

Big Data Tools for Enterprise Information Systems: a systematic review

Choquehuanca León Renzo Esteban¹; Palomino Vivanco Sebastian Valentin²

¹Universidad Tecnológica del Perú, Perú, renzort2002@gmail.com

²Universidad Tecnológica del Perú, Perú, sebastianpalomino912@gmail.com

Abstract— This systematic review examines the use of Big Data management tools in enterprise information systems for decision-making, aiming to analyze how these tools mitigate risks in critical decision-making processes. Utilizing the Scopus database and employing the PICO and PRISMA methodologies, 34 relevant articles were selected based on inclusion and exclusion criteria. The findings highlight the viability of data management tools such as SQL Server, Power BI, and Apache Hadoop in enhancing performance, reducing costs, and improving decision-making in organizational systems. The review underscores the importance of integrating these tools to address complex data analysis and decision-making queries in Big Data environments. Key insights reveal that these tools significantly impact operational efficiency, data processing, and strategic planning, making them indispensable for modern enterprises. The study concludes that the effective use of Big Data tools not only optimizes business processes but also supports sustainable organizational growth by enabling data-driven decision-making.

Keywords—Big Data, Apache Hadoop, Power BI, SQL Server, decision-making, business information systems.

Big Data Tools for Enterprise Information Systems: a systematic review

Choquehuanca León Renzo Esteban¹; Palomino Vivanco Sebastian Valentin²

¹Universidad Tecnológica del Perú, Perú, renzort2002@gmail.com

²Universidad Tecnológica del Perú, Perú, sebastianpalomino912@gmail.com

Resumen— Esta revisión sistemática examina el uso de herramientas de gestión de Big Data en los sistemas de información empresarial para la toma de decisiones, con el objetivo de analizar cómo estas herramientas mitigan los riesgos en los procesos de toma de decisiones críticas. Utilizando la base de datos Scopus y aplicando las metodologías PICO y PRISMA, se seleccionaron 34 artículos relevantes basados en criterios de inclusión y exclusión. Los hallazgos destacan la viabilidad de herramientas de gestión de datos como SQL Server, Power BI y Apache Hadoop son fundamentales para mejorar el rendimiento en los sistemas de información, optimizar el análisis de datos y reducir costos operativos, perfeccionando la toma de decisiones en los sistemas organizacionales. La revisión subraya la importancia de integrar estas herramientas para abordar consultas complejas de análisis de datos y toma de decisiones en entornos de Big Data. Los resultados clave revelan que estas herramientas tienen un impacto significativo en la eficiencia operativa, el procesamiento de datos y la planificación estratégica, convirtiéndolas en elementos indispensables para las empresas modernas. El estudio concluye que el uso efectivo de las herramientas de Big Data no solo optimiza los procesos empresariales, sino que también apoya el crecimiento sostenible de las organizaciones al permitir una toma de decisiones basada en datos.

Palabras clave: Big Data, Apache Hadoop, Power BI, SQL Server, toma de decisiones, sistemas de información empresarial.

I. INTRODUCCIÓN

En los últimos años, las empresas que gestionan grandes volúmenes de información han buscado implementar soluciones tecnológicas que permitan un análisis de datos rápido y preciso. El acelerado avance tecnológico ha posicionado la gestión eficiente de la información como uno de los pilares fundamentales para el éxito de una organización [1]. En este contexto, frente a una competencia comercial cada vez más globalizada, las empresas requieren herramientas de Big Data capaces de analizar grandes cantidades de datos, de diversos tipos, a alta velocidad, para resolver problemas de gran escala y complejidad [2]. El uso de Big Data en la toma de decisiones, junto con el avance del aprendizaje automático en los procesos empresariales, ha permitido una reducción generalizada de la carga de trabajo y una mayor agilidad en la toma de decisiones, logrando una disminución del tiempo de respuesta entre un 5% y un 30% a gran escala [2]. En este sentido, aplicaciones especializadas

en el procesamiento de datos, como Power BI, han contribuido a mejorar la toma de decisiones al implementar soluciones integrales en sistemas empresariales, utilizando ETL y KPI para facilitar la interpretación de los resultados [1]. Asimismo, herramientas como Apache Hadoop y SQL Server han permitido reducir los tiempos de análisis, disminuir costos y optimizar las respuestas orientadas a la mejora de procesos [3]. Por ejemplo, se ha utilizado un sistema de gestión en empresas manufactureras con el objetivo de reducir costos y mejorar el tiempo de respuesta [4].

