

# Temas de investigación sobre Big Data en el sector salud

Bochra Ahidar-Tarhouchi

*Universidad Rey Juan Carlos*

*b.abidar.2018@alumnos.urjc.es*

Marta Ortiz-de-Urbina-Criado

*Universidad Rey Juan Carlos*

*marta.ortizdeurbina@urjc.es*

<https://orcid.org/0000-0001-7527-6798>

Recibido: 25-07-2023; Aceptado: 20-12-2023; Publicado: 29-12-2023

## Resumen

**Objetivo:** El sector sanitario tiene en la actualidad un crecimiento rápido y sostenido que genera gran volumen de datos complejos y sensibles. El uso de Big Data ayuda a afrontar retos como la reducción de costes, así como la mejora del uso, la eficacia y la calidad de la información. Sin embargo, todavía hay pocas investigaciones que muestren cuál es la estructura de conocimiento de la investigación sobre la aplicación del Big Data en el sector salud. Por ello, el objetivo de este trabajo es, en primer lugar, identificar las principales temáticas estudiadas y su evolución en el tiempo. Y, en segundo lugar, realizar un análisis reflexivo para proponer temas futuros de investigación.

**Metodología:** Se ha utilizado una metodología mixta combinando un análisis bibliométrico de co-palabras con un análisis reflexivo.

**Resultados:** Se observa que los temas estudiados tanto en el periodo completo como en los dos subperiodos analizados (2012-2019 y 2020-2023) se pueden clasificar en cuatro bloques: contexto, herramientas, medidas y efectos. A partir de ellos se proponen temas futuros de investigación.

**Limitaciones:** Se ha utilizado solo una base de datos y una de las técnicas de análisis bibliométrico.

---

Email de contacto:  
[marta.ortizdeurbina@urjc.es](mailto:marta.ortizdeurbina@urjc.es)

ISSN 0212-1867 / e-ISSN 1989-3574

CC-BY

ESIC University

<https://revistasinvestigacion.esic.edu/esicmarket>

## Cómo citar:

Ahidar-Tarhouchi, B. & Ortiz-de-Urbina-Criado, M. (2023). Temas de investigación sobre Big Data en el sector salud. *ESIC Market. Economics and Business Journal*, 54(2), e316. DOI: 10.7200/esicm.54.316

**Implicaciones prácticas:** Se destaca la necesidad de realizar una aplicación adecuada de las herramientas de Big Data en temas de salud que permita avanzar en asuntos tan relevantes como la prevención, el pronóstico y la investigación sobre enfermedades y sus tratamientos.

**Palabras clave:** Big data; salud; análisis bibliométrico; inteligencia artificial; innovación.

**Códigos JEL:** M15; I10

# Research topics on Big Data in the Health sector

Bochra Ahidar-Tarhouchi

*Universidad Rey Juan Carlos*

*b.abidar.2018@alumnos.urjc.es*

Marta Ortiz-de-Urbina-Criado

*Universidad Rey Juan Carlos*

*marta.ortizdeurbina@urjc.es*

<https://orcid.org/0000-0001-7527-6798>

Received: 25-07-2023; Accepted: 20-12-2023; Published: 29-12-2023

## Abstract

**Objective:** The healthcare sector is currently experiencing rapid and sustained growth, generating huge amounts of complex and sensitive data. The use of big data can help to address challenges such as reducing costs and improving the use, efficiency and quality of information. However, there is still a lack of research on the knowledge structure of research on the application of big data in the health sector. Therefore, the aim of this paper is, firstly, to identify the main topics studied and their evolution in recent years. Secondly, to carry out a reflective analysis to propose future research topics.

**Methodology:** A mixed methodology was used, combining a bibliometric analysis of keywords with a reflexive analysis.

**Results:** It is found that the themes studied, both in the whole period and in the two sub-periods analysed (2012-2019 and 2020-2023), can be grouped into four blocks: context, tools, actions and impact. Based on this, future research topics are proposed.

**Limitations:** Only one database and one of the bibliometric analysis techniques were used.

---

Contact email:  
[marta.ortizdeurbina@urjc.es](mailto:marta.ortizdeurbina@urjc.es)

ISSN 0212-1867 / e-ISSN 1989-3574

CC-BY

ESIC University

<https://revistasinvestigacion.esic.edu/esicmarket>

## How to cite:

Ahidar-Tarhouchi, B. & Ortiz-de-Urbina-Criado, M. (2023). Temas de investigación sobre Big Data en el sector salud. *ESIC Market. Economics and Business Journal*, 54(2), e316. DOI: 10.7200/esicm.54.316

**Practical implications:** It highlights the need for the appropriate application of big data tools in health issues, enabling progress in relevant areas such as prevention, prognosis and research into diseases and their treatments.

**Keywords:** Big data; healthcare; bibliometric analysis; artificial intelligence; innovation.

**JEL Codes:** M15; I10

# 有关卫生领域大数据的研究课题

Bochra Ahidar-Tarhouchi

*Universidad Rey Juan Carlos*

*b.abidar.2018@alumnos.urjc.es*

Marta Ortiz-de-Urbina-Criado

*Universidad Rey Juan Carlos*

*marta.ortizdeurbina@urjc.es*

<https://orcid.org/0000-0001-7527-6798>

Received: 25-07-2023; Accepted: 20-12-2023; Published: 29-12-2023

## 文章摘要

**研究目标:** 医疗行业目前正经历着快速而持续的增长, 产生了大量复杂而敏感的数据。大数据的使用有助于解决医疗行业的各种挑战, 如降低成本以及提高信息的使用率、效率和质量。然而, 研究在卫生领域应用大数据的知识结构仍然很少。因此, 本文的目的首先是确定所研究的主要课题及其随时间的演变。其次, 进行反思分析, 以便提出未来的研究课题。

**分析方法:** 本研究采用了一种混合方法, 将共词分析得文献计量学与反思性分析相结合。

**调查结果:** 本研究的调查结果表明在整个时期和所分析的两个时期(2012-2019 年和 2020-2023 年)研究的主题可分为四个部分: 背景、工具、措施和效果。在此基础上, 提出了未来的研究课题。

**研究局限:** 本研究只使用了一个数据库和一种文献计量分析学。

**实际应用:** 本研究强调了在健康问题上适当应用大数据工具的必要性, 这将有助于在预防、预后和疾病及其治疗研究等相关问题上取得进展。

**关键词:** 大数据; 健康; 文献计量分析; 人工智能; 创新

JEL 代码: M15; I10

---

Contact email:  
[marta.ortizdeurbina@urjc.es](mailto:marta.ortizdeurbina@urjc.es)

ISSN 0212-1867 / e-ISSN 1989-3574

CC-BY

ESIC University

<https://revistasinvestigacion.esic.edu/esicmarket>

## How to cite:

Ahidar-Tarhouchi, B. & Ortiz-de-Urbina-Criado, M. (2023). Temas de investigación sobre Big Data en el sector salud. *ESIC Market. Economics and Business Journal*, 54(2), e316. DOI: 10.7200/esicm.54.316

## 1. Introducción

El sector sanitario es una de las industrias que más rápido crece a nivel mundial y que genera un gran volumen de datos complejos y sensibles. La caracterización de Big Data (BD) se basa en tres aspectos: volumen, velocidad y variedad (Raja *et al.*, 2020). Hoy en día, la utilización de Big Data Analytics (BDA) es muy demandada en el sector sanitario, ya que favorece la mejora de la calidad asistencial a la vez que contribuye a la reducción de los costes sanitarios (De-Pablos-Heredero *et al.*, 2019). El BDA puede utilizarse para la supervisión de la atención sanitaria, predicción de riesgos, detección y prevención del fraude, sistemas de apoyo a la toma de decisiones clínicas, sistemas de recomendación de atención sanitaria personalizada, descubrimiento de fármacos y ensayos clínicos (Raja *et al.*, 2020). Por ejemplo, *Prognos Health* es una base de datos compuesta por datos de diagnóstico clínico proveniente de diversas fuentes. Mediante el uso de los datos de los pacientes, Prognos ha desarrollado una plataforma de inteligencia artificial que tiene la capacidad de predecir la aparición de enfermedades, identificar necesidades terapéuticas, proporcionar oportunidades para ensayos clínicos y señalar posibles riesgos relacionados con más de 30 enfermedades (Prognos Health, 2022).

