

# APLICACIÓN DE *BIG DATA* EN SISTEMAS DE SALUD PÚBLICA

## APPLICATION OF BIG DATA IN PUBLIC HEALTH SYSTEMS



<sup>1</sup>Andrés Felipe López Copete,  
<sup>2</sup>Libardo Rodríguez Martínez,  
<sup>3</sup>Daniel Alexander Ramírez Gómez

<sup>1,3</sup>Universidad de Medellín, Colombia,  
Universidad Católica de Colombia  
<sup>2</sup>Universidad de San Carlos de Guatemala,

Recibido: 20/10/2022 Aprobado 22/12/2022

### RESUMEN

En este artículo se mencionan una serie de estrategias y casos de éxito que mediante su correcta aplicación permiten mejorar la atención de los pacientes en los sistemas de información en salud pública, a su vez con la optimización de recursos se aumentan los niveles de eficiencia en todos los procesos de atención como lo son, manejo de inventario, gestión de personal, prevención y diagnóstico oportuno de infecciones. Todo esto se logra mediante el uso de herramientas *big data*, las cuales permiten pronosticar y prevenir infecciones o enfermedades por medio de componentes analíticos, con millones de datos que son capturados desde diversas fuentes, almacenados en repositorios y procesados con diferentes algoritmos que facilitan su interpretación, para la eficiente toma de decisiones y gestión de recursos, no sólo a nivel de salud, sino a nivel gubernamental. Por otra parte, se generan una serie de desafíos investigativos los cuales están asociados en su mayoría a las políticas de privacidad, seguridad y leyes de la información en donde el rol de los entes gubernamentales es fundamental para poder alcanzar una relación de madurez entre la tecnología y salud.

**Palabras clave:** big data, Salud Pública, Sistemas de información, aplicaciones, privacidad, repositorios.

### ABSTRACT

*This article mentions a series of strategies and success stories that, through their correct application, allow improving patient care in public health information systems, in turn, with the optimization of resources, efficiency levels are increased in all care processes such as inventory management, personnel management, prevention and timely diagnosis of*

---

Citación: Lopez Copete, A. F., Rodriguez Martinez, L. ., & Ramirez Gómez, D. A. (2023). Aplicación de Big Data en Sistemas de Salud pública. *Publicaciones E Investigación*, 17(1). <https://doi.org/10.22490/25394088.6446>

<sup>1</sup><https://orcid.org/0000-0002-1280-8854> / [andrescopete9370@gmail.com](mailto:andrescopete9370@gmail.com)

<sup>2</sup><https://orcid.org/0000-0001-7066-1208> / [l.rodriguez881107@gmail.com](mailto:l.rodriguez881107@gmail.com)

<sup>3</sup><https://orcid.org/0000-0003-1448-0212> / [danielramirezgomez244@gmail.com](mailto:danielramirezgomez244@gmail.com)

---

<https://110.22490/25394088.6446>

*infections. All this is achieved using big data tools, which allow forecasting and preventing infections or diseases through analytical components, with millions of data that are captured from diverse sources, stored in repositories, and processed with different algorithms that facilitate their interpretation, for efficient decision-making and resource management, not only at the health level, but also at the government level. On the other hand, a series of investigative challenges are generated which are mostly associated with privacy policies, security, and information laws where the role of government entities is essential to be able to achieve a mature relationship between technology and health.*

**Key words:** : *Big data, Public Health, Information Systems, applications, privacy, repositories.*



## 1. INTRODUCCIÓN<sup>1,2</sup>

El cambio más significativo en la investigación en salud es la aceptación cada vez más frecuente del uso de datos a gran escala, lo que conlleva a una reconfiguración de los sistemas de información en salud, los cuales presentan déficit en la forma en la cual se genera la atención de los pacientes, por ejemplo, en los países de la OCDE (Organización para la Cooperación y Desarrollo Económico) un 20 por ciento de sus gastos en salud son ineficientes (Haring, 2021), dado que un alto número de pacientes no son atendidos de forma oportuna y no son medicados correctamente. Esto ocasiona que aproximadamente 2 millones de personas en todo el mundo no tengan acceso a medicamentos esenciales y muchos de ellos son de una calidad inferior al momento de ser recibidos (Ozawa *et al.*, 2019), por lo cual se generan cuellos de botella en los procesos de distribución y atención y pone en riesgo vidas humanas por no tener implementados sistemas de información adecuados.

La implementación de *big data* en sistemas de información de salud pública permite tanto a las organizaciones involucradas en temas de salud pública, como a los ministerios de salud, avanzar con la detección temprana de condiciones clínicas, identificación de patrones de comportamiento de pacientes, entre otros beneficios, encaminados a la construcción de políticas

de salud pública a nivel nacional. Esto se lograría con una serie de estrategias haciendo uso de herramientas *big data*, por ejemplo: el procesamiento por lotes MapReduce, Apache Spark, entre otros, alimentados por medio de repositorios centrales de datos médicos históricos. Las fuentes de información de *big data* que se aplican en los sistemas de salud pública son diversas con infinidad de variables asociadas a indicadores dependientes del tema de salud a tratar. Los datos son robustos y útiles si se les da un manejo adecuado, pero aún existen múltiples desafíos investigativos por enfrentar, como lo son: el uso adecuado de tecnologías que garanticen la calidad de los datos, contar con arquitecturas que permitan la captura, transformación y análisis de los datos y garantías de privacidad y seguridad de la información.

La implementación de estrategias de big data cuenta con una serie de casos de éxito que reporta la literatura, como lo son los casos del Brigham and Women's Hospital, Wissenschaftliches Institut, Rizzoli Orthopedic Institute entre otros que se detallarán más adelante.

## 2. DESARROLLO

A continuación, se mencionan algunos conceptos clave presentes en este artículo:

### **Capa de gobierno de datos:**

Se compone por la gestión de data maestros (MDM), en el cual se encuentran los procesos,

<sup>1</sup>Declaración de intereses: declaramos no tener ningún conflicto de intereses, que puedan haber influido en los resultados obtenidos o las interpretaciones propuestas.