Como consecuencia de lo anterior, el uso inadecuado de las herramientas de gestión de datos puede generar dificultades en el entorno empresarial, afectando la administración y la obtención de resultados óptimos en la toma de decisiones. Debido a este uso incorrecto, se implementó un almacén de datos mediante el proceso ETL (Extract, Transform, Load, por sus siglas en inglés, que implica extraer datos de diversas fuentes, transformarlos para su análisis y cargarlos en un almacén de datos), y posteriormente se analizó la información a través de paneles dinámicos que permitieron visualizar las métricas comerciales [1]. Por lo tanto, la falta de herramientas adecuadas o su uso incorrecto compromete la toma de decisiones críticas en las industrias, ya que no se cuenta con soluciones que permitan optimizar y reducir estos riesgos.

El objetivo de esta revisión sistemática de literatura es analizar y describir el uso de herramientas de Big Data en el marco de los sistemas de información empresarial para la toma de decisiones, con el propósito de que estas decisiones estén fundamentadas en evidencias y se minimicen los riesgos asociados a la toma de decisiones críticas. La relevancia de este estudio radica en que el uso adecuado del Big Data no solo permite optimizar y agregar valor a los procesos empresariales, sino que también contribuye a una mejor planificación estratégica y al crecimiento sostenible de las organizaciones. Asimismo, resulta fundamental examinar los enfoques actuales que integran técnicas de Aprendizaje Automático en el análisis de datos, con el fin de identificar sus limitaciones y explorar oportunidades de mejora [5]. Además de abordar los desafíos actuales en la detección de anomalías dentro del análisis de Big Data con fines empresariales, esta investigación busca identificar propuestas de algoritmos más eficientes y precisos, lo que podría impulsar avances significativos en este campo.

El estudio está organizado de la siguiente manera. La sección 2, Metodología, presenta los métodos utilizados para la elaboración de la revisión sistemática de literatura, incluyendo la estrategia PICO y un protocolo transparente de búsqueda y selección de literatura basado en la declaración PRISMA. En la sección 3, Resultados, se describen los datos recopilados a partir de la revisión de artículos científicos indexados, señalando los resultados clave para alcanzar el objetivo de esta revisión. En la sección 4 se plantea la discusión de los resultados obtenidos, proporcionando una visión crítica del uso de las herramientas Big Data en el marco de los sistemas de información empresarial. Finalmente, en la sección 5, se plantean las conclusiones del estudio.

II. METODOLOGÍA

A. Metodología PICO

Este estudio se ha desarrollado mediante una metodología de revisión sistemática de literatura. Como parte del proceso de sistematización de la búsqueda, selección y análisis de la información bibliográfica, se utilizó la estrategia PICO (acrónimo de Problema, Intervención, Comparación y Resultados, Contexto) para enunciar una pregunta de revisión estructurada, que abarque todas las dimensiones clave del problema planteado. La pregunta de revisión planteada es la siguiente: ¿Cómo se están utilizando las herramientas Big Data en el marco de los sistemas de información empresarial para la toma de decisiones? A partir de esta pregunta de revisión, se identificaron palabras clave para cada componente de PICOC, excluyéndose el componente Comparación, por no ser pertinente al problema (ver Tabla I).

TABLA I. COMPONENTES PICO

Siglas	Significado	Palabra general	Palabra clave
P	Problema / Población	Sistema de información empresarial	"business information system", "management systems"
I	Intervención	Big data	"big data", "data analytics", "data science", "large data", "Apache Hadoop", "Power BI", "SQL Server"
C	Comparación	Sin comparación	---
O	Resultados	Toma de decisiones	"decision making"
C	Contexto	Empresa	business, enterprise, corporate

A partir de los componentes de PICO se propusieron las siguientes subpreguntas de revisión:

- RQ1: ¿En qué contexto y para qué sistemas de información empresarial se están utilizando herramientas Big Data?
- RQ2: ¿Cómo se utilizan las herramientas Big Data para la toma de decisiones?

- RQ3: ¿Qué características presentan las herramientas de gestión de datos?
- RQ4: ¿Cuáles son los resultados obtenidos en la toma de decisiones usando las herramientas Big Data?