El uso de BD en el sector sanitario ofrece otras ventajas como la predicción de epidemias, mejorar la calidad de vida o mejorar la toma de decisiones de los profesionales. Además, puede ayudar a solucionar las limitaciones técnicas existentes. Sin embargo, su adopción en las empresas supone un reto para resolver problemas de gestión, almacenamiento, procesamiento y seguridad de la información (Singh *et al.*, 2023). En la actualidad, hay muchas empresas del sector salud que han adoptado el uso del BD. Esta práctica no solo impulsa en la creación de nuevos servicios sino también contribuye a generar empleo. Se estima que en el ámbito del BD en el sector de la salud experimentará un crecimiento del 15% en la generación de puestos de trabajo entre 2014 y 2024 (Mack, 2022). Algunos ejemplos de empresas que lo están usando son: Flatiron Health lo utiliza para brindar apoyo en la lucha contra el cáncer a hospitales, investigadores y médicos en todo el país (Flatiron, 2023); Embleema, especializada en red de blockchains, utiliza el BD para contribuir al descubrimiento de fármacos (Embleema, 2023); Hu-manity.co, empresa que ofrece crear un propio perfil médico de investigadores que estén buscando personas específicas para participar en estudios o ensayos médicos (Hu-manity.co, 2018). Por otra parte, encontramos servicios como *K Health*, que es una aplicación sanitaria gratuita, que ofrece a los consumidores una experiencia de telesalud basada en la inteligencia artificial. Los usuarios pueden describir sus síntomas y a partir de esta información, el algoritmo de la aplicación genera un diagnóstico probable y sugiere posibles tratamientos, así como otros síntomas a los que se debe prestar atención (K Health, 2023).

A medida que las empresas continúan aprovechando las ventajas del BD, se espera que el sector salud siga evolucionando y creciendo en términos de innovación y generación de empleo. Por ello, este tema es muy interesante también en el ámbito

académico. De la Torre Díez *et al.* (2016) concluyen que el crecimiento exponencial de la investigación sobre BD en sanidad tiene un impacto positivo en el sector. El BD ayuda a generar cantidades masivas de información que los métodos tradicionales no pueden analizar, por lo que se necesitan métodos alternativos. El sector salud se enfrenta a retos como el aumento de costes, la ineficacia y la mala calidad (Díaz Muñoz, 2005). El uso eficiente de los datos y la aplicación de análisis cuantitativos y cualitativos puede ayudar a afrontar estos retos (Islam *et al.*, 2018). Hammad *et al.* (2020) analizan el concepto de BD sanitario, que consiste en la recopilación, integración, gestión, procesamiento y análisis de datos médicos diversos. Se propone la web semántica como una solución a estos retos, gracias a su capacidad para buscar, reutilizar, integrar y compartir información.

De la Torre Díez *et al.* (2016) realizan una de las primeras revisiones de la literatura a partir de la selección de 46 artículos publicados en cuatro bases de datos diferentes, entre 2005 y enero de 2016. Sus resultados muestran la importancia del BD en la sanidad, en particular en la medicina personalizada y el control de epidemias. La bibliografía revisada concluye que se requieren inversiones en bioinformática, bio-matemáticas y bioestadística para el desarrollo del análisis y la integración de BD. Alonso *et al.* (2017) presentan una revisión de la literatura, desde 2006 hasta 2018, de las fuentes y técnicas de BD en el sector salud e identifican cuáles de estas técnicas son las más utilizadas en la predicción de enfermedades crónicas. Por su parte, Islam *et al.* (2018) hacen una revisión sistemática de la bibliografía relacionada con las perspectivas teóricas y aplicadas de la minería de datos en la sanidad. Una de las líneas de investigación que identifican es la que analiza la importancia de los informes estructurados y del uso de un lenguaje estándar para facilitar el procesamiento del lenguaje natural, así como de la necesidad de una base de datos de conocimientos para las tecnologías de apoyo a la toma de decisiones automatizadas y basadas en datos. También destacan la importancia de garantizar la calidad de la información a lo largo de todo el ciclo de vida de los datos sanitarios y la necesidad de desarrollar un marco sistemático para la gestión de la calidad de los datos.

Unos años después, Kasten (2020) analiza cómo las técnicas del BD se han extendido del sector financiero a otras áreas como el transporte o la sanidad y realizan una revisión de la literatura previa identificando sus características. El autor explica que se puede seguir investigando para explorar cómo las herramientas centradas en los datos pueden afectar al personal y a los procedimientos de la organización. En ese mismo año, se publica el trabajo de Raja *et al.* (2020) en el que se abordan seis preguntas clave de investigación, y se realiza una revisión de la literatura publicada entre 2015 y 2019 para analizar diferentes perspectivas, examinar fuentes de información, identificar herramientas y técnicas, resaltar beneficios y aplicaciones potenciales y abordar desafíos. Los principales retos que encuentran del BD sanitario se clasifican en gobernanza de datos, retos económicos, retos tecnológicos y problemas de seguridad y confidencialidad. Strang y Sun (2020) llevan a cabo una revisión bibliográfica sobre BD y su relación con la privacidad y la seguridad en la asistencia sanitaria. Relacionado con el anterior, Favaretto *et al.* (2020) revisan, en la literatura

previa, los problemas éticos que se derivan de la creciente digitalización en odontología. Recientemente, Singh *et al.* (2023) proponen varios métodos para ayudar a las organizaciones a aprovechar el poder del BDA en la atención sanitaria como la implementación de la gobernanza de datos, la formación de los empleados, el uso de técnicas de computación en la nube o la aplicación de restricciones de privacidad.

Por otra parte, algunos de los trabajos de revisión han utilizado técnicas bibliométricas. Por ejemplo, Galetsi y Katsaliaki (2020) hacen un análisis bibliométrico y de co-citación de 804 artículos publicados sobre BDA en salud, entre 2000 y 2016, utilizando el software de análisis de texto NVivo. Lo que buscan con estos análisis es identificar a las partes interesadas de BDA en salud y los sistemas innovadores de apoyo a la toma de decisiones. Liao *et al.* (2018) presentan el estado actual de los macrodatos médicos (MBD, por sus siglas en inglés) para explorar el uso de Medical Big Data mediante un análisis de visualización de los artículos utilizando varios programas: GraphPad Prism 5, VOSviewer y CiteSpace. Con ello analizan diferentes aspectos de los trabajos como las citas y el índice H por países, la distribución de las palabras clave, los trabajos más citados, las coautorías, y las revistas y autores más influyentes.

Si bien, hay algunos trabajos que utilizan técnicas bibliométricas para hacer revisiones de la literatura, se observa que hay aspectos que no han sido desarrollados o considerados en estos trabajos. Los estudios previos no incluyen la literatura más actual, quedándose la mayoría de ellos en años anteriores a la pandemia. Los términos que se han utilizado para hacer las búsquedas son variados y diferentes como BDA en salud, Medical Big data, macrodatos médicos. Por ello, las revisiones realizadas no son comparables porque muchos de los trabajos previos no recogen todos los aspectos de las aplicaciones del BD en el sector salud. Por otro lado, salvo el trabajo de Liao *et al.* (2018), que analiza las palabras claves de los artículos, aunque de forma descriptiva, los demás trabajos no hacen análisis de relaciones entre palabras ni identifican grupos de temas y redes temáticas.

Este artículo se centra en el estudio de BD en el sector salud. El primer objetivo de este trabajo es analizar la estructura de conocimiento de este tema y ver su evolución. Para ello, se plantea la primera pregunta de investigación: ¿Cuáles son los principales temas de investigación sobre BD en el sector salud? El segundo objetivo es ofrecer un mapa de los temas estudiados e identificar temas futuros de investigación. Se propone la segunda pregunta de investigación: ¿Qué temas pueden ser interesantes para desarrollar nuevas investigaciones sobre BD en el sector salud?

Para responder a la primera pregunta se utiliza una técnica bibliométrica, el análisis de co-palabras, que se centra en las estructuras conceptuales y permite descubrir los principales temas estudiados en un campo y las interacciones entre ellos (Cobo *et al.*, 2011). Y, para responder a la segunda pregunta, se realiza un análisis reflexivo a partir de los resultados obtenidos, para clasificar los principales temas estudiados e identificar, a partir de ellos, temas interesantes para futuras investigaciones.

Este artículo constituye una contribución original porque combina dos temas actuales y muy relevantes -BD y salud- y aplica una metodología mixta -técnicas bibliométricas y cualitativas-. Sus principales aportaciones son: a) presentar un

panorama de los principales temas analizados sobre BD en sector salud; b) proponer una agrupación lógica de estos temas en cuatro bloques: contexto, herramientas, medición y efectos; y c) presentar un mapa de la investigación presente y futura. De esta forma, ofrece una imagen las principales temáticas estudiadas y da un orden lógico a los conocimientos adquiridos sobre este tema hasta la fecha. También muestra la importancia de la multidisciplinariedad y de sus posibles desarrollos, en especial, desde una perspectiva empresarial. En definitiva, este trabajo presenta una nueva reflexión sobre las aplicaciones del BD para ayudar a conseguir avances productivos y valiosos en el campo de la salud.

El trabajo se ha estructurado de la siguiente manera. Tras la introducción, se describe la metodología y se detallan los pasos y procedimientos empleados para la recopilación de los datos y su análisis. Los resultados son presentados en el tercer apartado, en el que se analizan los diagramas estratégicos, la evolución de los temas en el tiempo, así como las redes temáticas de los grupos más relevantes. En el cuarto apartado se realiza un proceso de reflexión que permite proponer un mapa de la investigación presente y futura. Y finalmente, en las conclusiones se exponen las principales contribuciones del trabajo, sus limitaciones y se plantean algunas líneas futuras de investigación.

