidad, requieren aún un mayor desarrollo para madurar. Entre los términos presentes están *"Big Data"*, *"machine learning"*, *"e-commerce"*, y *"data analytics"*, que representan las bases sobre las cuales se construyen gran parte de las investigaciones actuales. Estos términos son fundamentales porque sirven como pilares en la investigación sobre el análisis de grandes volúmenes de datos y su relación con el comportamiento del consumidor. Su presencia como temas centrales indica que, aunque ampliamente investigados, aún existen oportunidades para profundizar en su desarrollo a través de nuevas metodologías o enfoques más especializados.

Ubicados en la parte superior izquierda del gráfico, los Temas de Nicho son aquellos bien desarrollados pero que tienen menor centralidad, lo que significa que son áreas muy especializadas o técnicas. Entre ellos se encuentran términos como *"neural network"*, *"fraud detection"*, *"time series"*, y *"Big Data analysis"*. Aunque estas áreas están altamente desarrolladas, su contribución se limita a aplicaciones muy específicas o industrias concretas, como la detección de fraudes o el análisis de series temporales. Este grupo de temas suele estar más vinculado a investigaciones técnicas o con un alto nivel de especialización, que no necesariamente se cruzan con las áreas de investigación más generales, pero que son decisivos en ciertos contextos.

En la parte superior derecha, los Temas Motores son los más relevantes y desarrollados, aquellos que impulsan la investigación y tienen un impacto significativo. Aquí se encuentran términos como *"data mining"*, *"customer behavior"*, *"clustering"*, *"classification"* y *"customer relationship management"*. Estos temas son clave porque tienen un alto nivel de desarrollo, son fundamentales para el progreso de la investigación en *Big Data* y su relevancia refleja que son áreas con una alta productividad científica y con aplicaciones directas en la industria, especialmente en la segmentación de clientes y la optimización de relaciones con los consumidores.

Este análisis proporciona una visión clara de cómo se organizan y evolucionan los temas en el campo del *Big Data* y el comportamiento del consumidor, destacando tanto las áreas centrales como las emergentes y de nicho.

Gráfico 3. 11. Mapa Temático en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus.

En el Gráfico 3.12, se presenta un “*Three-Field Plot*”, comúnmente utilizado en análisis bibliométrico para visualizar las relaciones entre tres campos de interés en un conjunto de documentos. Este gráfico ayuda a identificar las conexiones entre autores, instituciones y temas clave dentro del campo del comportamiento del consumidor y el *Big Data*, revelando la importancia de ciertas instituciones y autores en la promoción y el desarrollo de la investigación en estas áreas.

En la columna de la izquierda, los autores más destacados en el campo de *Big Data* y Comportamiento del Consumidor están visualizados. Los autores como Wang Y., Zhang Y., y Li Y. sobresalen, reflejando su notable producción científica en el área. Estos autores tienen conexiones directas con conceptos clave en la investigación, lo que indica su rol central en avanzar el conocimiento sobre estos temas. Por ejemplo, Wang Y. y Zhang Y. muestran múltiples conexiones con palabras clave como *data mining* y *consumer behavior*, destacando su contribución significativa en estos aspectos.

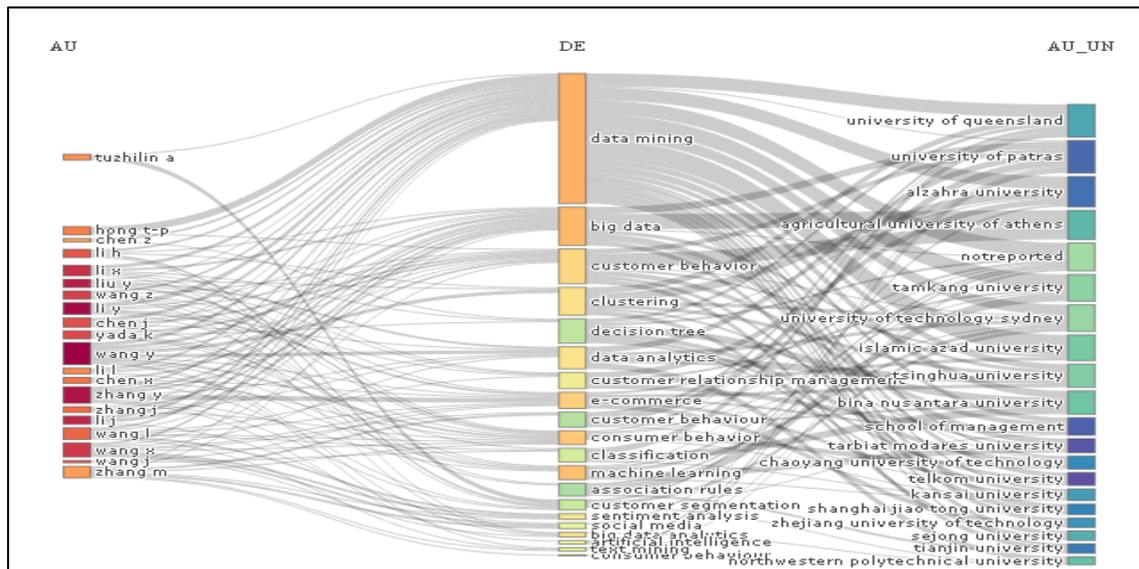
En la columna central, se representan los temas más investigados, que incluyen términos como *data mining*, *big data*, *customer behavior*, y *clustering*. Esto subraya la importancia de la minería de datos y el análisis del comportamiento del consumidor en la investigación actual. El término *customer behavior* está especialmente bien conectado con autores clave, lo que sugiere un enfoque constante en comprender cómo los datos masivos pueden informar sobre los patrones de comportamiento de los consumidores.

En la columna derecha, se visualizan las instituciones más activas en la producción de artículos sobre *Big Data* y Comportamiento del Consumidor. Tsinghua University, Islamic Azad University, y University of Queensland son algunas de las instituciones más prominentes, lo que refleja su fuerte enfoque en la investigación en estos temas. Estas instituciones, a menudo en colaboración con los autores destacados mencionados en el campo izquierdo, desempeñan un papel primordial en la generación de conocimiento sobre la intersección entre tecnología y comportamiento del consumidor.

Este gráfico permite ver cómo las instituciones, los autores y los temas clave están interrelacionados, mostrando una sólida red de investigación que respalda el avance en el campo de *Big Data* y Comportamiento del Consumidor. La importancia de términos como *data mining*, *customer behavior*, y *big data* en combinación con autores e instituciones influyentes indica que el enfoque de la investigación está centrado en la explotación de grandes volúmenes de datos para predecir y analizar patrones de comportamiento de los consumidores.

Esta representación visual también destaca que muchos de los temas clave están vinculados a tecnologías avanzadas como la inteligencia artificial y la minería de datos, lo que refuerza la tendencia actual de utilizar estas herramientas para mejorar el entendimiento y la toma de decisiones basada en datos dentro del marketing y otras áreas comerciales.

Gráfico 3. 12.. Relación entre Autores, Palabras Clave y Afiliaciones en la Investigación de Big Data y Comportamiento del Consumidor en el campo del Big Data y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

El segundo análisis realizado se basa en la clasificación proporcionada por el Tree of Science (ToS), como se muestra en la Tabla 3.7. Este enfoque organiza la literatura en categorías de *ROOT*, *TRUNK*, *BRANCH* y *LEAVES*, reflejando la evolución y estructura del campo del *Big Data* aplicado al comportamiento del consumidor.

ROOT hace referencia a los trabajos fundamentales que establecen las bases teóricas y metodológicas esenciales, por ejemplo, el trabajo pionero de Fayyad U.M. (1996) en "*Advances in Knowledge Discovery and Data Mining*" sentó las bases para las técnicas de minería de datos, fundamentales para extraer información valiosa de grandes conjuntos de datos. Del mismo modo, Agrawal R. (1993, 1994) contribuyó significativamente al desarrollo de procesos de minería de datos y descubrimiento de conocimiento, especialmente en el contexto del análisis de cesta de mercado y reglas de asociación. Han J. (2001) amplió estos conceptos proporcionando un marco integral que ha sido instrumental para investigaciones posteriores en este campo.

Los artículos en la categoría *TRUNK* representan estudios que han estructurado el campo del *Big Data* y el comportamiento del consumidor. Las categorías centrales identificadas por ToS incluyen minería de datos, comportamiento del consumidor, *Big Data*, ventas y comercio. Zhang M. (2022) explora las ventas de marcas en compras en línea basadas en el

procesamiento de *Big Data* del *IoT*, proporcionando un marco para entender cómo los datos pueden optimizar las estrategias de marketing digital. Zhang R. (2023) investiga la minería de datos del comportamiento del consumidor en plataformas de comercio electrónico social, utilizando algoritmos mejorados de *clustering* espectral para descubrir patrones en las interacciones de los consumidores.

Li J. (2023) introduce un método de minería profunda para datos de comportamiento del consumidor de usuarios de comercio electrónico, combinando técnicas de *clustering* y aprendizaje profundo para mejorar la segmentación y personalización de usuarios. Zhao X. (2021) examina la clasificación del riesgo de accidentes individuales basándose en datos del Internet de Vehículos (IoV) y datos de comportamiento del consumidor fuera de línea, demostrando cómo el análisis de *Big Data* puede mejorar la seguridad vial y la experiencia del consumidor.

Chen L-S. (2013) se centra en la clasificación de manipulación de reseñas utilizando árboles de decisión, destacando los desafíos de la integridad de datos y la confianza en las plataformas en línea. Wang J. (2021) discute la fijación de precios basada en el comportamiento y la contratación al por mayor bajo la competencia en la cadena de suministro, enfatizando el uso estratégico de datos de comportamiento del consumidor en decisiones de precios.

Las categorías *BRANCH* revelan especializaciones y diversificaciones temáticas dentro del estudio del *Big Data* y el comportamiento del consumidor. En *BRANCH 1*, caracterizada por temas como *Big Data*, perfilado, ontología, ventas y comportamiento del consumidor, se incluyen estudios como el de Ganhewa N.B. (2021) sobre soluciones de optimización de ventas para el comercio minorista de moda, demostrando cómo el perfilado de clientes y las herramientas de *Big Data* pueden mejorar las estrategias de ventas.

Otros estudios en *BRANCH 1*, como los de Elyusufi Z. (2019, 2018), discuten la combinación de fortalezas de CRM con herramientas de *Big Data* para el análisis de perfiles de clientes y el uso de arquitectura CEP en un contexto de *Big Data*. Estos trabajos subrayan la importancia de integrar diversas fuentes de datos y técnicas analíticas para obtener *insights* profundos sobre el comportamiento del consumidor.