<sup>2</sup>Declaración de consentimiento informado: el estudio se realizó respetando el Código de ética y buenas prácticas editoriales de publicación.

gobiernos, normas o reglas y herramientas para gestionar los datos. Por otra parte, está la gestión de ciclo de vida del dato, en donde se aplica el proceso de gestión de información de empresarial a los datos, desde su inicio hasta el final de su ciclo útil o desecho; gestionar la vida útil de los datos proporciona a las empresas alta competitividad en el mercado. Por último, seguridad y privacidad del dato, en donde las compañías deben implementar reglas a los datos y otros mecanismos de control a nivel de seguridad dada la sensibilidad de los datos, la protección de la privacidad del dato y evitar cualquier tipo de violación a los mismos. En resumen, este concepto se enfoca en el “cómo” del dato “Cómo aprovechar los datos en la organización” (Wang *et al.*, 2018).

#### **Big data:**

Colección de datos cuyo contenido no puede ser procesado con herramientas de software convencionales. Se caracteriza por 5V: volumen, variedad, velocidad, valor y veracidad (Wang *et al.*, 2021).

#### **Sistemas de información:**

Según K. & J. Laudon (1996), se define un sistema de información como: “el conjunto de componentes interrelacionados que capturan, almacenan, procesan y distribuyen la información para apoyar la toma de decisiones, el control, análisis y visión de una organización”.

#### **Privacidad y seguridad de la información:**

Se compone de los procesos que incluyen sesgos inadvertidos para impedir la violación de la privacidad y seguridad, definiendo relaciones con la propiedad de los datos, transparencia de integración e intercambio con diferentes sistemas (Saunders *et al.*, 2020).

### **3. METODOLOGÍA**

La metodología utilizada para realizar esta revisión sistemática de literatura es la propuesta por Pérez\_Rave (2012; 2019).

El propósito de este artículo es brindar respuesta a las siguientes preguntas de investigación:

**P1:** ¿Qué estrategias de *big data* se aplican en los sistemas de salud pública?

**P2:** ¿Qué casos de éxito reporta la literatura frente a la implementación de estrategias de *big data* sobre sistemas de salud pública?

**P3:** ¿Cuáles arquitecturas de referencia existen para implementar un modelo *big data* enfocado a sistemas de información en salud pública?

**P4:** ¿Cómo desde *big data* se pueden optimizar recursos y tiempos de atención en sistemas de salud pública?

**P5:** ¿Cuáles son los desafíos investigativos en la aplicación de *big data* en sistemas de información en salud pública?

Según las preguntas mencionadas previamente, se realizó una búsqueda en la base de datos Scopus aplicando la siguiente ecuación.

*TITLE-ABS-KEY ( “big data” AND “application” AND “in” AND “public health” ) AND ( LIMIT-TO ( DOCTYPE , “ar” ) OR LIMIT-TO ( DOCTYPE , “cp” ) ) AND ( LIMIT-TO ( SRCTYPE , “j” ) )*

La búsqueda se realizó entre marzo y abril de 2022 y esta arrojó un total de 142 resultados.

Se limitaron los resultados a las publicaciones escritas en inglés, se excluyeron artículos de revisión (*review*) y sólo se incluyeron artículos de investigación y artículos de conferencia. Se limitaron los resultados a las publicaciones de los últimos cinco años (2018-2022).



Figura 1. Número de publicaciones por año en el tema *big data* en sistemas de salud pública.  
Fuente: elaboración propia.

A partir de los resultados obtenidos en la búsqueda se construyó el diagrama de Pareto mostrado en la Figura 2. La totalidad de artículos encontrados en la búsqueda fue de 25. Del diagrama, se concluye que

el 42.3 % de éstos (los 18 más citados) constituyen lo “poco vital” al sumar entre ellos el 95.1 % del total de citaciones.

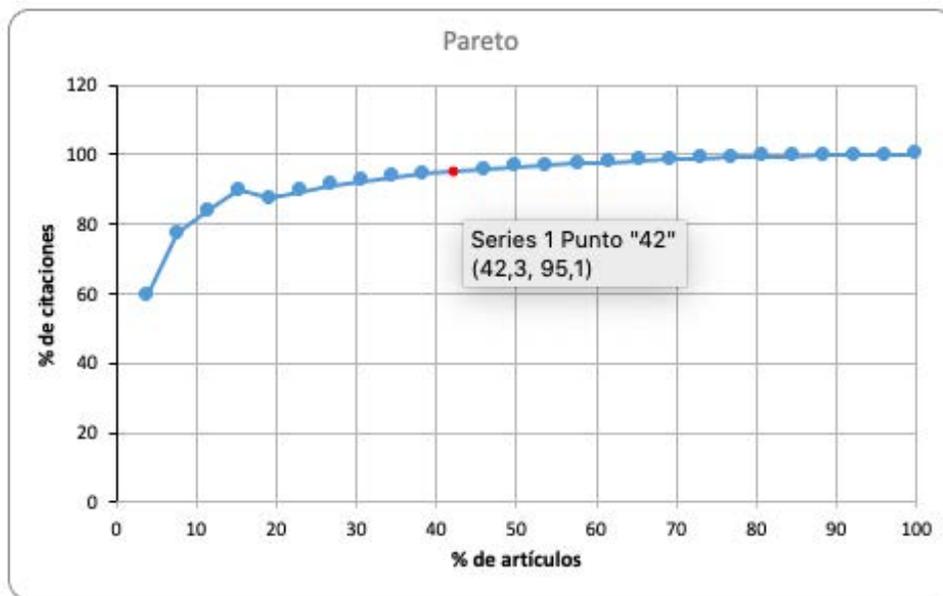


Figura 2. Diagrama de Pareto para los artículos encontrados.  
Fuente: Pérez-Rave (2019).