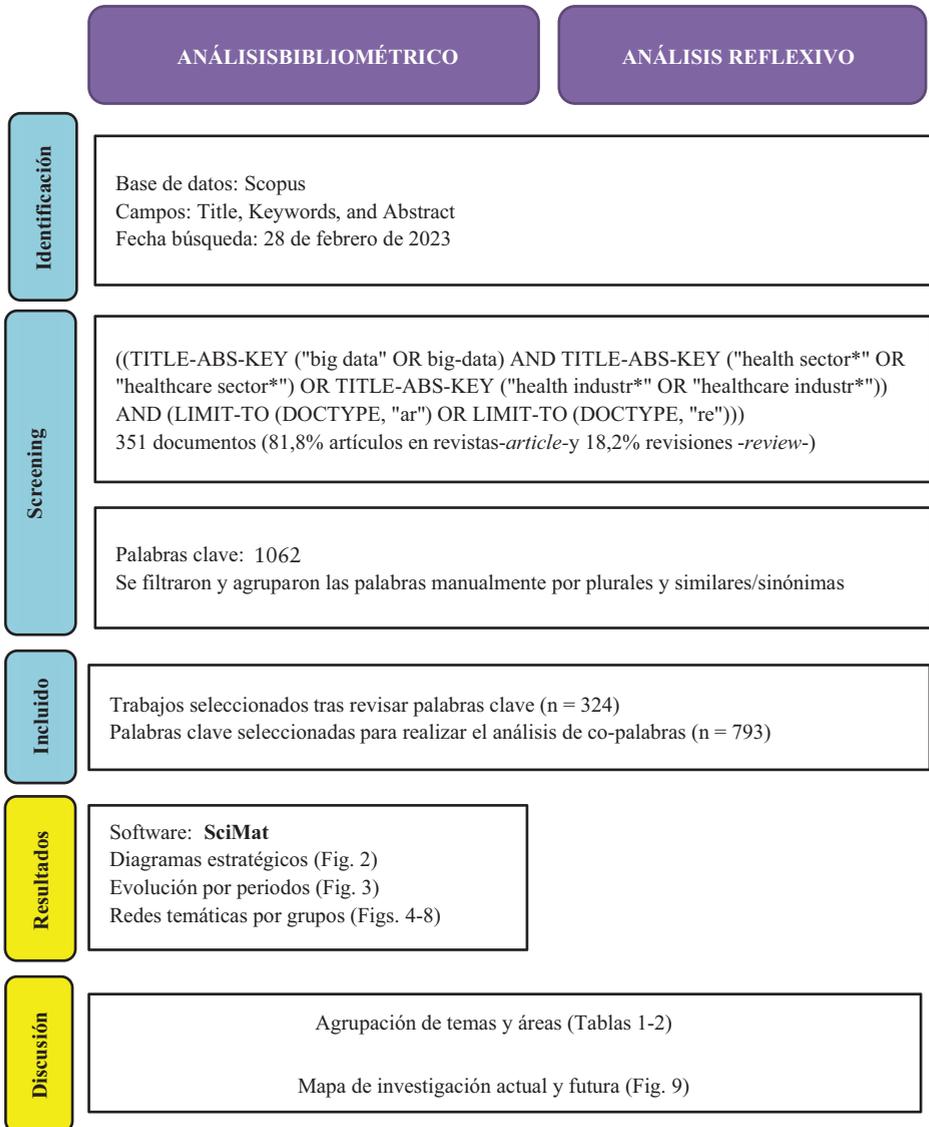
## 2. Metodología

La Figura 1 ilustra el proceso metodológico seguido y las distintas fases del estudio. En primer lugar, se ha realizado una búsqueda para presentar una imagen de los principales temas analizados en la literatura. Para ello, se ha realizado un análisis de co-palabras. Teniendo en cuenta los resultados obtenidos en los diagramas estratégicos y las redes temáticas, tras realizar un análisis reflexivo se presenta un mapa o esquema de los principales temas estudiados y se proponen temas para realizar nuevas investigaciones.

En cuanto a los años de publicación de los artículos, se puede observar que el 2022 fue el año con más publicaciones, 79 documentos (22,51%), seguido por el 2019, con 64 documentos (18,23%) y por el 2020, con 60 documentos (17,09%). Por otro lado, los años en los que hay menos documentos son 2012, 2013 y 2014, con porcentajes de 0,28%, y 1,42% respectivamente. Durante el periodo previo a la pandemia (2012-2019), se publicaron el 41,80% de los documentos analizados y el 58,40% a partir de 2020 y hasta 2023.

Para identificar los distintos temas tratados en la literatura sobre BD en el sector salud se utilizó una técnica bibliométrica, el análisis de co-palabras y el programa SciMat (Cobo *et al.*, 2012). Los principales temas de investigación se identificaron calculando la matriz de co-ocurrencias y el índice de asociación más común, el de equivalencia (Callon *et al.*, 1991; Coulter *et al.*, 1998). A partir de estos índices, se aplica un algoritmo de agrupación, el de los centros simples, para identificar subgrupos de palabras clave que estén fuertemente asociadas. Para identificar temas y

Figura 1. Procedimiento metodológico



Fuente: Elaboración propia inspirada en el modelo PRISMA

redes, se usan las medidas de centralidad y densidad, y se crean los diagramas estratégicos. El tamaño de un círculo representa la importancia del tema según el número de documentos en una red. Para cada tema, se crea su red temática, en la que el tema central es la palabra clave más significativa (mayor nivel de centralidad dentro de una red) y alrededor de ella se establecen sus relaciones con otras palabras. El

tamaño de cada nodo representa una medida cuantitativa como el número de documentos que incluyen la palabra, mientras que el grosor de los arcos suele representar la asociación entre palabras clave (es decir, el índice de equivalencia).

Para realizar el análisis de co-palabras, se ha considerado el período completo (2012-2023) y se ha comparado con dos subperiodos, el prepandemia (2012-2019, 41,60% de los documentos) y los últimos años evaluados (2020-2023, 58,40% de los documentos). Para analizar la evolución de los temas, Cobo *et al.* (2011) proponen utilizar el índice de inclusión de Sternitzke y Bergmann (2009) como medida para determinar el nivel de similitud entre dos redes temáticas con elementos en común a lo largo de periodos de tiempo consecutivos. A partir de este índice, se elabora un gráfico en el que cada tema está representado por un círculo y por líneas entre redes temáticas, cuyo grosor representa el índice de inclusión. Además, una línea continua relaciona las redes en las que el nodo central es el mismo o los casos en que el nombre de una red forma parte de otra (es decir, el nodo central de una red es un nodo no central en la otra), lo que indica que existe una fuerte relación entre ellos. Los demás casos se representan mediante una línea discontinua. El tamaño de un círculo representa el número de documentos que componen cada red.

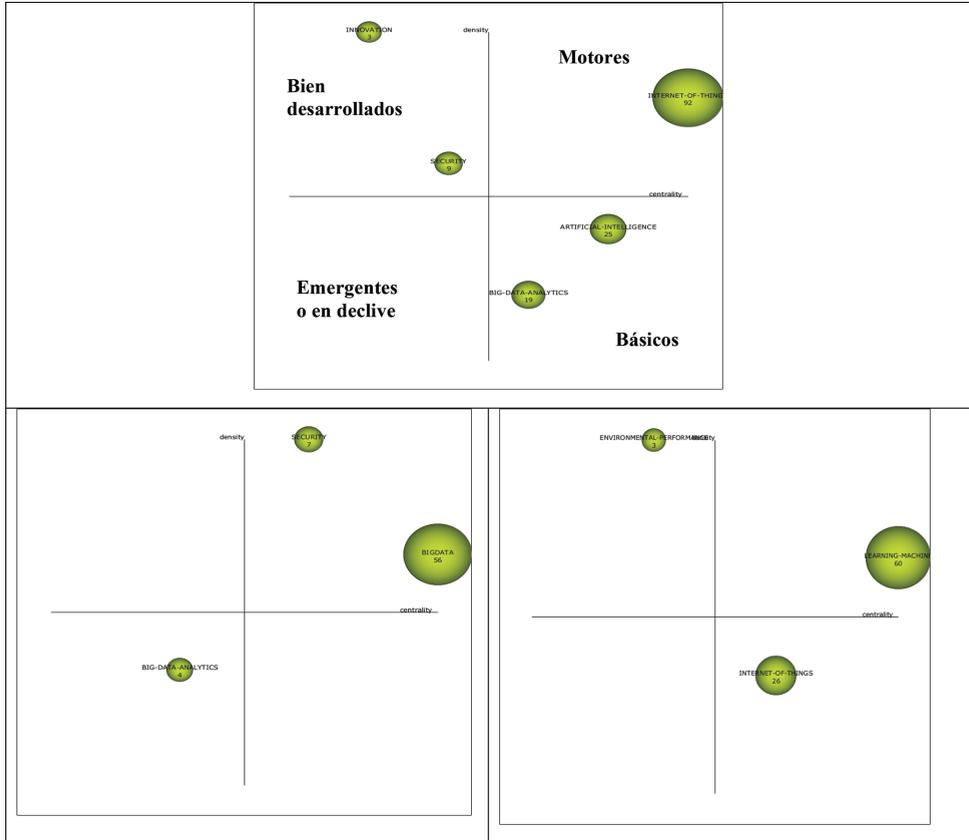
### 3. Resultados

Para identificar la estructura de conocimiento del tema estudiado, se presentan los resultados del análisis de co-palabras. Callon *et al.* (1991) proponen la clasificación de cada red temática en uno de los siguientes grupos: temas bien desarrollados y aislados (cuadrante superior izquierdo), temas emergentes o en desaparición (cuadrante inferior izquierdo), temas básicos y transversales (cuadrante inferior derecho) y temas motores (cuadrante superior derecho), teniendo en cuenta las medidas de centralidad y densidad y creando así un diagrama estratégico.