BRANCH 2 se concentra en estudios de estilo de vida, minería de datos, comportamiento del consumidor, grandes conjuntos de datos y marketing sostenible. Por ejemplo, Lazaroiu G. (2023) investiga los algoritmos de inteligencia artificial y las tecnologías de computación en

la nube en la gestión *fintech* basada en *blockchain*, destacando la intersección de tecnologías avanzadas y el comportamiento del consumidor en los servicios financieros. Wahdiniwaty R. (2023) discute los algoritmos de *Big Data* como habilitadores de prácticas de marketing sostenible, enfatizando el papel del análisis de datos en la promoción de la sostenibilidad.

BRANCH 3 se centra en el comportamiento del cliente, minería de datos, aprendizaje automático, *Big Data*, ventas y privacidad de datos. Estudios como el de López M.B.V. (2023) aplican modelos de minería de datos para predecir la deserción de clientes en la industria de telecomunicaciones, mostrando aplicaciones prácticas de analítica predictiva en la retención de clientes. Ahmadon M.A.B. (2021) explora la minería de servicios al consumidor basándose en las preferencias del usuario con minería asociativa y de procesos, destacando estrategias de marketing personalizadas.

Finalmente, los artículos en la categoría LEAVES representan investigaciones recientes que sintetizan y amplían los temas del comportamiento del consumidor y el *Big Data*. Las categorías principales identificadas son minería de datos, comportamiento del consumidor, *Big Data*, ventas y aprendizaje automático. Cao P. (2023) estudia el impacto de los estímulos de tecnología de inteligencia artificial en el comportamiento de consumo sostenible, proporcionando evidencia de usuarios de Ant Forest en China. Esta investigación subraya la influencia de las tecnologías de IA en la promoción de prácticas de consumo sostenible.

Wei L. (2023) examina la psicología y el comportamiento del consumidor basados en el Internet de las Cosas y el análisis de *Big Data*, ilustrando la integración de dispositivos interconectados en la comprensión de patrones de consumo. Awate A.S. (2023) ofrece una encuesta integral de técnicas de segmentación y clasificación en la era del *Big Data*, enfatizando la necesidad de métodos analíticos avanzados para manejar datos de consumidores a gran escala.

Este análisis demuestra que el campo del *Big Data* y el comportamiento del ha crecido en profundidad y especialización y continúa evolucionando rápidamente, abordando nuevos desafíos y adaptándose a los avances tecnológicos y sociales. La estructura proporcionada por el Tree of Science ofrece una vista holística de este campo dinámico, destacando tanto su desarrollo histórico como sus futuras trayectorias.

Los trabajos de *ROOT* establecen metodologías y conceptos esenciales, mientras que los artículos de *TRUNK* construyen sobre ellos para estructurar aún más el campo. Las categorías de *BRANCH* muestran diversificación temática, abordando tópicos especializados como el

perfilado, marketing sostenible y privacidad de datos. Los *LEAVES* representan investigaciones de vanguardia que integran tecnologías avanzadas como *IA*, *IoT* y aprendizaje automático para explorar nuevas dimensiones del comportamiento del consumidor.

Tabla 3. 7. Principales resultados de la Herramienta Tree of Science en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023

Categoría	Palabras clave ToS	Investigación
Root	<i>data mining, consumer behavior, big data, sales, customer behavior</i>	Fayyad UM. (1996). <i>Advances in Knowledge Discovery and Data Mining</i> .
		Agrawal R. (1993). <i>IEEE Transactions on Knowledge and Data Engineering, Vol. 5, P914</i> .
		Agrawal R. (1994). <i>Proceedings of the 20th VLDB Conference</i> .
		Han J. (2001). <i>Data Mining: Concepts and Techniques</i> .
		Ngai EWT. (2009). <i>Expert Systems with Applications, Vol. 36</i> .
		Duan L. (2009). <i>Annals of Operations Research, Vol. 168, P151</i> .
		Casillas J. (2005). <i>IEEE Transactions on Fuzzy Systems, Vol. 13, P13</i> .
		Bae JK. (2010). <i>Expert Systems with Applications, Vol. 37, P1821</i> .
		Chen L-F. (2012). <i>Intelligent Data Analysis, Vol. 16, P167</i> .
		Liang L. (2021). <i>Journal of Healthcare Engineering, Article ID 6673649</i>
Trunk	<i>data mining, consumer behavior, big data, sales, commerce</i>	Zhang M. (2022). <i>Online Shopping Brand Sales Based on IoT Big Data Processing</i> .
		Zhang R. (2023). <i>Consumer Behaviour Data Mining of Social E-commerce Platform Based on Improved Spectral Clustering Algorithm</i> .
		Li J. (2023). <i>A Deep Mining Method for Consumer Behaviour Data of E-commerce Users Based on Clustering and Deep Learning</i> .
		Zhao X. (2021). <i>Individual Driver Crash Risk Classification Based on IoV Data and Offline Consumer Behavior Data</i> .
		Chen L-S. (2013). <i>A Study on Review Manipulation Classification Using Decision Tree</i> .
		Wang J. (2021). <i>Behaviour-based Pricing and Wholesaling Contracting under Supply Chain Competition</i> .

		Shan Y. (2023). <i>Incorporating User Behavior Flow for User Risk Assessment</i> .
		Wang Y. (2022). <i>Big Data Mining Method of Marketing Management Based on Deep Trust Network Model</i> .
		Li T. (2022). <i>Influence of Content Marketing on Consumer Brand Behavior in the Era of Big Data</i> .
		Zhang M. (2021). <i>Data Mining Algorithm for Demand Forecast Analysis on Flash Sales Platform</i> .
Branch 1	<i>big data, profiling, ontology, sales, consumer behavior</i>	Ganhewa NB. (2021). <i>Sales Optimization Solution for Fashion Retail</i> .
		Artés J. (2019). <i>Tax Reforms and Google Searches: The Case of Spanish VAT Reforms During the Great Recession</i> .
		Elyusufi Z. (2019). <i>Combining CRM Strength and Big Data Tools for Customers Profile Analysis</i> .
		Gil M. (2018). <i>Nowcasting Private Consumption: Traditional Indicators</i> .
		Camacho M. (2018). <i>Tourism Economics, Vol. 24, P434</i> .
		Elyusufi Z. (2018). <i>Customer Profiling Using CEP Architecture in a Big Data Context</i> .
		Elliott D. (2018). <i>Tippers and Stiffers: An Analysis of Tipping Behavior in Taxi Trips</i> .
		Hossain ASMS. (2017). <i>3rd International Conference on Electrical Information and Communication Technology (EICT)</i> .
		Vellido A. (2017). <i>Electricity Rate Planning for the Current Consumer Market Scenario Through Segmentation of Consumption Time Series</i> .
		Petitjean F. (2016). <i>Knowledge and Information Systems, Vol. 47, P1</i> .
Branch 2	<i>lifestyle study, data mining, consumer behavior, large datasets, sustainable marketing</i>	Lăzăroiu G. (2023). <i>Artificial Intelligence Algorithms and Cloud Computing Technologies in Blockchain-based Fintech Management</i> .
		Jatain D. (2023). <i>International Journal of Information Management Data Insights, Vol. 3</i> .
		Wahdiniwaty R. (2023). <i>Big Data Algorithms as Enablers of Sustainable Marketing Practices: A Preliminary Research</i> .
		Zauskova A. (2022). <i>Visual Imagery and Geospatial Mapping Tools in the Metaverse Interactive Environment</i> .
		Zvarikova K. (2022). <i>Psychosociological Issues in Human Resource Management</i> .
		Zhao Z. (2022). <i>Analytics on the Cloud-Inner Melbourne</i> .

		Peters MA. (2022). <i>Analysis and Metaphysics, Vol. 21, P7.</i>
		Hancock K. (2022). <i>Geospatial Mapping and Remote Sensing Technologies across the Blockchain-based Metaverse.</i>
		Laviola E. (2022). <i>The International Journal of Advanced Manufacturing Technology, Vol. 119, P1769.</i>
Branch 3	<i>customer behavior, data mining, machine learning, big data, sales, data privacy</i>	López MBV. (2023). <i>Application of a Data Mining Model to Predict Customer Defection. Case of a Telecommunications Company in Peru.</i>
		Ahmadon MAB. (2021). <i>Mining Consumer Services Based on User Preference with Associative and Process Mining.</i>
		Sook Ling L. (2021). <i>Customer Churn Prediction for Telecommunication Industry: A Malaysian Case Study.</i>
		Wassouf WN. (2020). <i>Predictive Analytics Using Big Data for Increased Customer Loyalty: Syriatel Telecom Company Case Study.</i>
		Höppner S. (2020). <i>European Journal of Operational Research, Vol. 284, P920.</i>
		Nagarkar P. (2020). <i>Twitter Data Mining for Targeted Marketing.</i>
		Asniar. (2019). <i>Predictive Analytics for Predicting Customer Behavior.</i>
		Al-Zuabi IM. (2019). <i>Predicting Customer's Gender and Age Depending on Mobile Phone Data.</i>
		Kulkarni H. (2019). <i>Multi-Agent System for Customer Behavior Tracking Using Shoppers' Path or Traversal</i>
Leaves	<i>data mining, consumer behavior, big data, sales, machine learning</i>	Cao P. (2023). <i>The Impact of Artificial Intelligence Technology Stimuli on Sustainable Consumption Behavior: Evidence from Ant Forest Users in China.</i>
		Wei L. (2023). <i>Internet of Things and Big Data Based Analytics on Consumer Psychology and Behavior.</i>
		Lăzăroiu G. (2023). <i>Artificial Intelligence Algorithms and Cloud Computing Technologies in Blockchain-based Fintech Management.</i>
		Awate AS. (2023). <i>Understanding Customer Behaviour: A Comprehensive Survey of Segmentation and Classification Techniques in the Age of Big Data.</i>
		Sakas DP. (2023). <i>Leading Logistics Firms' Re-Engineering through the Optimization of the Customer's Social Media and Website Activity.</i>
		Hu W. (2023). <i>Construction of Purchase Intention Model of Digital Music Products Based on Data Mining Algorithm.</i>

		Xiaohong M. (2023). <i>A Hybrid Macro and Micro Method for Consumer Emotion and Behavior Research.</i>
		Lv H. (2023). <i>E-commerce Consumer Behavior Analysis Based on Big Data.</i>
		Huda I. (2023). <i>Design of Prediction Model Using Data Mining for Segmentation and Classification Customer Churn in E-Commerce Mall in Mall.</i>
		Kouarfaté BB. (2023). <i>Understanding Consumer Attitudes toward Cultured Meat: The Role of Online Media Framing</i>

Fuente: Tree of Science, base de datos Scopus. (2024).