Se realizó una búsqueda sistemática en la base de datos Scopus con el fin de encontrar artículos pertinentes, mediante una ecuación con las palabras clave previamente definidas:

TITLE-ABS-KEY ("big data" OR "data analytics" OR "data science" OR "large data" OR "Apache Hadoop" OR "Power BI" OR "SQL Server") AND ("business information system" OR "management systems") AND ("decision making") AND ("business" OR "enterprise" OR "corporate")

B. Identificación de estudios

La aplicación de la ecuación de búsqueda generó un total de 823 documentos recuperados. Estos fueron cribados en función del cumplimiento de todos los siguientes criterios de inclusión:

TABLA II. CRITERIOS (INCLUSIÓN Y EXCLUSIÓN)

Criterios de inclusión	
C.I.1	Estudios que apliquen o describan el uso de herramientas Big Data para la toma de decisiones en el marco de sistemas de información empresarial
C.I.2	Estudios que reporten resultados originales
Criterios de exclusión	
C.E.1	Estudios iniciales (conference papers), literatura científica menor (notas, cartas al editor) o revisiones de literatura
C.E.2	Estudios en idiomas distintos al inglés o español
C.E.3	Artículos publicados antes de 2020

El proceso de selección de artículos se realizó siguiendo de forma general los lineamientos descritos en la declaración PRISMA (ver Fig.1).

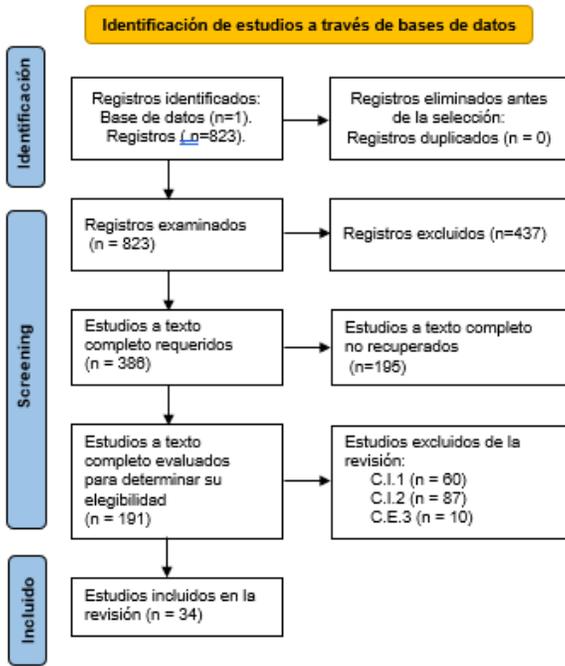


Fig. 1 Diagrama PRISMA

En el proceso de cribado, se pueden identificar tres fases:

- En una primera fase se evaluaron los 823 artículos recuperados a partir de una revisión de sus títulos, resúmenes y data bibliográfica. En esta fase, los registros excluidos por no ajustarse a los criterios de inclusión y exclusión fueron 437.
- En la siguiente fase se lograron recuperar a texto completo 191 estudios de los 386 inicialmente seleccionados.
- Los estudios recuperados (n=191) fueron evaluados a texto completo en función de los criterios establecidos previamente. 157 estudios no cumplieron con al menos uno de los criterios de inclusión o exclusión. Finalmente, fueron seleccionados 34 estudios para formar parte de la RSL.

III. RESULTADOS

Las subpreguntas de revisión se han utilizado para organizar los resultados obtenidos:

RQ1. ¿En qué contexto y para qué sistemas de información empresarial se están utilizando herramientas Big Data?

Como se aprecia en la tabla III, un 29% de los estudios no precisan el contexto productivo en que se desarrolla la investigación. Del resto de estudios, un 27% de las investigaciones se desarrollan en el marco de empresas proveedoras de servicios de salud, educación y cuidado animal [6],[10],[14]; un 18% en el sector de servicios de TI

[8],[15]; mientras que sólo 3 estudios se desarrollan en el marco de la industria manufacturera [12],[17],[19].