En la parte superior de la Figura 2 se muestra el diagrama estratégico obtenido a partir de este análisis para todo el periodo (2012-2023). En el cuadrante superior izquierdo se encuentran los términos “Innovation”, mencionado en tres artículos y “Security” (nueve artículos). Estos términos representan los conceptos bien desarrollados, lo que indica que han recibido una atención considerable y han sido objeto de estudio más detallado. En el cuadrante superior derecho, se ubica el término “Internet of Things”, tema motor que aparece en 92 artículos. Por otro lado, en el cuadrante inferior derecho se encuentran dos temas básicos y transversales: “Artificial Intelligence” (19 artículos) y “Big Data Analytics” (25).

En la parte inferior de la Figura 2 se presentan los diagramas para los subperiodos analizados. En el periodo 2012-2019, que corresponde al periodo pre-pandemia (Figura 2, izquierda), se puede observar dos temas motores: “Big Data” (56 artículos) y “Security” (siete artículos). Por otro lado, en lo que respecta a los temas emergentes durante este periodo, destaca el término “Big Data Analytics” (cuatro artículos).

**Figura 2.** Diagrama estratégico por número de documentos para todo el periodo (2012-2023) y por subperiodos (2012-2019 y 2020-2023)



El análisis del periodo actual -2020-2023- (Figura 2, derecha) revela cambios en las temáticas analizadas. En los temas bien desarrollados, destaca el término “Enviromental Performance” (tres artículos) que sugiere un enfoque renovado en la evaluación y mejora del desempeño ambiental. Por otro lado, en los temas básicos encontramos el término “Internet of Things” -IoT- (26 artículos). Y como tema motor tenemos el término “Learning Machine” (60 artículos).

La Tabla 1 resumen los grupos temáticos para todos los periodos analizados.

La Figura 3 muestra la evolución que se ha producido en estos periodos. Considerando todos los años (2012-2023), las temáticas giran en torno a cinco grandes temas. En el caso de IoT es, en los últimos años, cuando se puede considerar un tema muy relevante, aunque durante todo el periodo también está presente en muchos trabajos. La inteligencia artificial es un tema que se analiza también durante todo el periodo de forma regular; mientras que los estudios sobre analíticas de

Tabla 1. Grupos temáticos por periodos

Periodos	Grupos temáticos	Grupos según Callon <i>et al.</i> (1991)
2012-2023	Internet of Things	Motor
	Artificial Intelligence	Básico-transversal
	Big Data Analytics	Básico-transversal
	Security	Bien desarrollado a motor
	Innovation	Bien desarrollado
2012-2019	Big Data	Motor
	Security	Motor
	Big Data Analytics	Emergente
2020-2023	Learning Machine	Motor
	Environmental Performance	Bien desarrollado
	Internet of Things	Básico-transversal

BD se han desarrollado más en el periodo pre-pandemia, al igual que ocurre con el análisis de los aspectos relativos a la seguridad. En cambio, la innovación es un aspecto que se presenta en todo el periodo y a esto se le añade el interés por la aplicación de *machine learning* y por el estudio de los resultados medioambientales en los últimos años.

Para cada grupo temático, SciMat hace un análisis más detallado creando una red de relaciones entre temas. Considerando los tres diagramas estratégicos, se observa que, por una parte, hay términos que definen o delimitan un contexto como es IoT. Por otra parte, aparecen otros temas que reflejan las herramientas que se están empleando como son la inteligencia artificial y *Matching Learning*. Aparecen, en tercer lugar, términos que reflejan el deseo de medir, como *Big Data Analytics*. Y, finalmente, podemos ver algunos de los efectos que puede tener en cuestiones como la innovación, la seguridad y los resultados medioambientales. De esta forma podemos agrupar estos temas en cuatro bloques: contexto, herramientas, medidas y efectos.

Durante el primer periodo (2012-2019) aparece una red muy interesante que resume los temas que se han relacionado con el concepto de BD (Figura 4). Por una parte, encontramos aspectos relativos a las tecnologías –“Learning machine”, “Data Mining”, “Cloud Computing”, Hadoop, “Data Analytics”- (p.ej., Aceto *et al.*, 2018) y su aplicación a temas de salud –“Health-Technology”, “Healthcare”- y Medicina –“Medical-Imaging”, “Genomics”- (p.ej., Mehta *et al.*, 2019). Y, por otra parte, aparecen los usos o utilidades del BD como “Prescribing” y “Predictive-Analytics” (p.ej., Alonso *et al.*, 2017; Kumar y Singh, 2019).

Figura 3. Evolución de temas por periodos (2012-2023, 2012-2019 y 2020-2023)

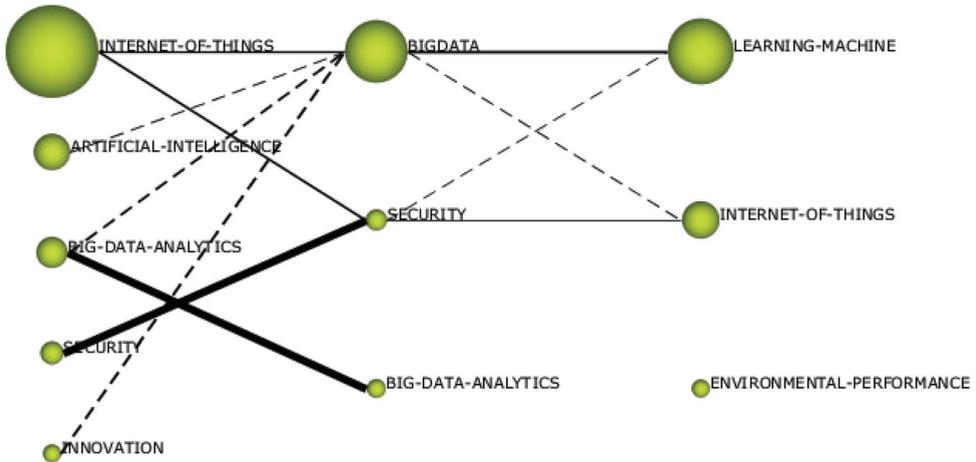
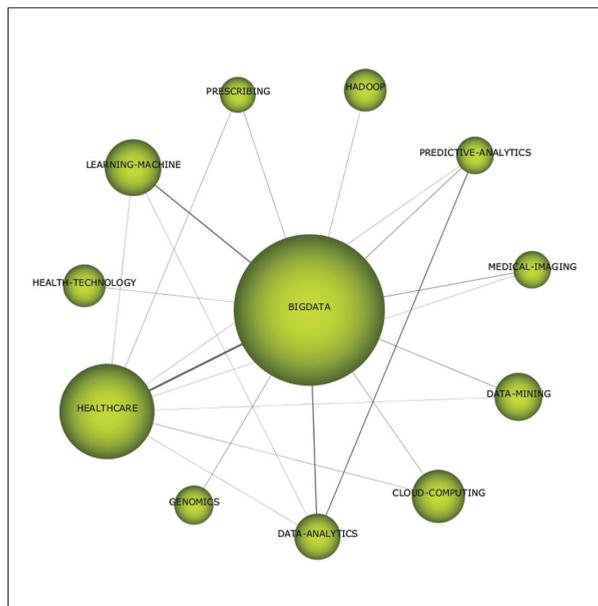


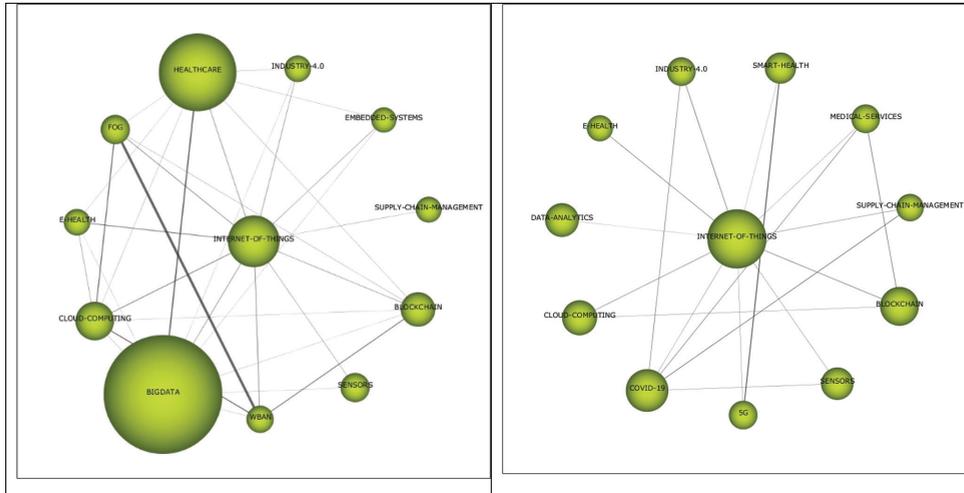
Figura 4. Red temática para Big Data (2012-2019)



### 3.1. Contexto

El tema “Internet of Things” es amplio y se conecta con una gran variedad de términos tanto en todo periodo completo como en los últimos años. Desde 2012 hasta 2023 (Figura 5, izquierda) aparecen once términos que están estrechamente

Figura 5. Red temática para Internet of Things (2012-2023 y 2020-2023)



relacionados entre sí. Se destaca la importancia (tamaño del nodo) y la fuerte conexión (grosor de línea) entre “Big Data” y “Healthcare”. También hay conexión directa y de cierta intensidad entre “FOG” (solución de código abierto para clonar y gestionar ordenadores en red) y “Wireless body area network (WBAN)”. Otros nodos que se relaciona son aspectos como “E-Health”, “Cloud Computing”, “Blockchain”, “Supply Chain Management” e “Industry 4.0”.