En el Gráfico 3.13, se presenta una Red de Colaboración utilizada en análisis bibliométrico para visualizar las relaciones de co-autoría entre los investigadores dentro de un campo de estudio específico. Este tipo de gráfico es útil para identificar las redes de colaboración y los principales actores dentro del ámbito de la investigación.

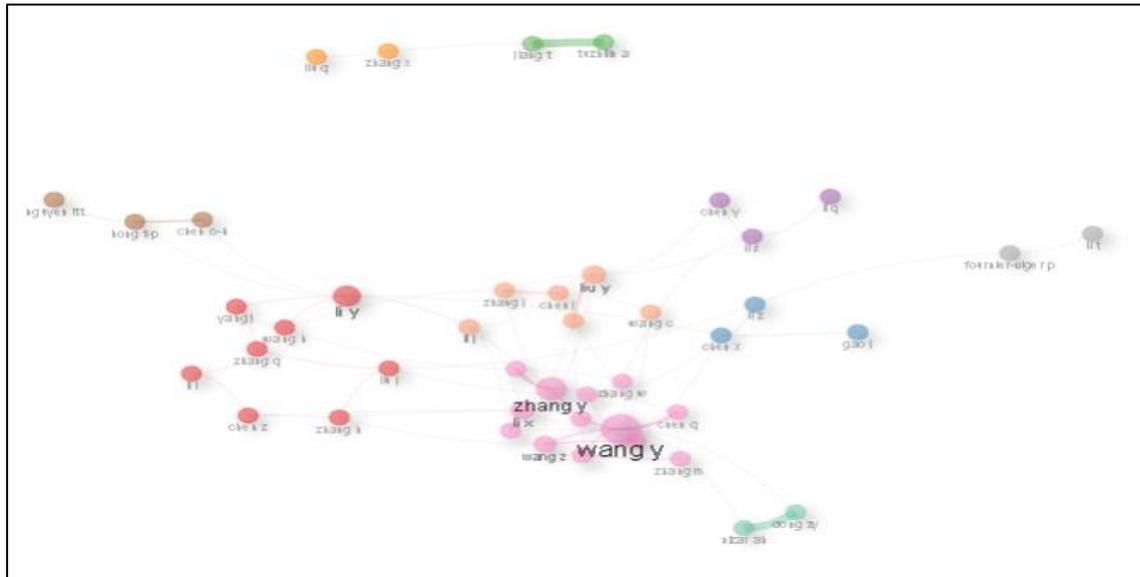
Los nodos en el gráfico representan a los autores, y el tamaño del nodo está relacionado con la cantidad de publicaciones o la centralidad del autor en la red. Los enlaces entre los nodos indican colaboraciones entre autores, es decir, la existencia de una línea entre dos nodos significa que esos dos autores han co-publicado al menos un documento. Los diferentes colores en el gráfico indican grupos de colaboración o clústeres dentro de la red, donde los autores dentro del mismo color tienden a colaborar más entre sí que con otros fuera del grupo.

Los clústeres identificados en el gráfico revelan agrupaciones de autores que colaboran estrechamente:

- Clúster 1, incluye a autores como Li Y, Zhang Y, Wang Y, Chen X, Wang L, Liu S, Yang J, Ma Y, Song L, Wang J, Yang S, Zhang W, Gao J, Guan Z, Hu MM, y Kang C. Este grupo muestra una alta conectividad interna y varios autores influyentes.
- Clúster 2, compuesto por autores como Li X, Liu X, Wang X, Wang Z, Huang J, Chau PYK, Huang H, y Jin Y, que, aunque menos centrales, tienen una notable presencia en la red.
- Clúster 3, incluye a Yang J, Ma Y, y Zhang X. Este grupo tiene una cohesión moderada y representa un subgrupo importante en la red.
- Clústeres 4 a 13, estos clústeres muestran autores con valores máximos de *closeness* (1.0) para todos sus integrantes, lo que indica una fuerte cohesión interna y una alta propensión a colaborar estrechamente dentro del clúster.

Este análisis de la red de colaboración destaca la estructura y dinámica de las relaciones entre investigadores en el campo del *Big Data* y el comportamiento del consumidor, proporcionando una visión clara de los principales actores y sus conexiones dentro de la comunidad científica.

Gráfico 3. 13. Red de Colaboración de autores en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024)

Nota: La tabla de datos de este gráfico se adjunta en el Anexo No. 4.6.

En el gráfico 3.14, se visualiza una Red de Colaboración de afiliaciones universitarias en el campo del "*Big Data*" y el "Comportamiento del Consumidor", abarcando el periodo de 1995 a 2023. Este gráfico, sirve para identificar las conexiones y colaboraciones entre diferentes instituciones académicas que contribuyen a la investigación en estas áreas.

Los nodos dentro del gráfico representan a las universidades, con el tamaño de cada nodo indicando la influencia o el volumen de publicaciones de la institución en los temas mencionados. Los enlaces entre los nodos demuestran la existencia de colaboraciones interuniversitarias, revelando que dos o más instituciones han co-publicado trabajos, lo que fortalece la red de conocimiento y la difusión científica.

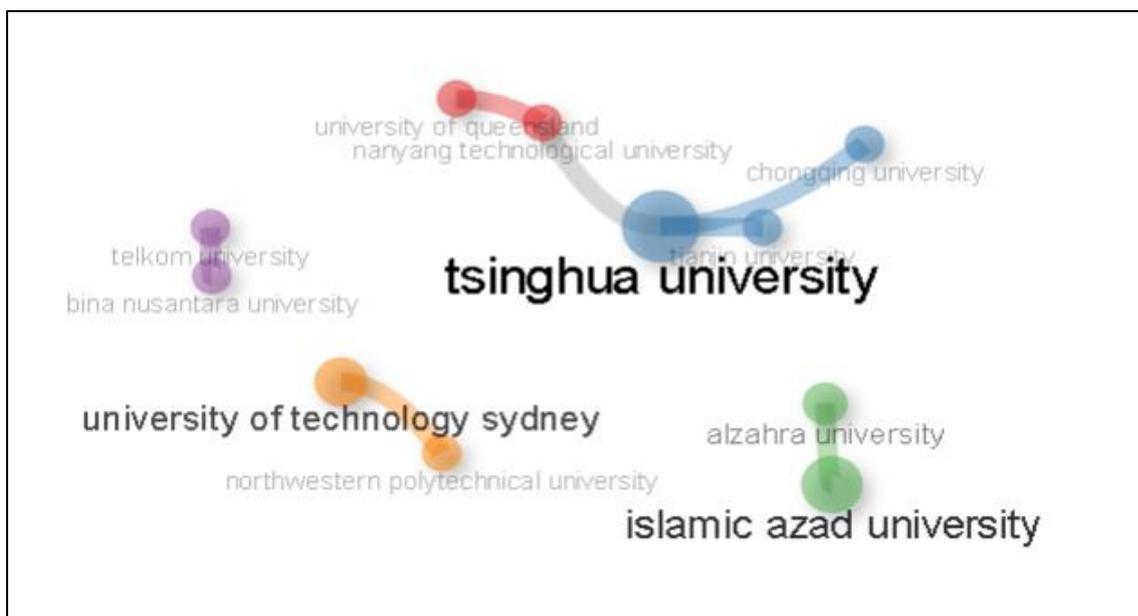
Las universidades centrales y de mayor tamaño en el gráfico, como Tsinghua University, indican su rol predominante en la investigación de *Big Data* y Comportamiento del Consumidor. Esta institución tiene un alto volumen de publicaciones y también forma el núcleo de la red de colaboración, mostrando numerosas conexiones con otras universidades,

lo que la posiciona como líder en la colaboración y generación de conocimiento en estos campos.

Los clústeres identificados en el gráfico revelan agrupaciones naturales de universidades que colaboran estrechamente. Por ejemplo, Tsinghua University, Central South University, Tianjin University y Chongqing University forman un clúster donde Tsinghua University domina, evidenciando su papel de liderazgo en la investigación. Además, Temple University, Northwestern University y Northeastern University conforman otro clúster, con Northeastern University destacándose por su alta centralidad y PageRank. Asimismo, Sichuan University, Chengdu University of Information Technology y Chongqing University of Posts and Telecommunications se agrupan, con Sichuan University como un nodo central en este grupo.

Este análisis resalta la interconexión y colaboración entre universidades en la investigación sobre *Big Data* y Comportamiento del Consumidor, ofreciendo una visión clara de los principales actores y sus vínculos en la comunidad académica.

Gráfico 3. 14. Red de Colaboración de afiliaciones en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

Nota: La tabla de datos de este gráfico se adjunta en el Anexo No. 4.8.

En el Gráfico 3.15, se presenta un Mapa de Colaboración entre Países que muestra las conexiones internacionales en la investigación sobre "Consumer Behavior" y "Big Data" durante el período analizado. El gráfico ilustra de manera visual las colaboraciones entre

diferentes naciones, utilizando líneas que conectan los países que han publicado trabajos de manera conjunta. La intensidad de estas colaboraciones se puede observar a través del grosor de las líneas, reflejando la frecuencia de las publicaciones compartidas entre los países.

Los países que participan más activamente en colaboraciones internacionales tienden a aparecer con nodos más grandes y colores más oscuros, lo que refleja su relevancia en la red de investigación. Por ejemplo, Estados Unidos y China son dos de los países más destacados, con una gran cantidad de líneas de conexión que los vinculan con otros países. Las colaboraciones más frecuentes, como las que existen entre Estados Unidos y China, se reflejan con líneas más gruesas, lo que indica un alto número de publicaciones conjuntas entre estas dos potencias de investigación.

En cuanto a las principales colaboraciones internacionales, Estados Unidos destaca como el país con el mayor número de colaboraciones globales. El gráfico muestra que tiene una fuerte relación con China, España, Hong Kong, y Corea, lo que refuerza su liderazgo en el campo de la investigación de *Big Data*. De manera similar, China mantiene relaciones importantes con el Reino Unido, Corea, y también colabora estrechamente con India, lo que demuestra la creciente influencia de Asia en el panorama global de la investigación.