TABLA III. TIPO DE EMPRESAS

N°	Empresas		
	Tipo de empresa	Cantidad	Porcentaje
1	Salud	6	18%
2	Cuidado animal	1	3%
3	Tecnología de la Información (TI)	5	18%
4	Ventas	6	18%
5	Manufactura	3	9%
6	Educación universitaria	2	6%
7	Transporte	1	3%
8	No precisa	10	29%

Más allá del contexto de estudio, se evidencia que las herramientas de Big Data se utilizan en el marco de diferentes sistemas de información empresarial. De acuerdo con la tabla IV, más del 50% del total de los artículos se centran en el análisis de datos en el marco de sistemas de gestión [3],[7],[11]. En cambio, su uso en sistemas empresariales de seguridad [5], clasificación [9] o almacenamiento [13] es marginal.

TABLA IV. TIPO DE SISTEMAS DE INFORMACIÓN EMPRESARIAL

Sistemas de Información Empresarial		
Sistema	Cantidad	Porcentaje
Sistemas de gestión	19	56%
Sistemas de apoyo a la toma de decisiones	7	21%
Sistemas de detección	5	15%
Sistemas de seguridad	1	3%
Sistemas de almacenamiento	1	3%
Sistemas de clasificación	1	3%

De igual manera de que hay diferentes tipos de sistemas de información, se cuenta con indicadores que desempeñan un papel importante en el uso que se le dará a esos datos recopilados llamados KPI (Key Performance Indicator), que permitirá medir el desempeño de un proceso o actividad en función de los objetivos estratégicos establecidos por una organización. Donde, su propósito es proporcionar información clara y objetiva. Estos indicadores tienen que ser medibles y alcanzables y ser usados como una herramienta eficaz para lograr las metas del sector empresarial.

RQ2. ¿Cómo se utilizan las herramientas Big Data en el marco de los sistemas de información empresarial?

Según la información recopilada, en la tabla V se describen los principales usos que se le da a las herramientas de Big Data.

TABLA V. USOS DADOS A LAS HERRAMIENTAS BIG DATA

Nº	Utilización	Cantidad
1	Identificación de datos	17
2	Toma de decisiones	14
3	Gestión de archivos electrónicos	4
4	Transformación de procesos	3
5	Identificación de procesos	2
6	Gestión de inventarios	2
7	Planificación de acciones	1

En relación con el uso de Big Data y según la tabla V, se identificaron 17 artículos que exploran su aplicación en la identificación y manejo de datos. Un estudio clave analiza cómo reducir la carga de trabajo y el tiempo de ejecución en un 5-30 % mediante la optimización del procesamiento de datos [2]. Otro artículo detalla la gestión de archivos electrónicos a gran escala, destacando las funciones y flujos operativos del sistema [9]. También se revisó un artículo que presenta estrategias para la implementación de Big Data en organizaciones, abordando infraestructura, análisis de datos y seguridad [14]. Asimismo, se encontró un esquema para estimar tiempos y minimizar recursos en sistemas Big Data basados en Apache Spark y Hadoop, considerando escenarios de fallos [15].

Así mismo, se encontraron múltiples artículos que enfatizan el papel de los sistemas de apoyo basados en Big Data en la toma de decisiones. Por ejemplo, un estudio destaca la relevancia del acceso en tiempo real a información selectiva como ventaja competitiva, sugiriendo la metodología de Vercellis para implementar un conjunto de KPI en la toma de decisiones [1]. Otro artículo describe un sistema de soporte de decisiones clínicas basado en Big Data para la prescripción de quimioterapia en niños con leucemia linfoblástica aguda (LLA), detectando 735 casos de desviaciones del protocolo y 57 errores en la prescripción (95 %, IC = 1,25-2,55) [8]. También se exploró un nuevo enfoque distribuido basado en TOPSIS para la toma de decisiones en contextos de Big Data, proporcionando una solución efectiva para problemas multicriterio [3]. Así mismo, se propone la incorporación de modelos predictivos utilizando Big Data [6],[12],[16] para reducir el riesgo en la toma de decisiones.

Otro estudio presenta un sistema de clasificación automática en 3D para optimizar procesos logísticos [16]. Además, se examina el impacto de tecnologías Big Data como Hadoop, Spark y Flink en la gestión de datos clínicos

[17]. También se analizó la implementación de un sistema experto basado en lógica difusa para mejorar la gestión de la cadena de suministro en TI, donde se combinan modelos de inteligencia artificial para optimizar la toma de decisiones y reducir riesgos en la cadena de abastecimiento [18]. Adicionalmente, se definieron requisitos para análisis de Big Data y su mapeo con herramientas del ecosistema Hadoop, resaltando la importancia de elegir herramientas adecuadas según la naturaleza de los datos y el tipo de análisis requerido [20]. Así mismo, se investigó la optimización del rendimiento de los frameworks Hadoop y Spark [23] y el uso del Internet de las Cosas (IoT) en plataformas de gobierno electrónico [24].