En los años 2020 a 2023 (Figura 5, derecha), en la red para “Internet of Things” el nodo más relevante es el término “Covid-19”. Internet of Things ha sido utilizado de varias maneras para hacer frente a la pandemia como para el monitoreo y rastreo de la propagación del virus, la optimización de la cadena de suministro y la habilitación de la telemedicina (Cabrera, 2021). Por otra parte, se observa la relación con otros términos como “Industry 4.0”, “Medical Services” y “Supply Chain Management”. El tema de los servicios médicos (*Medical Services*) pone de manifiesto la importancia que tiene el diagnóstico, tratamiento, atención médica y la prevención en temas de salud y más en situación de pandemia (Awad *et al.*, 2023). El concepto de *Supply Chain Management* hace referencia a aspectos relacionados con la capacidad de gestión de la cadena de suministro en todo el mundo. La crisis sanitaria generó una gran demanda de productos y servicios que hizo necesaria una gestión eficiente y adaptativa de la cadena de suministro (Francis, 2020).

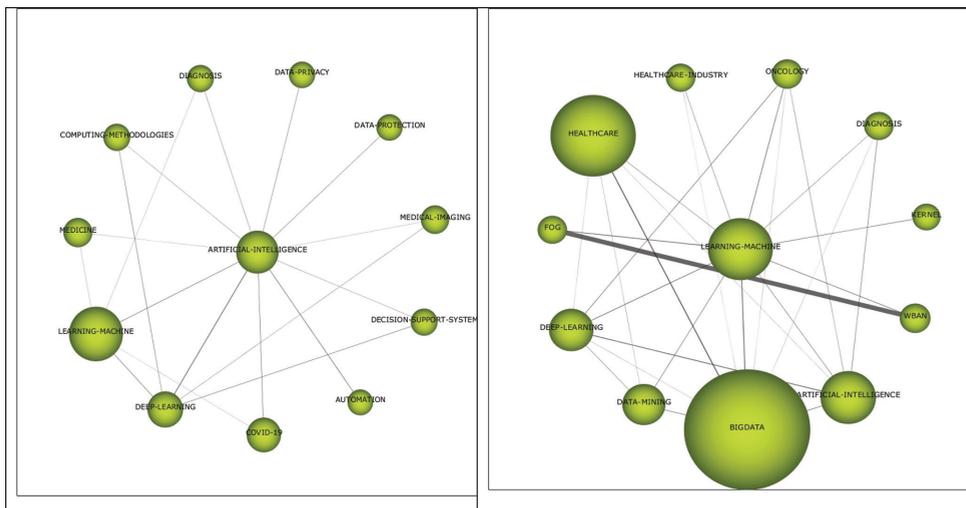
Dentro del grupo hay otros dos nodos destacables: “5G” y “Smart health”. La implementación del 5G en el ámbito de la salud permite una comunicación más rápida y confiable, mejorando así la vigilancia de la salud, la atención medica remota y el desarrollo de soluciones innovadoras en el campo de la medicina (Ahad *et al.*, 2020). También ayuda a que se hagan importantes avances en la adopción de tecnologías en este sector, desarrollándose denominaciones como *Smart health* (Firouzi *et al.*, 2022).

### 3.2. Herramientas

El tema “Artificial Intelligence” (Figura 6, izquierda) mantiene conexiones con once temas. Uno de los nodos destacados es el tema “Learning Machine”, que conecta con el tema principal y también comparte conexiones con los nodos “Medicine” y “Deep Learning”. Esta interconexión nos permite entender que esta rama de la inteligencia artificial facilita el análisis de datos médicos (Herrero-González, 2022), la realización de diagnósticos a partir de esta información y con ello la mejora de la toma de decisiones a nivel clínico (Murari *et al.*, 2022). Otro nodo para destacar es el tema “Computing methodologies” que comparte conexión con el nodo “Learning Machine” y con el tema principal “Artificial Intelligence”. Esta relación destaca el hecho de que la inteligencia artificial se apoya en técnicas y metodologías de computación avanzadas para lograr su funcionamiento y desarrollo (Durga *et al.*, 2019). Por otro lado, vemos la conexión del nodo “Covid-19” con el tema principal del grupo, “Artificial Intelligence”. En este sentido, la inteligencia artificial puede ser una buena herramienta en situaciones de crisis porque proporciona soluciones innovadoras y controles de seguimiento de esta enfermedad a nivel mundial (Vaishya *et al.*, 2020).

Otro de los grandes temas es “Learning Machine” (Figura 6, derecha). La industria de la salud genera una gran cantidad de datos y con ello entra el BD, que permite manejar y analizar esta cantidad de información, mientras que gracias al *learning machine* se utilizan estos datos para identificar patrones, realizar diagnósticos y mejorar los tratamientos médicos (Bohr y Memarzadeh, 2020). Otra relación para destacar es la que se establece entre los términos “Deep Learning”, “Artificial Intelligence” y “Oncology”. En este sentido, los algoritmos y modelos de aprendizaje

Figura 6. Red temática para Artificial Intelligence (2012-2023) y Learning Machine (2020-2023)

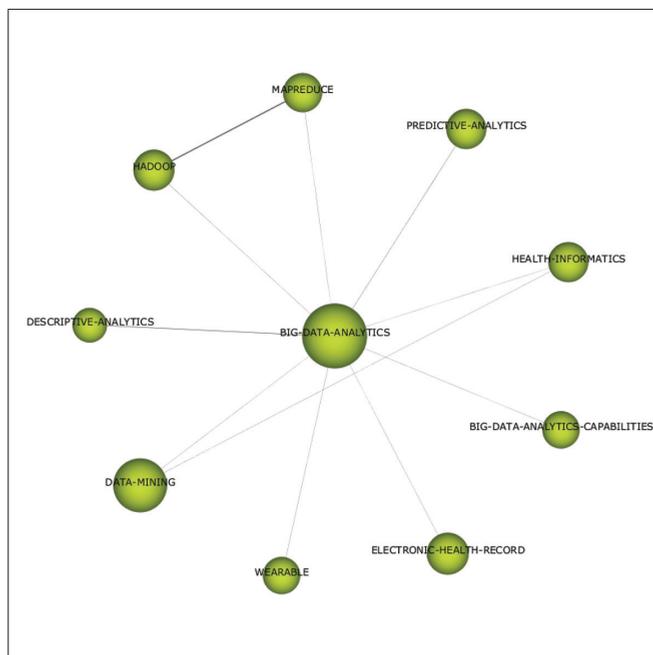


automático y el aprendizaje profundo incluido se aplican para temas de inteligencia artificial y pueden influir en la mejora de los diagnósticos y tratamientos oncológicos (Shimizu y Nakayama, 2020). También es destacable la relación entre “FOG” y “WBAN” (Yıldırım *et al.*, 2023), que radica en su enfoque común de aprovechar las redes inalámbricas para habilitar aplicaciones avanzadas en el campo de la salud.

### 3.3. Medidas

El tema “Big Data Analytics” (Figura 7) comparte relación con otros nueve temas. Entre ellos, destacan la conexión con “Data Mining” y “Health informatics”, ya que el uso de estas herramientas dentro del BDA facilita la transmisión de grandes volúmenes de datos en los sistemas informáticos sanitarios (Herland *et al.*, 2013). Destaca también la relación de “Descriptive Analytics” y “Big Data Analytics” para realizar análisis de datos y generar de información para la toma de decisiones (Nisar *et al.*, 2021). Por su parte, los temas “Hadoop” y “MapReduce” están estrechamente relacionados y se complementan a la hora de procesar grandes volúmenes de datos. “Hadoop” es un framework de código abierto que permite el almacenamiento y distribución de datos, mientras que “MapReduce” es un modelo de programación utilizado en el procesamiento paralelo de datos (Oscarfmdc, 2022).