Con base en los insumos obtenidos desde la medición del Pareto, se realizó un análisis de los títulos y resúmenes de algunos de los artículos, para verificar

que estuvieran relacionados con el tema de estudio y finalmente el resultado fue de diez y ocho (18) artículos, los cuales se listan a continuación:

**TABLA 1.**  
**Artículos seleccionados según relevancia, vigencia y acceso**

Artículo	Citaciones
Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations.	583
Big data in healthcare: management, analysis and future prospects.	175
Big data science: Opportunities and challenges to address minority health and health disparities in the 21st century. Ethnicity and Disease.	66
Taking the pulse of COVID-19: a spatiotemporal perspective.	58
Big Data Analytics in the Fight against Major Public Health Incidents (Including COVID-19): A Conceptual Framework.	17
Integrating Digital Technologies and Public Health to Fight Covid-19 Pandemic: Key Technologies, Applications, Challenges and Outlook of Digital Healthcare.	12
Bangladesh's digital health journey: reflections on a decade of quiet revolution.	12
A review of Big Data analytics and potential for implementation in the delivery of global neurosurgery.	8
Beyond Public Health Genomics: Can Big Data and Predictive Analytics Deliver Precision Public Health?	7
Application of Big Data to Support Evidence-Based Public Health Policy Decision-Making for Hearing.	7
Using big data analytics to improve HIV medical care utilization in South Carolina: A study protocol.	6
Meaningful Big Data Integration for a Global COVID-19 Strategy.	5
Modeling and tracking Covid-19 cases using Big Data analytics on HPCC system platform.	5
Big data opportunities: System health monitoring and management.	4
The Premise and Promise of Big Data for Tracking Population Health: Big Deal or Big Disappointment? Digestive Diseases and Sciences.	1
The application framework of big data technology in the COVID-19 epidemic emergency management in local government-a case study of Hainan Province, China.	1
Big tech, big data and the new world of digital health.	1
Big Data Warehouse for Healthcare-Sensitive Data Applications.	1

Fuente: elaboración propia.

**TABLA 2.**  
**Artículos seleccionados de los últimos 5 años**

<b>Id</b>	<b>Artículo</b>	<b>Año</b>
1	Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations.	2018
2	Big data in healthcare: management, analysis and future prospects.	2019
3	Big data science: Opportunities and challenges to address minority health and health disparities in the 21st century. Ethnicity and Disease.	2017
4	Taking the pulse of COVID-19: a spatiotemporal perspective.	2020
5	Big Data Analytics in the Fight against Major Public Health Incidents (Including COVID-19): A Conceptual Framework.	2020
6	Integrating Digital Technologies and Public Health to Fight Covid-19 Pandemic: Key Technologies, Applications, Challenges and Outlook of Digital Healthcare.	2021
7	Bangladesh's digital health journey: reflections on a decade of quiet revolution.	2019
8	A review of Big Data analytics and potential for implementation in the delivery of global neurosurgery.	2018
9	Beyond Public Health Genomics: Can Big Data and Predictive Analytics Deliver Precision Public Health?	2018
10	Application of Big Data to Support Evidence-Based Public Health Policy Decision-Making for Hearing.	2020
11	Using big data analytics to improve HIV medical care utilization in South Carolina: A study protocol.	2019
12	Meaningful Big Data Integration for a Global COVID-19 Strategy.	2020
13	Modeling and tracking Covid-19 cases using Big Data analytics on HPCC system platform.	2021
14	Big data opportunities: System health monitoring and management.	2019
15	The Premise and Promise of Big Data for Tracking Population Health: Big Deal or Big Disappointment? Digestive Diseases and Sciences.	2017
16	The application framework of big data technology in the COVID-19 epidemic emergency management in local government-a case study of Hainan Province, China.	2021
17	Big tech, big data and the new world of digital health.	2021
18	Big Data Warehouse for Healthcare-Sensitive Data Applications.	2021

Fuente: elaboración propia.

#### 4. RESULTADOS Y DISCUSIÓN

Respecto a las preguntas de investigación, se brinda respuesta a cada una de ellas a continuación:

**P1: ¿Qué estrategias de *big data* se aplican en los sistemas de salud pública?**

Conforme a la revisión sistemática de literatura para la pregunta P1, se identifican características comunes en los artículos [1, 2, 3, 4, 5, 6, 7, 10, 12, 18], que individualizan la estrategia más utilizada, la cual se centra

en las fuentes de datos obtenidas por medio de registros de salud electrónicos como insumo primordial de información (Figura 3), destacando la importancia de estos como proveedor determinante en aspectos sociales y de vigilancia en salud pública. También se estableció que en la era de la información los almacenes de datos (*data warehouse*) son puntos estratégicos en la recolección de datos obtenidos de diferentes fuentes como lo son, inteligencia artificial, redes sociales, aplicaciones de recopilación de datos, sensores entre otros, con datos estructurados y no estructurados para el análisis y la toma de decisiones.

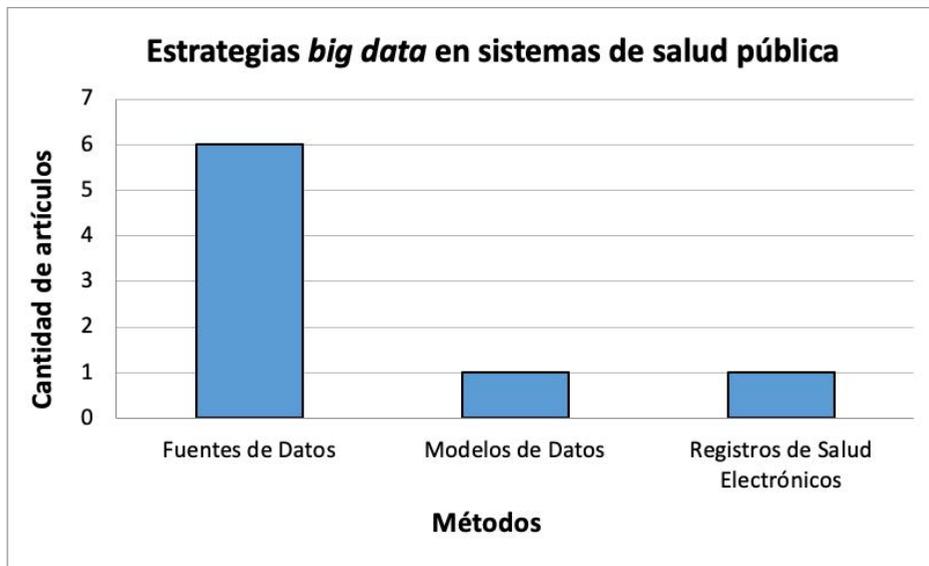


Figura 3. Estrategias big data sistemas de salud pública.