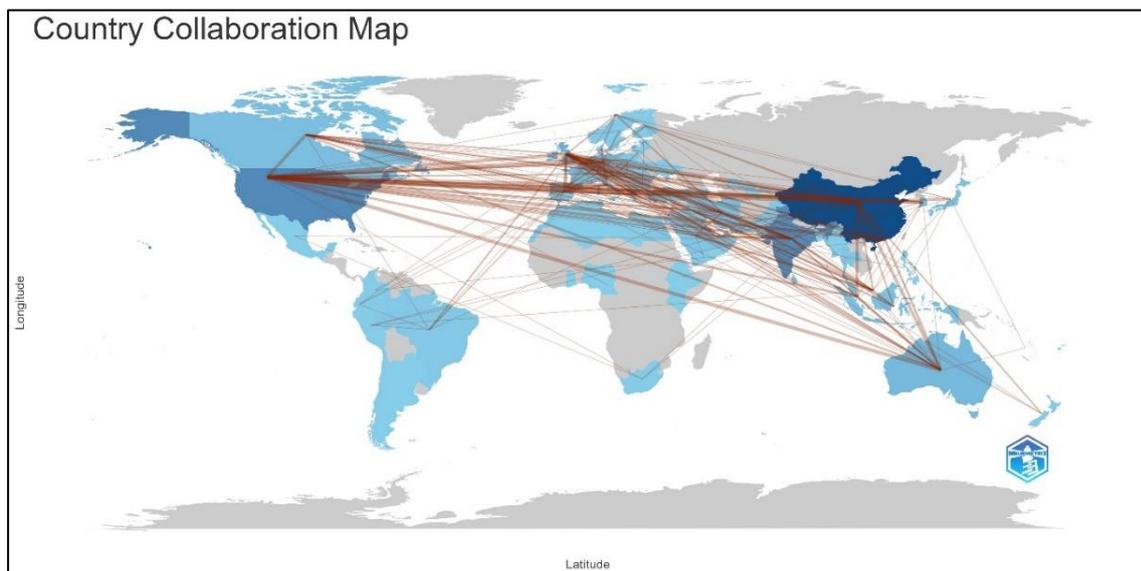
En el ámbito regional, Europa también muestra una densa red de conexiones internas, con países como Alemania, Francia, y el Reino Unido colaborando entre sí. Además, estos países tienen vínculos importantes con naciones fuera de Europa, como Estados Unidos y China. En Asia y Oceanía, China, India, y Australia lideran las colaboraciones, con conexiones significativas tanto entre ellos como con países en otros continentes, como Estados Unidos y los principales actores europeos.

América Latina, aunque menos densa en cuanto a colaboraciones internacionales, también está representada en el gráfico, con Brasil mostrando importantes vínculos de colaboración con países como Ecuador, Perú, Canadá, y Francia. Esto resalta el papel de Brasil como un líder en la investigación de *Big Data* en la región latinoamericana, además de sus crecientes colaboraciones con otros países fuera de la región.

El análisis del Mapa de Colaboración entre Países revela una tendencia clara hacia la colaboración transcontinental, especialmente entre América del Norte, Europa, y Asia, que forman los principales centros de producción científica en el campo de *Big Data* y el comportamiento del consumidor. Las colaboraciones tienden a concentrarse en países con fuertes capacidades de investigación, lo que refuerza el papel de las economías avanzadas en

liderar la producción y el intercambio de conocimiento científico a nivel mundial. Este mapa proporciona una visión integral de las dinámicas de colaboración internacional y destaca las principales redes geográficas que están impulsando el avance en estos campos de estudio.

Gráfico 3. 15. Mapa Mundial de Colaboración entre Países en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023



Fuente: Bibliometrix de R Studio, base de datos Scopus (2024)

Nota: La tabla de datos de este gráfico se adjunta en el Anexo No. 4.9.

3.1. Discusión de Resultados

Este capítulo presenta el análisis y la discusión de los resultados obtenidos a través de métodos bibliométricos aplicados al estudio del *Big Data* y el comportamiento del consumidor entre 1995 y 2023. Los hallazgos revelados proporcionan una comprensión detallada de las dinámicas de investigación, las tendencias emergentes, los actores clave y las estructuras de colaboración que caracterizan este campo de estudio. Cada sección está diseñada para responder a las preguntas de investigación planteadas en el marco introductorio y contribuye al cumplimiento de los objetivos específicos del estudio.

3.1.1. Evolución e Impacto de la Investigación sobre *Big Data* y Comportamiento del Consumidor

Los resultados evidencian un crecimiento exponencial en la producción científica relacionada con *Big Data* y comportamiento del consumidor, especialmente a partir de 2012. Este incremento coincide con la masificación de tecnologías digitales y la transformación del

consumidor en un agente activo en entornos digitales, conocido como Consumidor 2.0 (Muñoz y Aguilar, 2020). Este fenómeno refleja una evolución en el comportamiento del consumidor, donde las interacciones en línea y la generación de datos masivos se convierten en elementos clave.

Los resultados muestran un crecimiento exponencial en la producción científica relacionada con *Big Data* y comportamiento del consumidor, especialmente a partir de 2012. Este aumento coincide con la proliferación de tecnologías digitales y la creciente disponibilidad de datos generados por los consumidores en línea. La tasa de crecimiento anual del 21,51% refleja el interés creciente en este campo, indicando que el *Big Data* se ha convertido en una herramienta esencial para comprender las complejas dinámicas del comportamiento del consumidor en la era digital.

Este incremento en la producción científica ha tenido un impacto significativo en la comunidad académica y en la industria. El promedio de 16,5 citas por documento sugiere que las investigaciones en este campo son altamente relevantes y están siendo utilizadas como referencia para nuevos estudios. Además, la diversidad de tipos de documentos, que incluye artículos, libros, capítulos de libros y actas de conferencias, indica una amplia difusión del conocimiento y una transferencia efectiva entre la academia y la práctica empresarial.

La evolución temática identificada a través del análisis bibliométrico, evidencia una transición desde enfoques centrados en la minería de datos y las técnicas de procesamiento de información hacia aplicaciones más especializadas y orientadas al consumidor. La aparición de términos como "*Machine Learning*", "*Artificial Intelligence*" y "*Sentiment Analysis*" en años recientes sugiere que las investigaciones están incorporando tecnologías avanzadas para profundizar en la comprensión del comportamiento del consumidor.

Este crecimiento y evolución reflejan la necesidad de las empresas y organizaciones de adaptarse a un entorno competitivo y dinámico, donde el análisis de grandes volúmenes de datos permite identificar patrones y tendencias que antes eran difíciles de detectar. El impacto práctico de estos hallazgos se manifiesta en la capacidad de las empresas para personalizar ofertas, mejorar la experiencia del cliente y optimizar estrategias de marketing basadas en datos empíricos.

Además, la evolución de las citas promedio por año muestra fluctuaciones que podrían estar relacionadas con la rapidez con la que los nuevos estudios están siendo citados, lo que a su vez puede indicar la velocidad con la que se generan nuevos conocimientos y se incorporan a

la práctica. Esta dinámica sugiere que el campo es altamente innovador y está en constante actualización, lo que presenta desafíos y oportunidades para investigadores y profesionales.

3.1.2 Autores, Instituciones y Redes de Investigación Influyentes

El análisis bibliométrico ha permitido identificar a los autores más influyentes en el campo, siendo Wang Y., Li Y. y Zhang Y. quienes encabezan la lista en términos de producción científica y contribución al conocimiento. Estos autores han desempeñado un papel fundamental en el desarrollo y aplicación de técnicas de *Big Data* para el análisis del comportamiento del consumidor, lo que ha tenido un impacto significativo tanto en la academia como en la industria.

Las instituciones más destacadas incluyen la University of Technology Sydney y Tsinghua University, lo que indica una fuerte presencia de Australia y China en este campo de investigación. La prominencia de estas instituciones refleja su capacidad investigativa y también la inversión y prioridad que estos países han otorgado al desarrollo del *Big Data* y su aplicación en el análisis del comportamiento del consumidor.

Las redes de colaboración identificadas muestran una tendencia hacia la cooperación nacional más que internacional, especialmente en países como China e India. Esto puede deberse a factores culturales, lingüísticos o estratégicos, donde las instituciones prefieren fortalecer sus redes internas. Sin embargo, países como Estados Unidos y el Reino Unido presentan una mayor proporción de colaboraciones internacionales, lo que puede enriquecer la diversidad metodológica y teórica en el campo.

Es preciso identificar estas redes de colaboración para comprender cómo se difunde el conocimiento y cómo se generan las innovaciones. Las colaboraciones entre autores e instituciones pueden acelerar el desarrollo de nuevas técnicas y enfoques, y permiten abordar problemas complejos desde perspectivas multidisciplinarias. En términos prácticos, las empresas y organizaciones pueden beneficiarse al asociarse con estos investigadores e instituciones líderes para acceder a las últimas tecnologías y metodologías.

3.1.3 Principales Modelos y Estudios que Conforman la Base del Conocimiento

El análisis mediante el Tree of Science (ToS) ha sido fundamental para identificar la estructura conceptual y la evolución temática del campo. En la categoría ROOT, se han identificado estudios seminales como los de Fayyad et al. (1996) y Agrawal et al. (1993, 1994), que sentaron las bases teóricas y metodológicas de la minería de datos y el

descubrimiento de conocimiento en bases de datos. Estos trabajos proporcionan los fundamentos técnicos que han permitido el desarrollo de herramientas y técnicas para el análisis de grandes volúmenes de datos.

En la categoría TRUNK, se destacan investigaciones que han estructurado el campo del *Big Data* aplicado al comportamiento del consumidor. Estudios como los de Zhang M. (2022) y Li J. (2023) aplican las técnicas fundamentales a contextos específicos como el comercio electrónico y el análisis del comportamiento de compra en línea. Estos trabajos representan la consolidación de las técnicas de *Big Data* en aplicaciones prácticas que tienen un impacto directo en cómo las empresas entienden y se relacionan con sus clientes.

Las categorías BRANCH y LEAVES reflejan la diversificación y especialización temática del campo. En BRANCH, se encuentran estudios que integran conceptos como el perfilado de clientes, el marketing sostenible y la privacidad de datos. Por ejemplo, investigaciones como las de Lăzăroiu (2023) abordan la intersección entre *Big Data*, inteligencia artificial y *fintech*, mostrando cómo estas tecnologías pueden aplicarse en sectores emergentes.

En LEAVES, se identifican tendencias emergentes y estudios que exploran nuevas aplicaciones y desafíos, como el impacto de la inteligencia artificial en el consumo sostenible (Cao P., 2023) y el análisis del comportamiento del consumidor en el contexto del Internet de las Cosas (Wei L., 2023). Estos trabajos indican hacia dónde se dirige la investigación y cuáles serán las áreas de enfoque en el futuro.

La vinculación con el marco teórico es evidente en el uso de técnicas de *Big Data* para profundizar en las teorías del comportamiento del consumidor. La incorporación de tecnologías avanzadas permite analizar factores emocionales, sociales y psicológicos que influyen en las decisiones de compra, alineándose con teorías como la del condicionamiento operante de Skinner (1938) y la teoría de la perspectiva de Kahneman y Tversky (1979). Estas aplicaciones prácticas enriquecen el marco teórico al proporcionar evidencia empírica y nuevas formas de medir y entender el comportamiento del consumidor.

3.1.4 Tendencias Actuales y Emergentes en la Literatura

Las temáticas emergentes identificadas, como el análisis de sentimientos, la inteligencia artificial y el Internet de las Cosas, reflejan una evolución hacia enfoques más sofisticados y multidisciplinarios en el estudio del comportamiento del consumidor. Estas tendencias se relacionan con el marco teórico en varios aspectos.