En relación con la transformación de procesos, otros estudios proponen el uso de Big Data para alinear las prácticas empresariales a las necesidades objetivas identificadas mediante el análisis de datos. Así, un estudio abordó la implementación de un sistema de gestión de suministro en el que se utilizaba el análisis de datos para moldear los KPI al contexto organizacional durante su desarrollo y uso, siguiendo la lógica del método Investigación de diseño-acción (ADR por sus siglas en inglés), con el fin de obtener más eficiencia y rentabilidad [7]. Así mismo, el uso de Big Data facilita la planificación de procesos complejos. Por ejemplo, su uso permitió optimizar la planificación y programación de mantenimiento y reemplazo de componentes de materiales compuestos en trenes de alta velocidad, mejorando significativamente su eficiencia operativa [4].

RQ3. ¿Qué características presentan las herramientas de gestión de Big Data?

Según los datos recopilados, se identificaron diferentes herramientas de gestión de Big Data, que cuenta con funcionalidades distintas. En la tabla VI, se muestran las herramientas (o su combinación) más utilizadas en los estudios revisadas: Apache Hadoop, Power BI y SQL Server y, en menor medida, otras herramientas del ecosistema Hadoop, como MapReduce, Pig, Hive y Spark.

TABLA VI. HERRAMIENTAS DE GESTIÓN DE DATOS UTILIZADOS

Herramienta (o su combinación)	Cantidad	Porcentaje
Hadoop	13	38%
Power BI	11	32%
SQL Server	9	26%
SQL Server + Power BI	1	3%

De igual manera, en la figura 2 se muestran las principales características de estas herramientas.

- a. **Medición de datos:** Las herramientas de Big Data permiten recopilar, almacenar y procesar grandes volúmenes de información en tiempo real, facilitando la generación de métricas clave para la toma de decisiones estratégicas ([4], [7]).

- b. **Uso de dashboards:** Power BI y otras plataformas de análisis permiten la creación de tableros interactivos que visualizan información en tiempo real, mejorando la capacidad de los usuarios para interpretar datos complejos ([9], [11]).
- c. **Sistema interactivo:** Herramientas como SQL Server y Spark proporcionan interfaces interactivas que permiten realizar consultas dinámicas y personalizadas para obtener información relevante de grandes volúmenes de datos ([13], [15]).
- d. **Control de información:** Hadoop y Hive permiten gestionar y controlar la calidad de los datos, asegurando su integridad y reduciendo riesgos asociados a la pérdida de información crítica ([5], [10]).
- e. **Procesamiento de datos:** Tecnologías como MapReduce y Spark destacan por su capacidad de procesar datos en paralelo, optimizando los tiempos de respuesta y mejorando la eficiencia operativa de los sistemas de información empresarial ([6], [12]).

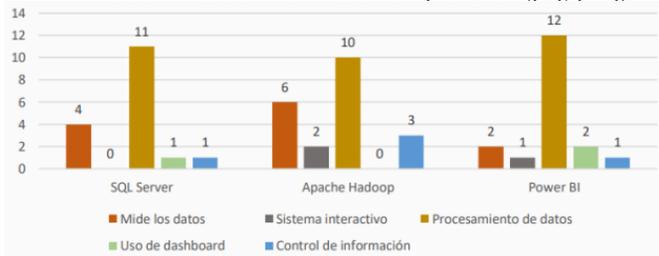


Fig. 2 Características de las herramientas de gestión de datos

Entre las herramientas revisadas, Power BI destaca en la visualización y procesamiento de datos, con estudios que muestran su aplicación en la predicción de demanda de mercado [10] y en marketing de ventas mediante ETL y KPI [1]. Así mismo, se ha demostrado su utilidad en la mejora de la gestión de recursos humanos, en la medida en que su uso para el análisis de datos facilita la gestión y seguimiento de métricas clave de recursos humanos, modelado predictivo y análisis de desempeño laboral [21, 22]. Además, se identificó el uso de BIM y herramientas de BI en la gestión de proyectos de construcción, explorando su integración con metodologías ágiles para mejorar la planificación, control de costos y eficiencia en el uso de materiales [29]. Así mismo, se investigó el uso de Power BI en empresas de servicios digitales en Perú [31] y la combinación de software de monitoreo de comportamiento con BI para mejorar el bienestar animal [26, 32].