**P2: ¿Qué casos de éxito reporta la literatura frente a la implementación de estrategias de *big data* sobre sistemas de salud pública?**

De acuerdo con la revisión realizada se encontraron casos de éxito asociados a la implementación de *big data* en sistemas de salud pública visualizados en la siguiente tabla:

**TABLA 3.**  
Casos de éxito encontrados en la literatura

Artículo	Descripción	Año	Fuente
Plataforma MIDAS	Herramienta que se ha ejecutado en redes de autoridades sanitarias, la cual puede extraer datos de fuentes públicas o privadas, implementado para la toma de decisiones en personas con probable HF.	2022	<i>MIDAS – Meaningful Integration of Data Analytics and Services.</i> (n.d.). Retrieved May 18, 2022, from <a href="http://www.midasproject.eu/">http://www.midasproject.eu/</a>
Wissenschaftliches Institut der AOK (WIdO) (IBM)	Optimización de tiempo de respuesta al consultar información de pacientes en una base de datos de 24 millones de registros con 10tb de consumo	2022	<i>WIdO – Wissenschaftliches Institut der AOK.</i> (n.d.). Retrieved May 18, 2022, from <a href="https://www.wido.de/">https://www.wido.de/</a>
Brigham and Women's Hospital	Prevención y tratamiento mediante análisis de datos para enfermedades reumáticas	2022	<i>BWH Awards, Honors &amp; Grants - Brigham and Women's Hospital.</i> (n.d.). Retrieved May 18, 2022, from <a href="https://www.brighamandwomens.org/about-bwh/newsroom/awards-honors-grants-detail?id=3916">https://www.brighamandwomens.org/about-bwh/newsroom/awards-honors-grants-detail?id=3916</a>
Premier healthcare alliance	Mediante métodos de análisis de datos, se desea crear un plan para disminuir el gasto en salud y disminuir hospitalizaciones	2022	<i>Premier Healthcare Alliance trusts IBM to deliver comprehensive healthcare solution - IBM Media Center.</i> (n.d.). Retrieved May 18, 2022, from <a href="https://mediacenter.ibm.com/media/Premier+Healthcare+Alliance+trusts+IBM+to+deliver+comprehensive+healthcare+soluti on/1_p3weduvd">https://mediacenter.ibm.com/media/Premier+Healthcare+Alliance+trusts+IBM+to+deliver+comprehensive+healthcare+soluti on/1_p3weduvd</a>
Rizzoli Orthopedic Institute	Se usó big data para diagnósticos ortopédicos	2022	<i>IBM: Research Helps Italian Orthopedic Institute Perform Deep Analytics to Treat Rare Skeletal Diseases   Market-Screener.</i> (n.d.). Retrieved May 18, 2022, from <a href="https://www.marketscreener.com/quote/stock/INTERNATIONAL-BUSINESS-MA-4828/news/IBM-Research-Helps-Italian-Orthopedic-Institute-Perform-Deep-Analytics-to-Treat-Rare-Skeletal-Dise-13281379/">https://www.marketscreener.com/quote/stock/INTERNATIONAL-BUSINESS-MA-4828/news/IBM-Research-Helps-Italian-Orthopedic-Institute-Perform-Deep-Analytics-to-Treat-Rare-Skeletal-Dise-13281379/</a>
Fondazione IRCCS Istituto Nazionale dei Tumori	Estudio del cáncer de cabeza y cuello mediante ciencia de datos con <i>big data</i> .	2022	<i>IBM's Biomedical Analytics Platform Helps Doctors Personalize Treatment   Fierce Healthcare.</i> (n.d.). Retrieved May 18, 2022, from <a href="https://www.fiercehealthcare.com/it/ibm-s-biomedical-analytics-platform-helps-doctors-personalize-treatment">https://www.fiercehealthcare.com/it/ibm-s-biomedical-analytics-platform-helps-doctors-personalize-treatment</a>
Fraunhofer FOKUS	Desarrolla e investiga nuevas herramientas para procesamiento y análisis de altos volúmenes de datos.	2022	<i>IBM and Fraunhofer bring Quantum Computing to Germany.</i> (n.d.). Retrieved May 18, 2022, from <a href="https://www.fraunhofer.de/en/press/research-news/2020/march/ibm-and-fraunhofer-bring-quantum-computing-to-germany.html">https://www.fraunhofer.de/en/press/research-news/2020/march/ibm-and-fraunhofer-bring-quantum-computing-to-germany.html</a>
Leeds Teaching Hospitals	Mediante los datos de pacientes se crean tendencias de posibles emergencias y atención que puedan requerir.	2013	<i>CASE STUDY.</i> (2013). <a href="http://www.intel.com/performance">http://www.intel.com/performance</a> .
Beth Israel Deaconess Medical Center	Se crearon aplicaciones con <i>big data</i> para prevención y control de enfermedades de forma oportuna.	2022	<i>Using reinforcement learning to identify high-risk states and treatments in healthcare - Microsoft Research.</i> (n.d.). Retrieved May 18, 2022, from <a href="https://www.microsoft.com/en-us/research/blog/using-reinforcement-learning-to-identify-high-risk-states-and-treatments-in-healthcare/">https://www.microsoft.com/en-us/research/blog/using-reinforcement-learning-to-identify-high-risk-states-and-treatments-in-healthcare/</a>
Atlantic Health System	Se usan los datos para mejorar el cuidado de los pacientes.	2022	<i>Atlantic Health System Realizes Greater Efficiency Using EMC Backup and Recovery Solutions with Intel® Xeon® Processors.</i> (n.d.).
Texas Health Harris Methodist Hospital	Aplicación con datos de los pacientes, que apoya la toma de decisiones de los médicos en pro de mejorar la calidad de vida de los pacientes.	2022	<i>Modern Healthcare.</i> (n.d.). Retrieved May 18, 2022, from <a href="https://www.modernhealthcare.com/">https://www.modernhealthcare.com/</a>