El análisis de sentimientos y la minería de opiniones permiten entender las emociones y percepciones de los consumidores, elementos clave en la teoría del comportamiento del consumidor que destaca la influencia de factores psicológicos y emocionales en las decisiones de compra (Solomon, 2006). La capacidad de analizar grandes volúmenes de datos no estructurados, como comentarios en redes sociales, proporciona una visión más completa y precisa de las actitudes del consumidor.

La inteligencia artificial y el aprendizaje automático están transformando la forma en que se modela y predice el comportamiento del consumidor. Estas tecnologías permiten manejar la complejidad y diversidad de los datos actuales, abordando las limitaciones de los modelos tradicionales basados en la racionalidad perfecta. Esto se alinea con la teoría de la racionalidad limitada de Simon (1956), reconociendo que las decisiones de los consumidores están influenciadas por múltiples factores y que los modelos deben adaptarse a esta realidad.

El Internet de las Cosas amplía el alcance de los datos disponibles, permitiendo un análisis más detallado del comportamiento en contextos físicos y digitales. Esto crea oportunidades para integrar teorías del comportamiento del consumidor con datos contextuales y en tiempo real, enriqueciendo la comprensión de cómo y por qué los consumidores toman ciertas decisiones.

La preocupación por la privacidad de los datos y la ética en su manejo, reflejada en las tendencias emergentes, también conecta con el marco teórico al resaltar la importancia de la confianza y la percepción de seguridad en las relaciones entre consumidores y empresas. La teoría de la responsabilidad social corporativa y la ética en los negocios subraya la necesidad de prácticas transparentes y responsables, aspectos que son cada vez más relevantes en el contexto del *Big Data*.

3.1.5 Aplicaciones Prácticas y Relevancia para la Industria

Los hallazgos de esta investigación tienen implicaciones prácticas significativas para la industria y el mercado. La evolución y tendencias identificadas permiten a las empresas:

- Mejorar la personalización y segmentación, utilizando técnicas avanzadas de *Big Data* y aprendizaje automático, las empresas pueden desarrollar perfiles más precisos de sus clientes y ofrecer productos y servicios que se ajusten mejor a sus necesidades y preferencias.

- Optimizar estrategias de marketing, el análisis de sentimientos y la minería de opiniones permiten adaptar mensajes y campañas de marketing de manera más efectiva, aumentando el impacto y retorno de la inversión.
- Anticipar tendencias y comportamientos, las herramientas de predicción basadas en *Big Data* permiten a las empresas anticipar cambios en el mercado y adaptarse rápidamente, manteniendo una ventaja competitiva.
- Abordar desafíos éticos y regulatorios, la creciente atención a la privacidad y ética en el manejo de datos requiere que las empresas implementen políticas y prácticas que cumplan con regulaciones y respondan a las expectativas de los consumidores, fortaleciendo la confianza y lealtad.
- Innovar en productos y servicios, la integración de tecnologías como el Internet de las Cosas y la inteligencia artificial abre nuevas oportunidades para desarrollar ofertas innovadoras y modelos de negocio disruptivos.

Estos aspectos resaltan la importancia de que las empresas estén al tanto de las tendencias y avances en el campo del *Big Data* y el comportamiento del consumidor, y que inviertan en capacidades y alianzas que les permitan aprovechar estos conocimientos.

Tabla 3. 8. Síntesis de técnicas y enfoques analíticos en el campo del *Big Data* y Comportamiento del Consumidor en el periodo 1995-2023

Herramienta de Análisis	Resultados Clave
Nube de Palabras	Palabras clave más frecuentes:
	<i>Data Mining, Consumer Behavior, Sales, Big Data, Customer Behavior, Machine Learning, Artificial Intelligence, Data Analytics, E-commerce, Customer Satisfaction, Social Networking (Online).</i>
	Estas palabras clave resaltan las tecnologías y conceptos principales en la recolección y análisis de datos, subrayando su uso extendido en la personalización de servicios y mejora de la experiencia del usuario.
Red de Co-ocurrencia de Palabras Clave	Clústeres identificados:
	Clúster 1 (Rojo):
	<i>- Machine Learning, Artificial Intelligence, Deep Learning, Data Analytics, Big Data, Data Mining, Clustering, Classification, Predictive Analytics.</i>
	Enfoque: Técnicas avanzadas de análisis de datos y su aplicación en marketing y estudio del comportamiento del consumidor.
	Clúster 2 (Azul):
	<i>- Social Media, Consumer Behavior, E-commerce, Marketing, Customer Satisfaction, Customer Behavior, Sentiment Analysis, Online Shopping.</i>
Mapa de Estructura Conceptual	Agrupaciones conceptuales identificadas:
	Tecnología y Análisis de Datos:
	<i>- Cloud Computing, Big Data Analytics, Data Mining, Business Intelligence, MapReduce, Hadoop.</i>
	Aplicaciones en Redes Sociales y Comportamiento del Consumidor:
	<i>- Twitter, Social Media, Consumer Behavior, Mobile Marketing, Sentiment Analysis, Opinion Mining.</i>
	Estas agrupaciones reflejan cómo las tecnologías de <i>Big Data</i> se aplican en diversas áreas para analizar y predecir el comportamiento del consumidor.
Diagrama de Evolución Temática (Diagrama de Sankey)	Evolución de temas entre periodos:
	1995-2012:
	<i>- Temas centrales: Data Mining, Web Mining, Fraud Detection, Business Intelligence.</i>
	2013-2023:
	<i>- Emergencia de nuevos temas: Big Data Analytics, Sentiment Analysis, Internet of Things, Artificial Intelligence, Machine Learning, Social Media, E-commerce.</i>
La transición indica un enfoque más especializado y práctico del <i>Big Data</i> en el análisis del comportamiento del consumidor.	
Mapa Temático (Basado en Relevancia y Desarrollo)	Distribución de temas en cuadrantes:
	Temas Emergentes o en Declive (Inferior Izquierdo):
	<i>- Trust, Purchase Intention.</i>
	Temas Básicos (Inferior Derecho):
	<i>- Big Data, Machine Learning, Consumer Behavior, E-commerce, Data Analytics.</i>
Temas de Nicho (Superior Izquierdo):	
<i>- Neural Network, Fraud Detection, Time Series, Behavioral Analytics.</i>	

	Temas Motores (Superior Derecho):
	- <i>Data Mining, Customer Behavior, Clustering, Classification, Customer Relationship Management.</i>
	Este mapa muestra la relevancia y el desarrollo de los temas dentro del campo, identificando áreas centrales y especializadas.
Análisis Tree of Science (ToS)	Clasificación de la literatura:
	ROOT (Raíz):
	- Trabajos fundacionales que establecen las bases teóricas y metodológicas, como <i>Fayyad U.M. (1996)</i> y <i>Agrawal R. (1993, 1994)</i> , esenciales para las técnicas de minería de datos.
	TRUNK (Tronco):
	- Estudios que han estructurado el campo, abordando categorías como minería de datos y <i>Big Data</i> , con autores como <i>Zhang M. (2022)</i> y <i>Wang Y. (2021)</i> explorando aplicaciones prácticas.
	BRANCHES (Ramas):
	- Especializaciones temáticas que muestran diversificación, abordando tópicos como <i>Emotion Analysis, Face Recognition, Customer Emotions, Diet, Exhibitions</i> , reflejando la diversificación del campo.
	LEAVES (Hojas):
- Investigaciones recientes que amplían temas actuales, integrando tecnologías avanzadas como <i>IA</i> y <i>IoT</i> , representadas por <i>Cao P. (2023)</i> y <i>Wei L. (2023)</i> .	

Elaborado por el Autor a partir de datos obtenidos en Bibliometrix y Tree of Science (2024).

Conclusiones y Recomendaciones

Conclusiones

El presente estudio subraya la influencia del *Big Data* en el análisis del comportamiento del consumidor, destacando cómo la capacidad de manejar grandes volúmenes de datos ha revolucionado las estrategias comerciales y el ámbito de investigación académica. A través de un meticuloso análisis bibliométrico realizado con herramientas como Bibliometrix y Tree of Science, este trabajo ha identificado tendencias claves y ha mapeado la evolución de la investigación publicada hasta el 2023.

- El análisis bibliométrico reveló un crecimiento significativo en la producción científica en este campo, especialmente a partir de 2012, coincidiendo con la expansión de las tecnologías digitales y el aumento en la generación de datos por parte de los consumidores. Con una tasa de crecimiento anual del 21,51% y un promedio de 16,5 citas por documento, se evidencia un incremento tanto en la cantidad de investigaciones, como en su calidad e impacto en la comunidad académica. Este crecimiento refleja la creciente importancia del *Big Data* en el estudio del comportamiento del consumidor y su relevancia en la agenda de investigación actual.
- Se identificaron a Wang Y., Li Y. y Zhang Y. como los autores más productivos e influyentes, así como a instituciones líderes como la University of Technology Sydney y Tsinghua University. China y Estados Unidos emergen como los países con mayor producción científica en este ámbito. Sin embargo, se observó que países como el Reino Unido a diferencia de China, presentan una mayor propensión a colaboraciones internacionales, lo que podría potenciar la diversidad y la innovación en futuras investigaciones.
- Las bases teóricas y metodológicas del campo de estudio del *Big Data* aplicado al comportamiento del consumidor tienen como pilares fundamentales los trabajos de Fayyad et al. (1996) y Agrawal et al. (1993, 1994) sobre minería de datos y descubrimiento de conocimiento en bases de datos. Estas investigaciones han proporcionado los cimientos conceptuales y prácticos que han guiado la evolución del campo. A través de la metodología de Tree of Science, se ha visualizado cómo estas teorías iniciales han dado lugar a nuevas líneas de investigación, adaptándose y expandiéndose hacia áreas emergentes como el análisis predictivo, la inteligencia artificial y la privacidad de datos. Estas tendencias reflejan una evolución clara hacia enfoques más avanzados y multidisciplinarios, caracterizados por la integración de

tecnologías innovadoras como el aprendizaje automático, el análisis de sentimientos y el Internet de las Cosas, lo que permite una comprensión más profunda y precisa del comportamiento del consumidor en contextos cada vez más complejos.