Figura 7. Red temática para Big Data Analytics (2012-2023)

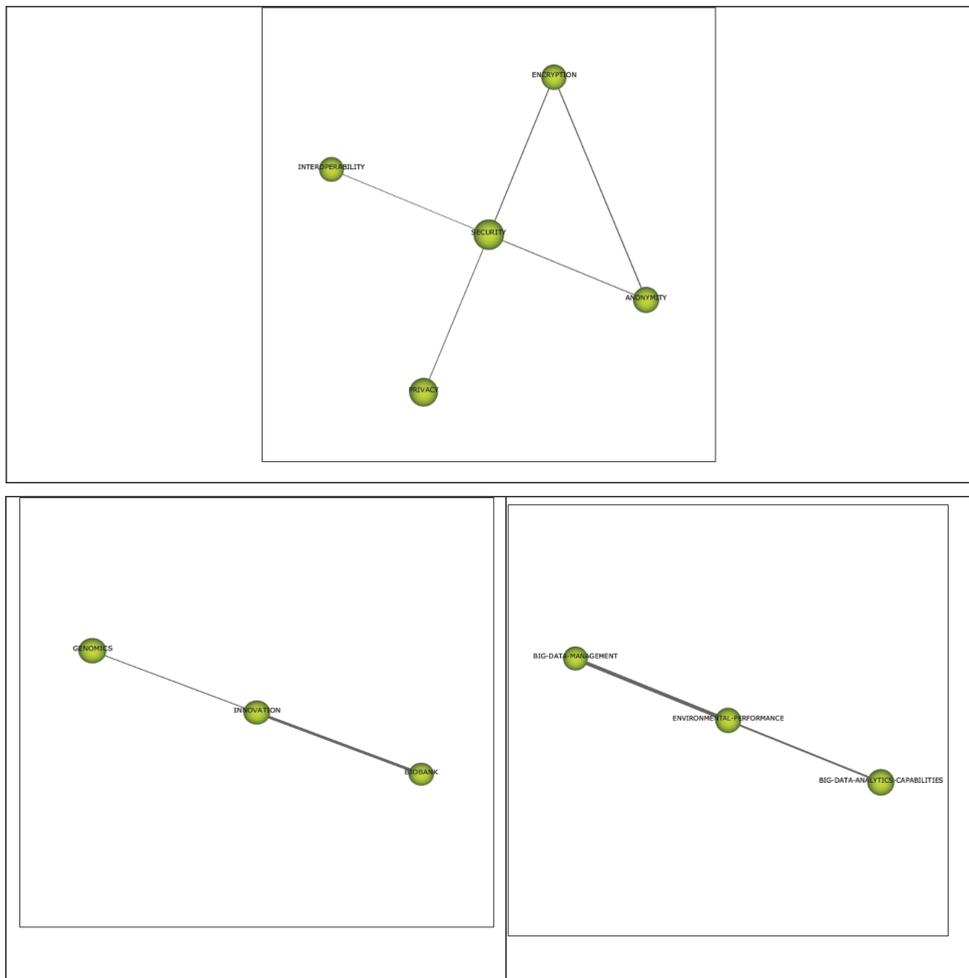


### 3.4. Efectos

El tema “Security” (Figura 8) mantiene conexiones con cuatro términos, destacando la relación de los temas “Encryption” y “Anonymity”, que son elementos que garantizan la protección de la información más relevante en entornos digitales. Los nodos “Privacy” e “Interoperability” se refieren más a la protección de la privacidad de la información mediante medidas de seguridad efectivas que permiten el intercambio de información entre sistemas en un entorno seguro (Abouelmehdi *et al.*, 2018).

El tema “Innovation” (Figura 8) se relaciona estrechamente con “Genomics” y “Biobank”. La genómica se centra en el estudio de los genes y sus funciones dentro

**Figura 8.** Red temática para Security (2012-2023), Innovation (2012-2023) y Environmental Performance (2020-2023)



del organismo lo que proporciona información valiosa sobre diversos aspectos relevantes para la salud y la innovación va de la mano con este concepto, ya que impulsa el desarrollo de nuevas tecnologías y abrir nuevas oportunidades en la mejora de resultados (Lee y DelVecchio, 2017). “Biobank” hace referencia a una entidad dedicada al almacenamiento y conservación de una variedad de muestras biológicas y la innovación desempeña un papel fundamental para impulsar el desarrollo de técnicas, análisis y otros métodos y a la vez amplia posibilidades de avances médicos (Kinkorová y Topolčan, 2020).

Por último, el tema “Environmental Performance” (Figura 8) se relaciona con dos términos muy similares, “Big Data Management” y “Big Data Analytics Capabilities”. Ambos términos son fundamentales para garantizar la calidad y accesibilidad de la información sobre aspectos ambiental. Además, la capacidad de análisis avanzado de esos datos permite comprender mejor el impacto ambiental de diferentes actividades y evaluar la efectividad de las medidas sanitarias para proteger el medio ambiente (Nisar *et al.*, 2021).

#### 4. Discusión y agenda futura

A partir de los resultados obtenidos en el apartado anterior, se ha realizado un proceso de análisis reflexivo para proponer un esquema / mapa de los temas estudiados y para definir algunos temas futuros (Figura 9).

Los temas que han sido estudiados en la literatura presentan una imagen muy cercana a lo que está pasando en la realidad empresarial. Las empresas del sector salud se encuentran con un avance imparable de la tecnología y con todas las utilidades que tiene *Internet of Things*, pero ahora se añade las grandes oportunidades y

Figura 9. Mapa de la investigación presente y futura

	Bien desarrollados	Básicos	Motores	Futuros
Contexto			<i>Internet of Things</i>	Multidisciplinariedad Gobierno corporativo Gestión del conocimiento
Herramientas	Inteligencia artificial		<i>Learning Machine</i>	Tecnologías como realidad virtual, ampliada, inteligencia artificial, etc. Metaverso
Medidas		<i>Big Data Analytics</i>		<i>Big Data Analytics</i> <i>Balance Scorecard</i>
Efectos	Seguridad Innovación Rendimiento medioambiental			Factores determinantes Obstáculos / problemas Coordinación relacional

amenazas que supone el desarrollo de la inteligencia artificial. No solo el uso de los datos y la información disponible alcanza otros niveles, sino que además se genera nueva información, más enriquecida que tiene muchas ventajas en el sector salud, pero también supone superar muchos retos y problemas debidos a la sensibilidad de sus datos. Las herramientas clásicas del Big Data como el *Learning Machine* están siendo desarrolladas y se está ampliando el alcance las Big Data Analytics, que han servido para innovar y mejorar los resultados, pero tienen como principales retos la seguridad y vigilancia de los datos.

Esta imagen / mapa de los temas principales de la investigación ofrece una buena base para plantearse las vías de avance en la investigación a futuro. Algunas preguntas de investigación interesantes que se pueden plantear son: ¿Qué efectos puede tener Chat GPT en las diferentes variables de uso de los BD en Ciencias de la Salud? ¿Cuáles son los riesgos específicos que se plantean en salud pública debido a la conectividad de la información (bases de datos) y a las brechas de ciberseguridad?

La agenda futura en temas de BD en el sector salud tiene que seguir reflejando claramente la realidad para abordar formas diferentes de afrontar los retos y resolver los problemas. La investigación se ha realizado desde diferentes áreas de conocimiento lo que permite reflejar el gran valor de la multidisciplinar. En la Tabla 2 se han resumido las áreas de conocimiento para todo el periodo y para el último año -2023-. En 2023 hay 40 documentos que están adscritos al menos a dos áreas diferentes de conocimiento. Tanto durante todo el periodo analizado como en el último año destacan los trabajos del área de Ciencias de la Computación, Ingeniería, Medicina y Dirección de Empresas.

En el área de la Dirección de Empresas se ha observado que todavía hay poco desarrollo de investigaciones sobre las aplicaciones de BD en el sector de la salud. En concreto, se ha realizado una búsqueda específica en Scopus (16/11/2023) limitando los resultados al campo de *Business, Management and Accounting* y se han obtenido solo 26 documentos, de los cuales siete son del año 2023 y el primero se publicó en el año 2017. Por eso, se propone que puede ser interesante profundizar en el estudio de este tema desde esta disciplina.

Desde el campo de la Dirección de Empresas se puede ofrecer fundamentación y modelos que ayuden a hacer la implementación de prácticas de análisis de grandes cantidades de datos en empresas de una forma más eficiente, ética y con responsabilidad social teniendo en cuenta a todos los grupos de interés. Además, los modelos de gestión del conocimiento pueden servir para entender mejor el manejo de grandes volúmenes de datos e información y las posibilidades de reutilización de estos datos. También, otros modelos del área de empresa pueden ser muy interesantes para el desarrollo de estudios futuros, es el caso de los modelos de coordinación relacional, de los modelos de cambio organizativo y de los de toma de decisiones.