Artículo	Descripción	Año	Fuente
Mount Sinai Medical Center	Uso de big data para el análisis de resultados de laboratorio para mejorar el cuidado del paciente.	2022	MIT Technology Review. (n.d.). Retrieved May 18, 2022, from <a href="https://www.technologyreview.com/">https://www.technologyreview.com/</a>
Mission Health	Mejorar los cuidados de la salud, crear nuevas regulaciones en conjunto con los gobiernos locales.	2022	Health Catalyst   Healthcare Data and Analytics Technology and Services. (n.d.). Retrieved May 18, 2022, from <a href="https://www.healthcatalyst.com/">https://www.healthcatalyst.com/</a>
MultiCare Health System	Según datos clínicos del paciente, se determina que tan probable es una hospitalización por problemas cardíacos.	2022	Health Catalyst   Healthcare Data and Analytics Technology and Services. (n.d.). Retrieved May 18, 2022, from <a href="https://www.healthcatalyst.com/">https://www.healthcatalyst.com/</a>

Fuente: elaboración propia.

Adicional a los casos mencionados previamente, se tienen otros adicionales (Wang *et al.*, 2018)

**P3: ¿Cuáles arquitecturas de referencia existen para implementar un modelo *big data* enfocado a sistemas de información en salud pública?**

Conforme a la revisión sistemática de cada uno de los artículos para la pregunta P3, se identifican características comunes en sus arquitecturas en los artículos [1, 2, 5, 6, 7, 10, 12,18] en las que se destacan,

el procesamiento por lotes MapReduce como modelo de programación implementado en grandes colecciones de datos, sistemas distribuidos Hadoop, desde su fichero distribuido HDFS por sus siglas en inglés (Hadoop Distributed File System) abordando la característica más significativa de este sistema de ficheros, la cual es su capacidad para almacenar los archivos en un clúster de varias máquinas, Apache Spark como framework de computación en clúster entre otros, alimentados por medio de repositorios centrales de datos médicos históricos (Figura 4).

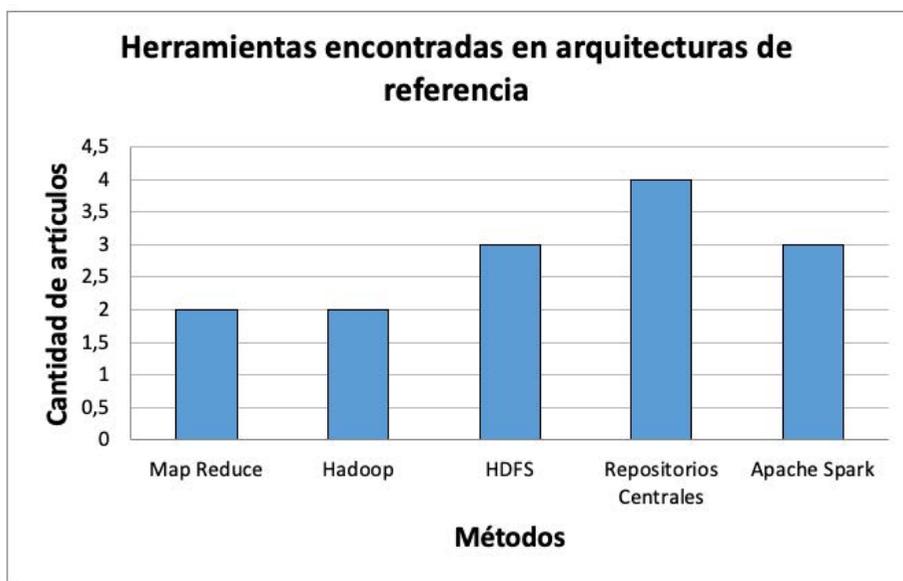


Figura 4. Herramientas encontradas en arquitecturas de referencia. Fuente: elaboración propia.

La Figura 5, ilustra un modelo de arquitectura en el cual se visualizan los componentes requeridos en cada de las capas de manera que permita la explotación de

información para la toma de decisiones, esto mediante reportes *ad-hoc* o aplicaciones que proporcionen datos en tiempo real.