- Las tendencias actuales apuntan hacia la integración de tecnologías avanzadas como inteligencia artificial, aprendizaje automático y análisis de sentimientos en el estudio del comportamiento del consumidor. Temáticas emergentes como el Internet de las Cosas, el análisis predictivo y las preocupaciones éticas sobre privacidad y manejo de datos indican que el campo está evolucionando hacia enfoques más sofisticados y multidisciplinarios. Estas tendencias sugieren que las futuras investigaciones se orientarán hacia la personalización de experiencias, la anticipación de necesidades del consumidor y la implementación de prácticas éticas en el uso de datos, lo que permitirá a las empresas y académicos desarrollar estrategias más efectivas y centradas en el consumidor.
- El impacto del *Big Data* sobre el comportamiento del consumidor ha sido significativo, con aplicaciones prácticas que permiten a las empresas personalizar ofertas, predecir comportamientos de compra y mejorar la experiencia del cliente. Sin embargo, también se identificaron desafíos importantes, como las deficiencias en la calidad de los metadatos en las publicaciones académicas y la necesidad de fortalecer las redes de colaboración internacional. Estos problemas deben ser abordados para maximizar el potencial de las investigaciones futuras y garantizar que el análisis del *Big Data* se realice de manera ética, transparente y responsable.

Recomendaciones

Las recomendaciones que se presentan a continuación están diseñadas para abordar estos retos y aprovechar las oportunidades identificadas, con el fin de enriquecer futuras investigaciones y prácticas en el campo.

- Es esencial que revistas y editoriales implementen políticas rigurosas para asegurar que los metadatos de las publicaciones sean completos y precisos. Esto incluye campos como "*Keywords Plus*" y "*Corresponding Author*". Una mejora en la calidad de los metadatos facilitará análisis bibliométricos más profundos y precisos, y aumentará la visibilidad y accesibilidad de las investigaciones, beneficiando a toda la comunidad académica.

- Es esencial promover y ampliar las redes de colaboración entre autores e instituciones de diferentes países y disciplinas. La creación de consorcios internacionales y la participación en proyectos colaborativos enriquecerán las perspectivas teóricas y metodológicas, impulsando la innovación y abordando desafíos complejos desde enfoques integrales. Dirigir esfuerzos de investigación hacia temáticas emergentes como "*Internet of Things*", "*Artificial Intelligence*", "*Machine Learning*" y "*Big Data Analytics*" permitirá explorar nuevas aplicaciones y metodologías en el análisis del comportamiento del consumidor, incluyendo el impacto de tecnologías disruptivas en las decisiones de consumo y la experiencia del cliente.
- Dado el aumento de preocupaciones relacionadas con la privacidad y la ética en el manejo de datos, es recomendable que los investigadores y las empresas incorporen consideraciones éticas desde el diseño de la investigación y el desarrollo de tecnologías. Esto fortalecerá la confianza del público y asegurará el uso responsable y sostenible del *Big Data*. Además, se sugiere aprovechar el *Big Data* no solo como una herramienta analítica, sino también como un medio para probar, refinar y desarrollar teorías existentes y emergentes sobre el comportamiento del consumidor.
- Reconociendo los posibles sesgos en los estudios bibliométricos, como el sesgo de selección de bases de datos, el sesgo lingüístico y el sesgo de citación, se sugiere que futuras investigaciones consideren la inclusión de múltiples bases de datos, publicaciones en diversos idiomas y análisis críticos de las prácticas de citación. Esto permitirá obtener una visión más equilibrada y representativa del campo, asegurando que las conclusiones reflejen verdaderamente la diversidad y amplitud de la investigación global.
- Se recomienda ampliar la investigación hacia sectores específicos como la salud, la educación, el medio ambiente y la economía colaborativa, y considerar diferentes contextos culturales. Esto permitirá comprender mejor cómo el *Big Data* puede aplicarse de manera efectiva en diversos entornos y responder a necesidades locales y globales. Finalmente, es importante que las investigaciones incluyan evaluaciones del impacto y la sostenibilidad de las aplicaciones de *Big Data* en el comportamiento del consumidor, analizando los efectos sociales, éticos y ambientales para asegurar el desarrollo de estrategias responsables y contribuir al bienestar de la sociedad.

Referencias

- Agrawal, R., T. Imielinski y A. Swami. 1993. "Mining Association Rules Between Sets of Items in Large Databases." En *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, 207–216. Nueva York: ACM.
- Aksnes, Dag W., Fredrik N. Piro y Kristoffer Rørstad. 2019. "Gender Gaps in International Research Collaboration: A Bibliometric Approach." *Scientometrics* 120(2): 747–774.
- Ahmadon, M. A. B. y M. H. F. Rahiman. 2021. "Mining Consumer Services Based on User Preference with Associative and Process Mining." *International Journal of Advanced Computer Science and Applications* 12(5): 324–330.
- Ambrosio, A. 2013. "¿Qué es el crossuser?" *Qué Aprendemos Hoy*. Consultado el 11 de febrero de 2024. <http://queaprendemoshoy.com/que-es-el-crossuser/>.
- Aoujil, Zakaria, Mohamed Hanine, Emmanuel Soriano y Abdus Samad. 2023. "Artificial Intelligence and Behavioral Economics: A Bibliographic Analysis of Research Field." *IEEE Access* 11: 139367–139394.
- Aragona, Biagio. 2022. "Tipos de Big Data y Análisis Sociológico: Usos, Críticas y Problemas Éticos." *EMPIRIA. Revista de Metodología de Ciencias Sociales* 53 (enero-abril): 15–30. <https://doi.org/10.1016/empiria.53.2022.32610>.
- Aria, Massimo y Corrado Cuccurullo. 2017. "Bibliometrix: An R-tool for Comprehensive Science Mapping Analysis." *Journal of Informetrics* 11(4): 959–975.
- Awate, A. S. y P. N. Mahalle. 2023. "Understanding Customer Behaviour: A Comprehensive Survey of Segmentation and Classification Techniques in the Age of Big Data." *Journal of Data, Information and Management* 5(1): 1–19.
- Bawack, Ransome, Samuel Wamba, Kevin André y Shahriar Akter. 2022. "Artificial Intelligence in E-Commerce: A Bibliometric Study." *Electronic Markets* 32: 297–338.
- Bienert, Igor, Rogério Carvalho, Pedro Beraldo y Carlos Caramori. 2015. "Bibliometric Indexes, Databases and Impact Factors in Cardiology." *Revista Brasileira de Cirurgia Cardiovascular* 30(3): 254–259.
- Blackwell, Roger D., Paul W. Miniard y James F. Engel. 2005. *Consumer Behavior*. 10ª ed. Mason, Ohio: Thomson/South-Western.
- Bordons, María y Ángeles Zulueta. 1999. "Evaluation of Scientific Activity through Bibliometric Indicators." *Revista Española de Cardiología* 52(10): 790–800.
- Broadus, R. 1987. "Toward a Definition of 'Bibliometrics'." *Scientometrics* 12: 373–379.
- Burnham, J. F. 2006. "Scopus Database: A Review." *Biomedical Digital Libraries* 3: 1. <https://doi.org/10.1186/1742-5581-3-1>.
- Callzilla. 2023. "15 Customer Behavior Models and Examples of Implementation." 30 de junio de 2023. <https://www.callzilla.cx/15-customer-behavior-models-and-examples-of-implementation/>. Consultado el 30 de febrero de 2024.
- Cao, P., X. Duan y Z. Yang. 2023. "The Impact of Artificial Intelligence Technology Stimuli on Sustainable Consumption Behavior: Evidence from Ant Forest Users in China." *Sustainability* 15(5): 1–19.

- Chang, Y., M. Huang y C. Lin. 2015. "Evolution of Research Subjects in Library and Information Science Based on Keyword, Bibliographical Coupling, and Co-citation Analyses." *Scientometrics* 105(3): 2071–2078.
- Chavez, Henry, Fernando Martín y María Belén Albornoz. 2022. "'Big Data' Research: A Bibliometric Analysis of the Scopus Database, 2009–2019." *Journal of Scientometric Research* 11(1): 64–78.
- Chen, L.-S., A. S. P. Chen y T.-H. Hsu. 2013. "A Study on Review Manipulation Classification Using Decision Tree." *Advanced Science Letters* 19(3): 851–854.
- Chen, Min, Shiwen Mao y Yunhao Liu. 2014. "Big Data: A Survey." *Mobile Networks and Applications* 19(2): 171–209.
- Cobo, M. J., A. G. López-Herrera, E. Herrera-Viedma y F. Herrera. 2011. "Science Mapping Software Tools: Review, Analysis, and Cooperative Study among Tools." *Journal of the American Society for Information Science and Technology* 62(7): 1382–1402. <https://doi.org/10.1002/asi.21525>.
- Colmenarejo Fernández, Rosa. 2018. "Ética Aplicada a la Gestión de Datos Masivos." *Anales de la Cátedra Francisco Suárez* 52: 113–129. https://www.academia.edu/4896164/Information_Ethics_II_Towards_a_Unified_Taxonomy.
- Core of Science. 2020. "Tree of Science." <https://tos.coreofscience.com/>. Consultado el 2 de julio de 2024.
- Coronado, Linda. 2019. *Teoría de Comportamiento del Consumidor*. Iquitos: UNAP.
- Dans, Enrique. 2012. "Big Data: Una Perspectiva Histórica." 12 de mayo de 2012. <https://www.enriquedans.com/2012/05/big-data-una-perspectiva-historica.html>. Consultado el 18 de enero de 2024.
- Donthu, N., S. Kumar, D. Mukherjee, N. Pandey y W. Lim. 2021. "How to Conduct a Bibliometric Analysis: An Overview and Guidelines." *Journal of Business Research* 133: 285–296.
- Elyusufi, Z., A. Erraissi y M. Bakkoury. 2018. "Customer Profiling Using CEP Architecture in a Big Data Context." En *Proceedings of the Mediterranean Conference on Information & Communication Technologies*, 221–228.
- Elyusufi, Z., A. Erraissi y M. Bakkoury. 2019. "Combining CRM Strength and Big Data Tools for Customers Profile Analysis." *International Journal of Computer Applications* 177(14): 23–29.
- Erevelles, Sunil, Nobuyuki Fukawa y Linda Swayne. 2016. "Big Data Consumer Analytics and the Transformation of Marketing." *Journal of Business Research* 69(2): 897–904.
- Falagas, Matthew E., Eugenia I. Pitsouni, George A. Malietzis y Georgios Pappas. 2008. "Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and Weaknesses." *FASEB Journal* 22(2): 338–342.
- Fawcett, Tom y Foster Provost. 1997. "Adaptive Fraud Detection." *Data Mining and Knowledge Discovery* 1(3): 291–316. <https://doi.org/10.1023/A:1009700419189>.
- Ferrer, Luigi. 2018. "Comportamiento del Consumidor 2.0: Nuevas Realidades en Entornos Digitales." *Marketing Visionario* 7(1): 141–155.