Desde la pandemia hay un mayor interés en la transparencia y apertura de datos sobre salud. Dash *et al.* (2019) realizan interesantes reflexiones sobre las fuentes de BD disponibles en registros hospitalarios, historiales médicos de pacientes, resultados de exámenes médicos y dispositivos que forman parte del Internet de

**Tabla 2.** Áreas de conocimiento de la investigación sobre Big Data en salud

Área de conocimiento (Scopus)	Muestra (2012-2023)	2023(1)	Área de conocimiento (Scopus)	Muestra (2012-2023)	2023(1)
Ciencias de la Computación	176	23	Farmacología, Toxicología y Farmacia	12	0
Ingeniería	131	18	Enfermería	9	2
Medicina	87	6	Multidisciplinar	7	0
Dirección de Empresas ( <i>Business, management</i> y contabilidad)	43	7	Ciencias Agrarias y Biológicas	6	0
Matemáticas	37	6	Energía	6	1
Ciencias de los Materiales	28	5	Química	5	1
Bioquímica, Genética y Biología Molecular	27	1	Inmunología y Microbiología	5	0
Ciencias de la Decisión	25	5	Artes y Humanidades	4	0
Ciencias Sociales	22	2	Economía, Econometría y Finanzas	4	0
Profesiones de la Salud	17	3	Psicología	4	0
Ciencias Medioambientales	16	2	Ciencias de la Tierra y Planetarias	2	0
Física y Astronomía	15	3	Odontología	1	0
Ingeniería Química	14	1	Neurociencia	1	1

(1) Scopus (16/11/2023, 40 documentos)

las cosas. Los autores comentan que estos datos requieren una gestión y un análisis adecuados para obtener información significativa. En este sentido, una opción es la publicación de datos abiertos sobre salud, por parte, tanto de empresas como de administraciones públicas, para ofrecer información que ayude a mejorar la toma de decisiones sanitarias y a crear nuevos productos y/o servicios basados en datos abiertos. Esto puede ayudar a prevenir enfermedades y epidemias, tomar decisiones sobre política sanitaria, gestionar la atención sanitaria y mejorar la investigación en temas de salud.

Por otra parte, parece oportuno, en un momento económico, social y político como el actual, que se amplie el conocimiento sobre los factores determinantes de la aplicación de BD en empresas y organizaciones (p. ej., Janyapoon *et al.*, 2021). Y que, en especial, se profundice en el efecto que pueden tener en este campo la inteligencia artificial, otras tecnologías como la realidad virtual, aumentada o mixta y aplicaciones tan actuales de estas como el metaverso. Otra línea futura de

investigación es la que tiene que ver con el análisis de los obstáculos y los problemas (p. ej., Kant Pal *et al.*, 2021) que puede generar si no se usan de forma adecuada, ética y moral esas grandes cantidades de datos sobre la salud de las personas y sobre aspectos sanitarios con gran impacto social y económico.

## 5. Conclusiones

En este trabajo se ha explorado la literatura sobre BD en el sector de la salud, se han identificado los principales temas analizados, sus interrelaciones y su evolución lo que contribuye a profundizar en su conocimiento. Y además ha permitido realizar propuestas de futuro para continuar avanzando en esta línea de investigación.

La acumulación y tratamiento de grandes cantidades de datos sobre salud tiene un impacto social, económico y empresarial. Por ello, es importante conocer su utilidad e identificar las oportunidades que ofrecen los conjuntos de datos. Además, una mayor apertura y posibilidades de acceso a estos datos puede revertir en una mayor transparencia, participación e innovación en la sociedad (Weerakkody *et al.*, 2017).

Este trabajo tiene varias contribuciones. En el ámbito académico, este trabajo representa un avance en el conocimiento del campo del BD en la salud. El análisis realizado ofrece una visión general de los temas más relevantes y de su evolución temporal. Se plantea la necesidad de disponer de datos sobre salud y de utilizarlos de forma adecuada para que puedan aportar valor social y económico de forma ética. También se presenta el interés de analizar las ventajas e inconvenientes de la apertura de estos datos (datos abiertos sobre salud) para desarrollar productos y servicios innovadores. El análisis reflexivo realizado es un punto de partida para futuras investigaciones y estudios relacionados con el uso del BD en el sector de la salud.

Desde un enfoque más práctico, no podemos olvidar que el sector sanitario tiene datos sensibles con una configuración de protección alta. Hay estructuras de datos inadecuadas que ocasionan complejidad en el tratamiento de los datos y limitaciones en su uso. En este trabajo se plantea el problema de la seguridad de los datos y se destaca la importancia que tiene la aplicación adecuada de las herramientas de BD en temas de salud ya que, entre otras, cosas permiten avanzar en asuntos tan relevantes como la investigación de enfermedades y sus tratamientos.

En el contexto empresarial, el análisis realizado brinda información que puede ayudar a mejorar el uso del BD en empresas del sector de la salud. Los resultados resaltan la importancia de la innovación y del rendimiento medioambiental, lo cual es fundamental para las organizaciones que quieren aprovechar el potencial del BD en sus operaciones y toma de decisiones. Las herramientas relacionadas con el BD ayudan a las empresas sanitarias a mejorar pronósticos, optimizar los tratamientos y generar recomendaciones personalizadas para los pacientes. También se plantea que hay que seguir dando pasos para aplicar modelos de gestión empresarial que ayuden a mejorar la adopción y uso de analíticas de BD en temas de salud.

Este trabajo presenta algunas limitaciones. Para la selección de la muestra se utilizó la base de datos Scopus, por lo que otros estudios futuros pueden ampliarlo añadiendo la información de otras fuentes. Por otro lado, este trabajo realiza una revisión de la literatura utilizando bibliometría, pero sería interesante complementarlo con otras metodologías como el análisis de eventos. Por otra parte, para desarrollar más los estudios desde un enfoque empresarial, se pueden analizar los factores determinantes y sus efectos utilizando técnicas cuantitativas como los modelos de ecuaciones estructurales. Además, trabajos futuros pueden plantear el estudio de temas relacionados como el gobierno de datos abiertos en el sector salud, así como temas de gestión del conocimiento y de conectividad de la información sobre salud. Finalmente, sería interesante ampliar el análisis a otros sectores relacionados con el de la salud para ampliar los efectos que puede tener la aplicación de las herramientas de BD en la prevención, control y abordaje de problemas sanitarios.

## Declaración de conflictos de interés

Los autores declaran que no presentan conflictos de interés en relación con la investigación, autoría o la publicación de este trabajo.

## Financiación

Los autores no han recibido apoyo financiero para la investigación, por autoría y/o la publicación de este trabajo.

## Referencias

- Abouelmehdi, K., Beni-Hessane, A. y Khaloufi, H. (2018). Big healthcare data: preserving security and privacy. *Journal of Big Data*, 5, 1. <https://doi.org/10.1186/s40537-017-0110-7>
- Aceto, G., Persico, V. y Pescapé, A. (2018). The role of information and communication technologies in healthcare: taxonomies, perspectives, and challenges. *Journal of Network and Computer Applications*, 107, 125-154. <https://doi.org/10.1016/j.jnca.2018.02.008>
- Ahad, A., Tahir, M., Sheikh, M. U., Ahmed, K. M., Mughees, A. y Numani, A. (2020). Technologies trend towards 5g network for smart health-care using IoT: a review. *Sensors*, 20(14), 4047. <https://doi.org/10.3390/s20144047>
- Awad, A. I., Fouda, M. M., Khashaba, M. M., Mohamed, E. R. y Hosny, K. M. (2023). Utilization of mobile edge computing on the internet of medical things: A survey. *ICT Express*, 9(3), 473-485. <https://doi.org/10.1016/j.icte.2022.05.006>

- Alonso, S. G., de la Torre Díez, I., Rodrigues, J. J. P. C., Hamrioui, S. y López-Coronado, M. A. (2017). Systematic review of techniques and sources of big data in the healthcare sector. *Journal of Medical Systems*, 41(11), 183. <https://doi.org/10.1007/s10916-017-0832-2>
- Bohr, A. y Memarzadeh, K. (2020). Chapter 1 - Current healthcare, big data, and machine learning. En Adam Bohr, A. y Memarzadeh, K. (Eds.), *Artificial Intelligence in Healthcare* (pp. 1-24). Academic Press. <https://doi.org/10.1016/B978-0-12-818438-7.00001-0>
- Cabrera, M. (17 de febrero de 2021). *El internet de las cosas (IoT) y la salud en la era de la COVID-19*. Conexión ESAN. <https://www.esan.edu.pe/conexion-esan/el-internet-de-las-cosas-iot-y-la-salud-en-la-era-de-la-covid-19-1#:~:text=La%20implementaci%C3%B3n%20del%20IoT%20ha,la%20reducci%C3%B3n%20de%20los%20contagios>
- Callon, M., Courtial, J. y Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics*, 22(1), 155-205. <https://doi.org/10.1007/bf02019280>
- Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E. y Herrera, F. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field: a practical application to the Fuzzy Sets Theory field. *Journal of Informetrics*, 5(1), 146-166. <https://doi.org/10.1016/j.joi.2010.10.002>
- Cobo, M., López-Herrera, A. G., Herrera-Viedma, E. y Herrera, F. (2012). SciMAT: A new science mapping analysis software tool. *Journal of the Association for Information Science and Technology*, 63(8), 1609-1630. <https://doi.org/10.1002/asi.22688>
- Coulter, N., Monarch, I. y Konda, S. (1998). Software engineering as seen through its research literature: a study in co-word analysis. *Journal of the American Society for Information Science*, 49, 1206-23. [https://doi.org/10.1002/\(SICI\)1097-4571\(1998\)49:13<1206::AID-ASI7>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1097-4571(1998)49:13<1206::AID-ASI7>3.0.CO;2-F)
- Dash, S., Shakyawar, S. K., Sharma, M. y Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of Big Data*, 6(1), 1-25. <https://doi.org/10.1186/s40537-019-0217-0>
- De la Torre Díez, I., Cosgaya, H. M., Garcia-Zapirain, B. y López-Coronado, M. (2016). Big data in health: a literature review from the year 2005. *Journal of Medical Systems*, 40, 209. <https://doi.org/10.1007/s10916-016-0565-7>
- De-Pablos-Heredero, C., López-Hermoso-Agius, J. J., Martín-Romo-Romero, S. y Medina-Salgado, S. (2019). *Organización y transformación de los sistemas de información en la empresa* (4ª edición). ESIC Editorial.
- Díaz Muñoz, R. (2005). Dimensiones de la calidad del servicio sanitario. *ESIC Market*, 36(121), 87-114. <https://revistasinvestigacion.esic.edu/esicmarket/index.php/esicm/article/view/44>
- Durga, S., Nag, R. y Daniel, E. (2019). Survey on machine learning and deep