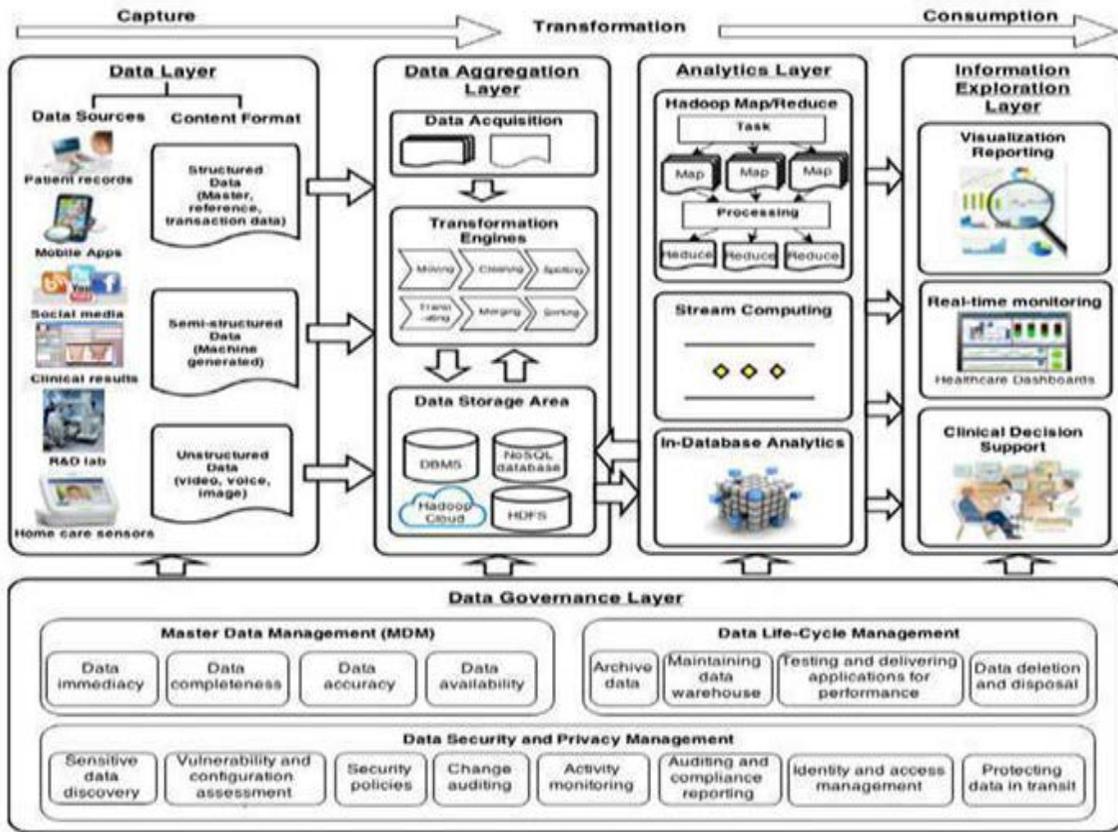


Figura 5. Big data analytics architecture in health care.  
 Fuente: Wang *et al.*, 2018.

La Figura 6, detalla el almacenamiento y procesamiento de datos mediante Map Reduce y *hadoop*,

donde finalmente en un *data warehouse* se consulta la información por parte de los usuarios finales.

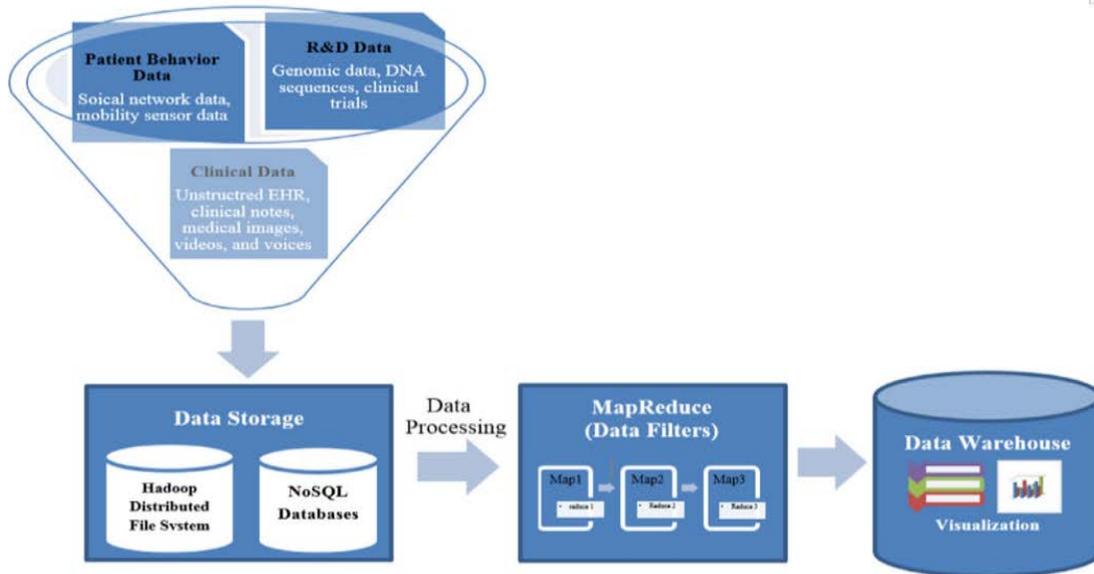


Figura 6. The process of analyzing unstructured data in health care organizations.  
 Fuente: Wang, *et al.*, 2018.

**P4: ¿Cómo desde *big data* se pueden optimizar recursos y tiempos de atención en sistemas de salud pública?**

Se pudo concluir la manera en la que se optimizan tiempos de respuesta para actuar de forma efectiva por medio de visualizaciones que facilitaron a gran medida el análisis para la toma de decisiones en la prevención y control de pandemias, dichas visualizaciones también funcionaron de una manera muy favorable en el área de diagnósticos y predicción de enfermedades, disminuyendo las tasas de reingreso hospitalario, la optimización de dotación de insumos médicos y mejoras notables en las cadenas de suministros. Dentro de las características comunes usadas para la optimización de recursos, se encontró que las técnicas de análisis y visualización más usadas en el ámbito de salud pública fueron las de enfoques usados con georreferenciación, para el monitoreo en tiempo real de tasas de contagio a niveles nacionales e interdepartamentales, aplicaciones de consultas médicas en tiempo real, creación de indicadores de salud para predicciones precisas y tasas de aprendizaje automático.

**P5: ¿Cuáles son los desafíos investigativos en la aplicación de *big data* en sistemas de información en salud pública?**

Conforme a la revisión sistemática de cada uno de los artículos para la pregunta P5 se identifican características comunes en los artículos [1, 2, 3, 4, 5, 7, 8, 18], en los que se determinan que los desafíos investigativos en la aplicación de *big data* en sistemas de información de salud pública centran su atención en el ámbito de la privacidad y la seguridad de la información, la actualización y veracidad de las fuentes de datos y la regulación por parte de los gobiernos de los diferentes países respecto al uso de los datos y un sistema de gobernanza de estos. Desafíos éticos y de confianza, acceso libre y compartido de los datos, políticas de privacidad, además de la necesidad recurrente de integración de sistemas de salud para evitar duplicidad, evitar la inexactitud de los datos por la diversidad y mejorar la calidad de la información para toma de decisiones con una mayor exactitud (Figura 7).