- Gan, Ye, Duo Li, Nicola Robinson y Ping Liu. 2022. "Practical Guidance on Bibliometric Analysis and Mapping Knowledge Domains Methodology – A Summary." *Journal of Informetrics* 16(4): 102203.
- Ganhewa, N. B. y W. P. J. D. Bandara. 2021. "Sales Optimization Solution for Fashion Retail." *Journal of Computer Science* 17(7): 620–633.
- Garfield, Eugene. 2004. "Citation Analysis as a Method of Historical Research into Science." *The Garfield Library*. <http://www.garfield.library.upenn.edu/ci/chapter7.pdf>. Consultado el 10 de enero de 2024.
- Goh, Khim-Yong, Cheng-Suang Heng y Zhijie Lin. 2013. "Social Media Brand Community and Consumer Behavior: Quantifying the Relative Impact of User- and Marketer-Generated Content." *Information Systems Research* 24(1): 88–107. <https://doi.org/10.1287/isre.1120.0469>.
- Gómez, Ángel. 2021. "Big Data, un Sistema de Gestión de Datos." *TA University*. <https://www.tauniversity.org/big-data-un-sistema-de-gestion-de-datos>.
- Gómez, María y María Sequería. 2015. *Estudio del Comportamiento del Consumidor*. Managua: Universidad Nacional Autónoma de Nicaragua.
- Granda, José, Adolfo Alonso, Francisco García, Segismundo Solano, Carlos Jiménez y Rafael Aleixandre. 2013. "Ciertas Ventajas de Scopus sobre Web of Science en un Análisis Bibliométrico sobre Tabaquismo." *Revista Española de Documentación Científica* 36(2).
- Haba, Herman, Christophe Bredillet y Omkar Dastane. 2023. "Green Consumer Research: Trends and Way Forward Based on Bibliometric Analysis." *Cleaner and Responsible Consumption*.
- Han, Jiawei, Micheline Kamber y Jian Pei. 2001. *Data Mining: Concepts and Techniques*. San Francisco: Morgan Kaufmann.
- Hernández, Emily y Néstor D. Duque. 2017. "Big Data: An Exploration of Research, Technologies and Application Cases." *Ciencia e Ingeniería Neogranadina* 20(39).
- Hervé, Catherine y Etienne Mullet. 2009. "Age and Factors Influencing Consumer Behaviour." *International Journal of Consumer Studies* 33(3): 302–308.
- Hofacker, Charles, Edward Malthouse y Fareena Sultan. 2016. "Big Data and Consumer Behavior: Imminent Opportunities." *Journal of Consumer Marketing* 33(2): 89–97.
- Jiménez, E. 1993. "Bibliometría (Análisis Bivariante)." *Revista Española de Documentación Científica* 16(4).
- Kahneman, Daniel y Amos Tversky. 1973. "On the Psychology of Prediction." *Psychological Review* 80(4): 237–251.
- Kahneman, Daniel y Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica* 47: 263–291.
- Khan, Khalid S., Regina Kunz, Jos Kleijnen y Gerd Antes. 2003. "Five Steps to Conducting a Systematic Review." *Journal of the Royal Society of Medicine* 96(3): 118–121.
- Kotler, Philip y Kevin Keller. 2006. *Dirección de Marketing*. Traducido por Clara Rivera. México: Pearson.

- Kusnetzky, Dan. 2010. "What Is Big Data?" *ZDNet*, 15 de febrero de 2010. <https://www.zdnet.com/article/what-is-big-data/>. Consultado el 25 de enero de 2024.
- Laney, Doug. 2001. "3D Data Management: Controlling Data Volume, Velocity, and Variety." *Application Delivery Strategies* 6.
- Lazaroiu, George y Ewa Rogalska. 2023. "How Generative Artificial Intelligence Technologies Shape Partial Job Displacement and Labor Productivity Growth." *Oeconomia Copernicana* 14(3): 703–706. <https://doi.org/10.24136/oc.2023.020>.
- Li, J. y Y. Wang. 2023. "A Deep Mining Method for Consumer Behaviour Data of E-Commerce Users Based on Clustering and Deep Learning." *Journal of Organizational and End User Computing* 35(2): 1–15.
- Liu, Stephanie y Anna Mattila. 2017. "Airbnb: Online Targeted Advertising, Sense of Power, and Consumer Decisions." *International Journal of Hospitality Management* 60: 33–41.
- López, M. B. V. y J. R. R. Rosas. 2023. "Application of a Data Mining Model to Predict Customer Defection: Case of a Telecommunications Company in Peru." *International Journal of Data Analysis Techniques and Strategies* 15(1): 57–70.
- Macías, Adrián. 2016. "El Crossuser." *Métodos de Información* 7(12): 5–16.
- MacRoberts, Michael H. y Barbara R. MacRoberts. 2018. "The Mismeasure of Science: Citation Analysis." *Journal of the Association for Information Science and Technology* 69(3): 474–482.
- Malpica, William, Melva Gómez y Álvaro Villa. 2021. "Study on Marketing Research and Its Importance in Export Processes: Bibliometric and Thematic Analysis in Scopus." *Revista Publicando* 8(2): 91–103.
- Mick, David, Susan Broniarczyk y Jonathan Haidt. 2024. "Choose, Choose, Choose, Choose, Choose, Choose, Choose: Emerging and Prospective Research on the Deleterious Effects of Living in Consumer Hyperchoice." *Journal of Business Ethics* 52: 207–211.
- Mongeon, Philippe y Adèle Paul-Hus. 2015. "The Journal Coverage of Web of Science and Scopus: A Comparative Analysis." *Scientometrics* 106(1): 213–228. <https://doi.org/10.1007/s11192-015-1765-5>.
- Montero, Donny y Omar El Kadi. 2017. "Marca Personal para el Posicionamiento de los Creadores Visuales." *Marketing Visionario* 5(2).
- Mostafa, Mohamed M. 2013. "More Than Words: Social Networks' Text Mining for Consumer Brand Sentiments." *Expert Systems with Applications* 40(10): 4241–4251. <https://doi.org/10.1016/j.eswa.2013.01.019>.
- Mullainathan, Sendhil y Eldar Shafir. 2013. *Scarcity: Why Having Too Little Means So Much*. Nueva York: Times Books – Henry Holt and Company, LLC.
- Muñoz, Fabrizio y Stephanny Aguilar. 2020. *Marketing Digital's Influence on the New Buy and Sell Tendencies of the Consumer 2.0*. Arequipa: Universidad Católica San Pablo.
- Muniady, Rajennd, Abdullah Al-Mamun, P. Yukthamarani Permarupan y Noor Raihani Binti Zainol. 2014. "Factors Influencing Consumer Behavior: A Study Among University Students in Malaysia." *Asian Social Science* 10(9): 18–25.

- Orellana, Fernando. 2022. *Theory of Consumer Behavior, Production and Costs*. Vol. 4. Guayaquil: E-IDEA 4.0 Revista Multidisciplinar.
- Ouyang, Wei et al. 2020. "Evolutionary Overview of Consumer Health." *Journal of Consumer Health on the Internet* 24(2): 107–123.
- Páez, Erika y Andrés Monroy. 2020. *Implementación de un Modelo de Análisis de Sentimientos con Respecto a la JEP Basado en Minería de Datos en Twitter*. Bogotá: Universidad Católica de Colombia.
- Paredes, Antonio. 2015. "Big Data: State of the Art." *International Journal of Information Systems and Software Engineering for Big Companies* 2(1): 38–59.
- Pavlov, Ivan. 1903. "The Experimental Psychology and Psychopathology of Animals." En *Proceedings of the 14th International Medical Congress*, 23–30.
- Peralta, María, Maylín Frías y Orlando Gregorio. 2015. "Criterios, Clasificaciones y Tendencias de los Indicadores Bibliométricos en la Evaluación de la Ciencia." *Revista Cubana de Información en Ciencias de la Salud* 26(3): 290–309.
- Raiteri, Melisa. 2016. *El Comportamiento del Consumidor Actual*. Mendoza: UNCUYO.
- Rangaiah, Mallika. 2023. "5 Factors Influencing Consumer Behavior." *AnalyticSteps*, 21 de noviembre de 2023. <https://www.analyticsteps.com/blogs/5-factors-influencing-consumer-behavior>. Consultado el 10 de febrero de 2024.
- Recalde, María Fernanda, Lizbeth Suárez y Juan Carlos Páez. 2021. *El Comportamiento del Consumidor en el Ámbito Económico y Psicológico*. Sangolquí: Revista Conectividad.
- Reina, María Dolores y Juan Carlos Rodríguez. 2023. "Main Theoretical Consumer Behavioural Models: A Review from 1935 to 2021." *Heliyon* 9.
- Reisch, Lucia y Min Zhao. 2017. "Behavioural Economics, Consumer Behaviour, and Consumer Policy." *Behavioural Public Policy* 1(2): 190–206.
- Rodríguez, María. 2021. "Data Visualization: Sankey Diagram." *Datasketch*, 29 de septiembre de 2021. <https://www.datasketch.co/es/blog/data-visualization-sankey-diagram/>. Consultado el 3 de julio de 2024.
- Saleem, Sumayya, Michal Perlman, Elizabeth Dhuey y Linda Blanca. 2024. "Understanding 21st Century Skills Needed in Response to Industry 4.0: Exploring Scholarly Insights Using Bibliometric Analysis." En *Telematics and Informatics Reports*.
- Scopus. 2014. "The Scopus h-index, What's It All About? Part I." 14 de agosto de 2014. <https://blog.scopus.com/posts/the-scopus-h-index-what-s-it-all-about-part-i>. Consultado el 23 de junio de 2024.
- Simon, Herbert. 1956. "Rational Choice and the Structure of the Environment." *Psychological Review* 63: 129–138.
- Skinner, B. F. 1938. *The Behavior of Organisms: An Experimental Analysis*. Minneapolis: Department of Psychology, University of Minnesota.
- Small, Henry. 1973. "Co-citation in the Scientific Literature: A New Measure of the Relationship Between Two Documents." *Journal of the American Society for Information Science* 24(4): 265–269.
- Solomon, Michael R. 2006. *Comportamiento del Consumidor*. 7ª ed. México D.F.: Pearson Educación.