- learning algorithms used in internet of things (IOT) healthcare. 2019 3rd *International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, pp. 1018-1022. <https://doi.org/10.1109/iccmc.2019.8819806>
- Embleema (2023). *About us*. Embleema. Recuperado 4 de julio de 2023, de <https://www.embleema.com/>
- Favaretto, M., Shaw, D. E., De Clercq, E. M., Joda, T. y Elger, B. S. (2020). Big data and digitalization in dentistry: a systematic review of the ethical issues. *International Journal of Environmental Research and Public Health*, 17(7), 2495. <https://doi.org/10.3390/ijerph17072495>
- Firouzi, F., Farahani, B., Barzegari, M. y Daneshmand, M. (2022). AI-driven data monetization: the other face of data in iot-based smart and connected health. *IEEE Internet of Things Journal*, 9(8), 5581-5599. <https://doi.org/10.1109/JIOT.2020.3027971>
- Flatiron (2023). *Flatiron health*. Recuperado 4 de julio de 2023, de <https://flatiron.com/>
- Francis, J. B. (2020). COVID-19: Implications for supply chain management. *Frontiers of health services management*, 37(1), 33-38. <https://doi.org/10.1097/hap.0000000000000092>
- Galetsi, P. y Katsaliaki, K. (2020). Big data analytics in health: An overview and bibliometric study of research activity. *Health Information & Libraries Journal*, 37(1), 5-25. <https://doi.org/10.1111/hir.12286>
- Hammad, R., Barhoush, M. y Abed-Alguni, B. H. (2020). A semantic-based approach for managing healthcare big data: a survey. *Journal of Healthcare Engineering*, 2020, 1-12. <https://doi.org/10.1155/2020/8865808>
- Herland, M., Khoshgoftaar, T. M. y Wald, R. (2013). *Survey of clinical data mining applications on big data in health informatics*. <https://doi.org/10.1109/icmla.2013.163>
- Herrero-González, A. (2022). The value of data and its applicability in the Health Sector. *Revista Española de Medicina Nuclear e Imagen Molecular (English Edition)*, 41(1), 39-42. <https://doi.org/10.1016/j.remnie.2021.11.002>
- Hu-manity.co. (2018). *Hu-manity.co: Restoring trust in digital*. <https://hu-manity.co/>
- Islam, S., Hasan, M. M., Wang, X., Germack, H. D. y Noor-E-Alam, M. A. (2018). A systematic review on healthcare analytics: application and theoretical perspective of data mining. *Healthcare*, 6(2), 54. <https://doi.org/10.3390/healthcare6020054>
- Janyapoon, S., Liangrokapt, J. y Tan, A. (2021). Critical success factors of business intelligence implementation in Thai hospitals. *International Journal of Healthcare Information Systems and Informatics (IJHISI)*, 16(4), 1-21. <http://doi.org/10.4018/IJHISI.20211001.0a19>
- Kant Pal, S., Mukherjee, S., Baral, M. M., y Aggarwal, S. (2021). Problems of big data adoption in the healthcare industries. *Asia Pacific Journal of Health Management*, 16(4), 282-287. <https://doi.org/10.24083/apjhm.v16i4.1359>

- Kasten, J. E. (2020). Big Data applications in healthcare administration. *International Journal of Big Data and Analytics in Healthcare (IJBD AH)*, 5(2), 12-37. <https://doi.org/10.4018/ijbdah.2020070102>
- K Health (2023). *About Us | K Health*. <https://khealth.com/about>
- Kinkorová, J. y Topolčan, O. (2020). Biobanks in the era of big data: objectives, challenges, perspectives, and innovations for predictive, preventive, and personalised medicine. *EPMA Journal*, 11, 333-341 <https://doi.org/10.1007/s13167-020-00213-2>
- Kumar, S. y Singh, M. (2019). Big data analytics for healthcare industry: impact, applications, and tools. *Big Data Mining and Analytics*, 2(1), 48-57. <https://doi.org/10.26599/BDMA.2018.9020031>
- Lee, K. y DelVecchio, A. (2017). *Genomics*. Health IT. <https://www.techtarget.com/searchhealthit/definition/genomics>
- Liao, H., Tang, M., Luo, L., Li, C., Chiclana, F., y Zeng, X. J. (2018). A bibliometric analysis and visualization of medical big data research. *Sustainability*, 10(1), 166. <https://doi.org/10.3390/su10010166>
- Mack, J. (2022). *5 health care Big Data companies to watch*. University of San Diego Online Degrees. <https://onlinedegrees.sandiego.edu/5-health-care-big-data-companies-watch/>
- Mehta, N., Pandit, A. y Shukla, S. (2019). Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study. *Journal of Bio-medical Informatics*, 100, 103311. <https://doi.org/10.1016/j.jbi.2019.103311>
- Murari, T., Prathiba, L., Singamaneni, K. K., Venu, D., Nassa, V. K., Kohar, R. y Uparkar, S. S. (2022). Big data analytics with OENN based clinical decision support system. *Intelligent Automation & Soft Computing*, 31(2), 1241-1256. <https://doi.org/10.32604/iasc.2022.020203>
- Nisar, Q. A., Nasir, N., Jamshed, S., Naz, S., Ali, M. y Ali, S. (2021). Big data management and environmental performance: role of big data decision-making capabilities and decision-making quality. *Journal of Enterprise Information Management*, 34(4), 1061-1096. <https://doi.org/10.1108/JEIM-04-2020-0137>
- Oscarfmdc (2022). ¿Qué es Hadoop MapReduce? Introducción. *Aprender BIG DATA*. Recuperado 26 de junio de 2023, de <https://aprenderbigdata.com/hadoop-mapreduce/>
- Prognos Health (2022). *About us*. Recuperado 4 de julio de 2023, de <https://prognoshealth.com/about-us/>
- Raja, R., Mukherjee, I. y Sarkar, B. K. (2020). A systematic review of healthcare big data. *Scientific Programming*, 2020, 1-15. <https://doi.org/10.1155/2020/5471849>
- Shimizu, H. y Nakayama, K. I. (2020). Artificial intelligence in oncology. *Cancer science*, 111(5), 1452-1460. <https://doi.org/10.1111/cas.14377>
- Singh, R., Agrawal, S., Sahu, A. y Kazancoglu, Y. (2023). Strategic issues of big data analytics applications for managing health-care sector: a systematic literature review and future research agenda. *The TQM Journal*, 35(1), 262-291. <https://doi.org/10.1108/tqm-02-2021-0051>

- Sternitzke, C. y Bergmann, I. (2009). Similarity measures for document mapping: a comparative study on the level of an individual scientist. *Scientometrics*, 78(1), 113-30. <https://doi.org/10.1007/s11192-007-1961-z>
- Strang, K. D. y Sun, Z. (2020). Hidden big data analytics issues in the healthcare industry. *Health Informatics Journal*, 26(2), 981-998. <https://doi.org/10.1177/1460458219854603>
- Vaishya, R., Javaid, M., Khan, I. y Haleem, A. (2020). Artificial intelligence (AI) applications for COVID-19 pandemic. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, 14(4), 337-339. <https://doi.org/10.1016/j.dsx.2020.04.012>
- Weerakkody, V., Irani, Z., Kapoor, K., Sivarajah, U. y Dwivedi, Y. K. (2017). Open data and its usability: an empirical view from the citizen's perspective. *Information Systems Frontiers*, 19(2), 285-300. <https://doi.org/10.1007/s10796-016-9679-1>
- Yıldırım, E., Cicioğlu, M. y Çalhan, A. (2023). Fog-cloud architecture-driven internet of medical things framework for healthcare monitoring. *Medical & Biological Engineering & Computing*, 61, 1133-1147. <https://doi.org/10.1007/s11517-023-02776-4>