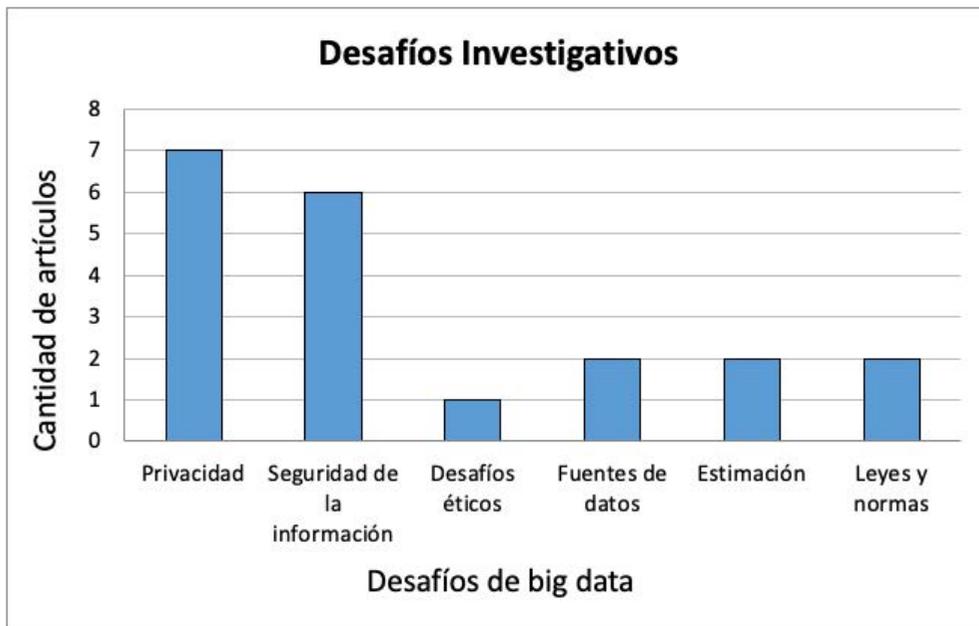


Figura 7. Desafíos investigativos.

Fuente: elaboración propia.

## 5. CONCLUSIONES

Según los resultados obtenidos por medio de la revisión sistemática de literatura según Pérez\_Rave (2012; 2019), la aplicación de *big data* integrada con los sistemas de salud pública permite el desarrollo de modelos analíticos con la integración de diversas fuentes de datos, apoyada con los registros de salud pública de las diferentes entidades, obteniendo una mejor identificación, segmentación y caracterización de los pacientes, generando análisis de tipo micro y macro sobre cualquier población de la cual se tenga información, esto con la finalidad de crear y proyectar una identificación temprana de enfermedades. Por esto es necesario tener toda la información de las diferentes fuentes y por parte de las entidades de salud pública, definir estándares globales en el registro que hagan más fácil su obtención para su posterior procesamiento.

Con relación a la implementación, esta determina un camino de mayor efectividad en el control de enfermedades con ayuda de diagnósticos como detecciones tempranas, dado que actualmente se encuentran muchos casos de éxito sobre la implementación de arquitecturas *big data*, teniendo en cuenta que la data se modela y transforma hacia la realidad a la cual se quiere proyectar, es necesario tener claro que los modelos de transformación permiten indagar hasta lo más específico que se desee y permita la data, con el fin de valorar caminos de mayor accesibilidad. En la población actual la deficiencia en el sistema de salud pública es muy alta, además que por la georreferenciación muchas poblaciones están apartadas de los sistemas de salud, lo cual genera panoramas no reconocibles en la transmisión de enfermedades; con la aplicación de *big data* se puede proveer soluciones tempranas, y la consolidación de un sistema de salud de inclusión global.

Respecto a las arquitecturas de *big data*, con la cantidad de información que se puede obtener desde las diferentes fuentes de datos, es necesario tener repositorios centralizados y estandarizados, esto con el fin de permitir que la ejecución de los modelos analíticos

sea más eficiente y de rápido acceso para la toma de decisiones, cabe resaltar que en la actualidad hay procesos de digitalización de la información que se han llevado a cabo en sistemas de salud de Estados Unidos, definidos como EHR (Registros de Salud Electrónico), EMR (Registro Médico Electrónico), PHR (Registro de Salud Personal) (Dash *et al.*, 2019) lo cual demuestra el direccionamiento actual y el manejo que se le quiere dar a la información poblacional.

Es evidente que la aplicación de *big data* implica mejoras en la calidad de atención, procesos logísticos de entrega de medicamentos y prevenciones tempranas, disminuyendo procesos de reincidencia en las consultas dado que se puede brindar un diagnóstico efectivo apoyado en información histórica categorizada. Adicionalmente se pueden orientar políticas dentro de los sistemas de atención actual, minimizando costos de operación y a la vez aumentando la eficiencia.

También se observa que hay muchos desafíos para lograr una implementación robusta con respecto a las políticas de seguridad de la información, las cuales son determinadas por entes gubernamentales, puesto que su nivel es de un alto grado de restricción. Es necesario definir políticas de acceso a la información que sean de obligatorio cumplimiento, para determinar prioridades en la protección de información en salud, en vista de que los datos son altamente heterogéneos y pueden obtenerse de diferentes tipos y fuentes.

Se busca vislumbrar como conclusión principal de esta investigación que, con la aplicación de *big data* en sistemas de salud pública, se genera un apoyo determinante en la mejora de la calidad de la atención poblacional, diagnósticos tempranos, reducción de costos operativos y efectividad en el accionar y toma de decisiones internas como externas. Además de ofrecer muchos rubros de acción con procesos de impacto a mediano y largo plazo que mejoren la integralidad y accesibilidad hacia el individuo como población en general.