- Stávková, Jana, Ladislav Stejskal y Zuzana Toufarová. 2008. "Factors Influencing Consumer Behaviour." *Agricultural Economics – Czech* 54(6): 274–284.
- Sun, Zhaohao, Lizhe Sun y Kenneth Strang. 2018. "Big Data Analytics Services for Enhancing Business Intelligence." *Journal of Computer Information Systems* 58(2): 162–169.
- Tapia, Mauro, Freddy Alvarado y Gary Cox. 2019. "Factores Externos que Influyen en el Comportamiento de los Consumidores." *Dominio de las Ciencias* 5(2).
- Thaler, Richard. 1985. "Mental Accounting and Consumer Choice." *Marketing Science* 4(3): 199–214.
- Thangavel, Packiara y Bibhas Chandra. 2023. "Two Decades of M-Commerce Consumer Research: A Bibliometric Analysis Using R Biblioshiny." *Sustainability* 15(15).
- Thayyib, P.V. et al. 2023. "State-of-the-Art of Artificial Intelligence and Big Data Analytics." *Sostenibilidad*.
- Thorndike, Edward Lee. 1911. *Animal Intelligence: Experimental Studies*. Míchigan: Universidad de Míchigan.
- UNESCO. 2020. *Ciudadanía Digital - Curriculum para la Formación Docente*. Montevideo: UNESCO.
- Valencia, Daniel, Sebastian Robledo, Ricardo Pinilla, Nestor Duque y Gerard Olivar. 2020. "Algoritmo SAP para Análisis de Citaciones: Una Mejora para Tree of Science." *Ingeniería e Investigación* 40: 45–49.
- Valencia, William, Juan E. López y Valentina López. 2020. "Tree of Science (ToS) Visualizations of Scientific Literature." *Ingeniería y Universidad* 24(2): 1–13.
- Van Leeuwen, Thed N., Henk F. Moed, Robert J. W. Tijssen, Martijn S. Visser y Anthony F. J. van Raan. 2001. "Language Biases in the Coverage of the Science Citation Index and Its Consequences for International Comparisons of National Research Performance." *Scientometrics* 51(1): 335–346.
- Vanhala, Mika, Chien Lu, Jaakko Peltonen, Sanna Sundqvist, Jyrki Nummenmaa y Kalervo Järvelin. 2020. "The Usage of Large Data Sets in Online Consumer Behaviour: A Bibliometric and Computational Text-Mining–Driven Analysis of Previous Research." *Journal of Business Research* 106: 46–59.
- Varian, Hal R. 2014. *Microeconomía Intermedia: Un Enfoque Actual*. Barcelona: Antoni Bosch.
- Wahdiniwaty, Rahma, Neng Sugiana, M. Syaifei, Yeffry Handoko Putra y Lili Wibowo. 2023. "Big Data Algorithms as Enablers of Sustainable Marketing Practices: A Preliminary Research." En *Proceedings of INCITEST 2023*, 1–8. <https://doi.org/10.1109/INCITEST59455.2023.10396871>.
- Wang, Y. y X. Zhang. 2022. "Big Data Mining Method of Marketing Management Based on Deep Trust Network Model." *Journal of Organizational and End User Computing* 34(6): 1–13.
- Wei, L. y J. Chen. 2023. "Internet of Things and Big Data Based Analytics on Consumer Psychology and Behavior." *Journal of Organizational and End User Computing* 35(2): 1–14.

- Wirtz, Jochen, Paul G. Patterson, Werner H. Kunz, Thorsten Gruber, Vinh Nhat Lu, Stefanie Paluch y Antje Martins. 2018. "Brave New World: Service Robots in the Frontline." *Journal of Service Management* 29(5): 907–931. <https://doi.org/10.1108/JOSM-04-2018-0119>.
- Witten, Ian H., Eibe Frank y Mark A. Hall. 2016. *Data Mining: Practical Machine Learning Tools and Techniques*. 4ª ed. Cambridge, MA: Morgan Kaufmann.
- Xiang, Zheng, Muzaffer Uysal, Zvi Schwartz y John Gerdes. 2015. "What Can Big Data and Text Analytics Tell Us About Hotel Guest Experience and Satisfaction?" *International Journal of Hospitality Management* 44: 120–130.
- Yu, Shidong y Dongsheng Yang. 2016. "The Role of Big Data Analysis in New Product Development." En *Proceedings of the International Conference on Network and Information Systems for Computers*, Yingkou, China.
- Zhang, M. y Y. Li. 2021. "Data Mining Algorithm for Demand Forecast Analysis on Flash Sales Platform." *Journal of Organizational and End User Computing* 33(6): 1–13.
- Zhao, X. y J. Li. 2021. "Individual Driver Crash Risk Classification Based on IoV Data and Offline Consumer Behavior Data." *IEEE Transactions on Intelligent Transportation Systems* 22(7): 4373–4382.
- Zuluaga, Martha, Sebastian Robledo, Oscar Arbelaez, German Osorio y Nestor Duque. 2022. "Tree of Science - ToS: A Web-Based Tool for Scientific Literature Recommendation. Search Less, Research More!" *Issues in Science and Technology Librarianship* 100.
- Zupic, Ivan y Tomaz Čater. 2015. "Bibliometric Methods in Management and Organization." *Organizational Research Methods* 18(3): 429–472.

Anexos

Anexo 4. 1. Citas promedio por año

Year	MeanTCperArt	N	MeanTCperYear	CitableYears
1995	16.00	1	0.53	30
1996	140.00	1	4.83	29
1997	339.50	2	12.12	28
1998	8.75	4	0.32	27
1999	14.50	6	0.56	26
2000	3.00	2	0.12	25
2001	79.38	8	3.31	24
2002	17.11	9	0.74	23
2003	18.54	13	0.84	22
2004	22.71	17	1.08	21
2005	29.96	25	1.50	20
2006	27.50	12	1.45	19
2007	8.83	23	0.49	18
2008	43.31	16	2.55	17
2009	16.75	40	1.05	16
2010	9.89	27	0.66	15
2011	13.48	46	0.96	14
2012	22.31	48	1.72	13
2013	40.22	55	3.35	12
2014	15.57	65	1.42	11
2015	24.11	75	2.41	10
2016	17.49	109	1.94	9
2017	20.76	114	2.60	8
2018	25.59	147	3.66	7
2019	15.37	172	2.56	6
2020	19.24	176	3.85	5
2021	12.37	191	3.09	4
2022	7.55	223	2.52	3
2023	2.51	234	1.25	2

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

Anexo 4. 2. Distribución y Colaboración Global en Publicaciones

Country	Articles	Articles %	SCP	MCP	MCP %
CHINA	336	18.1	290	46	13.7
INDIA	112	6	96	16	14.3
USA	87	4.7	74	13	14.9
UNITED KINGDOM	53	2.8	26	27	50.9
KOREA	41	2.2	32	9	22
IRAN	38	2	32	6	15.8

AUSTRALIA	31	1.7	22	9	29
JAPAN	26	1.4	23	3	11.5
CANADA	25	1.3	20	5	20
INDONESIA	24	1.3	23	1	4.2
MALAYSIA	21	1.1	14	7	33.3
GERMANY	20	1.1	15	5	25
FRANCE	19	1	14	5	26.3
GREECE	18	1	13	5	27.8
ITALY	18	1	13	5	27.8
SPAIN	18	1	16	2	11.1

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

Anexo 4. 3. Producción científica por país.

Country	Freq
CHINA	1299
INDIA	658
USA	580
UK	209
AUSTRALIA	188
IRAN	148
SOUTH KOREA	143
INDONESIA	119
JAPAN	116
MALAYSIA	114
GERMANY	112
CANADA	109
ITALY	102
FRANCE	101
SPAIN	89

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

Anexo 4. 4. Áreas temáticas

Área	Porcentaje
Computer Science	34.1%
Engineering	15.6%
Business, Management	10.0%
Mathematics	8.8%
Decision Sciences	8.2%
Other	8.2%
Social Sciences	4.8%
Economics, Econometrics	3.5%
Energy	2.9%
Physics and Astronomy	2.2%
Medicine	1.7%

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

Anexo 4. 5. Nube de palabras

Terms	Frequency
data mining	413
big data	272
machine learning	128
consumer behavior	106
e-commerce	86
customer behavior	78
clustering	69
data analytics	60
big data analytics	51
sentiment analysis	51
customer relationship management	49
consumer behaviour	47
artificial intelligence	46
social media	41
text mining	40
classification	39
customer segmentation	38
decision tree	36
customer behaviour	35
association rules	34
business intelligence	33
marketing	30
crm	25
deep learning	25
market basket analysis	24

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

Anexo 4. 6. Red de Co-citación entre autores

Node	Cluster	Betweenness	Closeness	PageRank
big data	1	180.05	0.016	0.085
machine learning	1	90.667	0.016	0.071
consumer behavior	1	26.8	0.013	0.041
e-commerce	1	16.631	0.013	0.036
data analytics	1	8.264	0.013	0.026
big data analytics	1	8.636	0.012	0.018
sentiment analysis	1	10.061	0.011	0.023
consumer behaviour	1	5.539	0.012	0.02
artificial intelligence	1	1.565	0.012	0.023
social media	1	0.772	0.011	0.019
text mining	1	0.816	0.011	0.017
business intelligence	1	20.003	0.012	0.018
marketing	1	0.08	0.011	0.012

deep learning	1	0.075	0.01	0.01
internet of things	1	0	0.01	0.007
digital marketing	1	0.123	0.011	0.012
opinion mining	1	0.251	0.01	0.011
process mining	1	0.008	0.008	0.004
customer satisfaction	1	3.218	0.011	0.011
data analysis	1	0.727	0.011	0.011
predictive analytics	1	0.008	0.011	0.008
hadoop	1	0.013	0.01	0.007
online shopping	1	0.171	0.011	0.008
data mining	2	547.484	0.018	0.161
customer behavior	2	42.964	0.013	0.036

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

Anexo 4. 7. Red de Colaboración entre autores

Cluster	Autores
1	li y, li l, chen z, liu j, wang h, yang j, zhang h, zhang q
2	chen x, gao j, li z
3	tuzhilin a, jiang t
4	li s, chen y, li q
5	liu q, zhang x
6	hong t-p, chen c-h, nguyen ltt
7	wang y, zhang y, li x, wang x, wang z, li h, wang j, wang l, zhang m, ma j, zhang c, chen q, zhang w
8	fournier-viger p, li t
9	nizar ah, dong zy
10	liu y, li j, chen j, zhang j, wang c, zhang d

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

Anexo 4. 8. Red de Colaboración entre Afiliaciones

Cluster	Universidades
1	university of queensland, nanyang technological university
2	tsinghua university, tianjin university, chongqing university
3	islamic azad university, alzahra university
4	bina nusantara university, telkom university
5	university of technology sydney, northwestern polytechnical university

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).

Anexo 4. 9. Red de Colaboración entre países

Cluster	Países
1	china, usa, united kingdom, australia, iran, korea, germany, canada, italy, france, spain, turkey, greece, hong kong, netherlands, singapore, united arab emirates, israel, switzerland, ireland, belgium, denmark, new zealand, norway, colombia
2	finland, sweden
3	romania, slovakia
4	poland, ukraine
5	india, indonesia, japan, malaysia, saudi arabia, pakistan, iraq
6	portugal, brazil, egypt, peru, ecuador
7	czech republic, hungary

Fuente: Bibliometrix de R Studio, base de datos Scopus (2024).