## REFERENCIAS

- <https://apps.who.int/iris/handle/10665/329331>
- Dash, S., Shakyawar, S. K., Sharma, M., & 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/FIGURES/6>
- Haring, R. (Ed.). (2021). *Handbook of Global Health*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-45009-0>
- Hernández Trasobares, A. (2003). Los sistemas de información: evolución y desarrollo. *Proyecto social: Revista de relaciones laborales*, 10-11, 149-165.
- Jia, Q., Guo, Y., Wang, G., & Barnes, S. J. (2020). Big Data Analytics in the Fight against Major Public Health Incidents (Including COVID-19): A Conceptual Framework. *International Journal of Environmental Research and Public Health* 2020, 17(17), 6161. <https://doi.org/10.3390/IJERPH17176161>
- Khan, M. A. H., Cruz, V. O., & Azad, A. K. (2019). Bangladesh's digital health journey: reflections on a decade of quiet revolution. *WHO South-East Asia Journal of Public Health*, 8(2), 71–76. <https://doi.org/10.4103/2224-3151.264849>
- Khoury, M. J., Engelgau, M., Chambers, D. A., & Mensah, G. A. (2018). Beyond Public Health Genomics: Can Big Data and Predictive Analytics Deliver Precision Public Health? *Public Health Genomics*, 21(5–6), 244–250. <https://doi.org/10.1159/000501465>
- Madanian, S., Parry, D. T., Airehrour, D., & Cherrington, M. (2019). mHealth and big-data integration: promises for healthcare system in India. *BMJ Health & Care Informatics*, 26(1), e100071. <https://doi.org/10.1136/BMJHCI-2019-100071>
- Mansoor, E., & Al-Kindi, S. G. (2017). The Premise and Promise of Big Data for Tracking Population Health: Big Deal or Big Disappointment? *Digestive Diseases and Sciences*, 62(3), 562–563. <https://doi.org/10.1007/S10620-017-4458-5>
- Mansoor, E., Sadeer, Al-Kindi, G., & Al-Kindi, S. G. (2017). The Premise and Promise of Big Data for Tracking Population Health: Big Deal or Big Disappointment? *Digestive Diseases and Sciences*, 62(3), 562–563. <https://doi.org/10.1007/S10620-017-4458-5>
- Mao, Z., Zou, Q., Yao, H., & Wu, J. (2021). The application framework of big data technology in the COVID-19 epidemic emergency management in local government—a case study of Hainan Province, China. *BMC Public Health*, 21(1), 1–19. <https://doi.org/10.1186/S12889-021-12065-0/TABLES/2>
- Olatosi, B., Zhang, J., Weissman, S., Hu, J., Haider, M. R., & Li, X. (2019). Using big data analytics to improve HIV medical care utilisation in South Carolina: A study protocol. *BMJ Open*, 9(7), e027688. <https://doi.org/10.1136/BMJOPEN-2018-027688>
- Ozawa, S., Shankar, R., Leopold, C., & Orubu, S. (2019). Access to medicines through health systems in low- and middle-income countries. *Health Policy and Planning*, 34(Supplement\_3), iii1–iii3. <https://doi.org/10.1093/HEAPOL/CZZ119>
- Pérez Rave, J. (2012). *Revisión sistemática de literatura en ingeniería*. Medellín: UdeA.
- Pérez Rave, J. (2019). *Revisión sistemática de literatura en ingeniería ampliada y actualizada*. Segunda edición. Medellín: UdeA.
- Saunders, G. H., Christensen, J. H., Gutenberg, J., Pontoppidan, N. H., Smith, A., Spanoudakis, G., & Bamiou, D. E. (2020). Application of Big Data to Support Evidence-Based Public Health Policy Decision-Making for Hearing. *Ear and Hearing*, 41(5), 1057–1063. <https://doi.org/10.1097/AUD.0000000000000850>
- Shahid, A., Nguyen, T. A. N., & Kechadi, M. T. (2021). Big Data Warehouse for Healthcare-Sensitive Data Applications. *Sensors*, 21(7), 2353. <https://doi.org/10.3390/S21072353>
- Tavazzi, L. (2019). Big data: Is clinical practice changing? *European Heart Journal, Supplement*, 21, B98–B102. <https://doi.org/10.1093/eurheartj/suz034>
- Thomason, J. (2021). Big tech, big data and the new world of digital health. *Global Health Journal*, 5(4), 165–168. <https://doi.org/10.1016/J.GLOHJ.2021.11.003>
- Villanustre, F., Chala, A., Dev, R., Xu, L., LexisNexis, J. S., Furht, B., & Khoshgoftaar, T. (2021). Modeling and tracking Covid-19 cases using Big Data analytics on HPCC system platform. *Journal of Big Data*, 8(1), 1–24. <https://doi.org/10.1186/S40537-021-00423-Z/FIGURES/14>
- Wang, Q., Su, M., Zhang, M., & Li, R. (2021). Integrating Digital Technologies and Public Health to Fight Covid-19 Pandemic: Key Technologies, Applications, Challenges and Outlook of Digital Healthcare. *International Journal of Environmental Research and Public Health* 2021, 18(11), 6053. <https://doi.org/10.3390/IJERPH18116053>
- Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, 126, 3–13. <https://doi.org/10.1016/j.techfore.2015.12.019>
- West, J. L., Fargen, K. M., Hsu, W., Branch, C. L., & Couture, D. E. (2018). A review of Big Data analytics and potential for implementation in the delivery of global neurosurgery. *Neurosurgical Focus*, 45(4), E16. <https://doi.org/10.3171/2018.7.FOCUS18278>
- Yang, C., Sha, D., Liu, Q., Li, Y., Lan, H., Guan, W. W., Hu, T., Li, Z., Zhang, Z., Thompson, J. H., Wang, Z., Wong, D., Ruan, S., Yu, M., Richardson, D., Zhang, L., Hou, R., Zhou, Y., Zhong, C., ... Ding, A. (2020). Taking the pulse of COVID-19: a spatiotemporal perspective. *International Journal of Digital Earth*, 13(10), 1186–1211. <https://doi.org/10.1080/17538947.2020.1809723>
- Zhang, X., Pérez-Stable, E. J., Bourne, P. E., Peprah, E., Duru, O. K., Breen, N., Berrigan, D., Wood, F., Jackson, J. S., Wong, D. W. S., & Denny, J. (2017). Big data science: Opportunities and challenges to address minority health and health disparities in the 21st century. *Ethnicity and Disease*, 27(2), 95–106. <https://doi.org/10.18865/ED.27.2.95>