Por otra parte, se ha usado SQL Server por su gran capacidad de procesamiento de datos. Tal es el caso de su uso para el análisis de datos de COVID-19 con bases de datos SQL y NoSQL [27]. Otro estudio ha investigado cómo SQL Server gestiona datos oncológicos masivos mediante modelos ETL, optimizando consultas en esquemas de constelación de hechos [12].

En cuanto a la medición de datos, doce estudios destacaban la efectividad de distintas herramientas para esta función. Por ejemplo, en el sector financiero, Apache Hadoop ha sido clave en la detección de fraudes al analizar grandes volúmenes de transacciones en tiempo real ([18]). En manufactura, el uso de Spark ha permitido mejorar la precisión de pronósticos de demanda mediante algoritmos de aprendizaje automático, reduciendo costos y tiempos de producción ([19]). Además, Power BI ha sido implementado en el sector educativo para monitorear el rendimiento de los estudiantes a través de dashboards interactivos que integran métricas de aprendizaje y asistencia ([20]). Estas aplicaciones demuestran cómo la medición de datos con herramientas de Big Data mejora la eficiencia en distintos sectores.

RQ4. ¿Cuáles son los resultados obtenidos en la toma de decisiones usando las herramientas Big Data?

En la tabla VII se muestran los principales resultados obtenidos con la aplicación de herramientas de análisis de Big Data en los sistemas de información empresarial.

TABLA VII. PRINCIPALES RESULTADOS OBTENIDOS

N°	Resultados	Cantidad estudios
1	Mejora operativa	11
2	Mejora de la gestión	7
3	Mejora del análisis de los datos	7
4	Mejora en la toma de decisiones	4
5	Mejora en la predicción de resultados	4
6	Mejora de la seguridad y protección contra ataques	3
7	Optimización de uso de recursos	3
8	Mejora de la integración de datos	2

Como resultado general de los artículos recopilados, se concluye que la mayor contribución del uso de estas tecnologías es la mejora operativa en la gestión de sistemas empresariales. Un artículo destaca cómo la optimización de procesos industriales en el sector textil contribuye a la eficiencia operativa mediante el diseño de modelos de producción mejorados [13]. De manera similar, en el sector logístico, el uso de Big Data ha permitido reducir tiempos de entrega y mejorar la gestión de inventarios mediante el análisis predictivo de la demanda y la optimización de rutas de distribución ([17]). Estas mejoras han resultado en una reducción del 20 % en costos operativos y un incremento del 15 % en la satisfacción del cliente al reducir retrasos en la cadena de suministro ([19]). Estos hallazgos refuerzan la importancia del análisis de datos en la toma de decisiones estratégicas y operativas, mejorando la eficiencia y competitividad de las organizaciones en diversos sectores.

También fue frecuente la mención a la mejora en los procesos de gestión y en el análisis de datos. Así, se encontró un artículo que explora la aplicación de BIM y BI en la gestión de proyectos de construcción [28]. También se exploró la aplicación de BIM y BI en la gestión de proyectos de construcción [29] y otro que propone un sistema de visualización de reportes para la gestión de incidentes por desastres en Sri Lanka [30]. Un caso adicional de mejora en el análisis de datos se observó en el sector financiero, donde el uso de Big Data permitió desarrollar modelos de detección de fraudes en tiempo real, identificando transacciones sospechosas con un 98 % de precisión mediante algoritmos de machine learning ([32]). Este tipo de análisis ha sido crucial para reducir pérdidas económicas y mejorar la seguridad en las operaciones bancarias. Estos hallazgos refuerzan la importancia del análisis de datos en la toma de decisiones estratégicas y operativas, mejorando la eficiencia y competitividad de las organizaciones en diversos sectores.

Respecto a resultados en gestión de la seguridad, un estudio describe la efectividad del análisis de datos mediante Apache Hadoop para mejorar el monitoreo de seguridad en correos electrónicos [11]. También se analizó mediante Big Data las condiciones de salud y el desempeño laboral de empleados mayores en empresas estatales [33]. Con respecto a la integración de datos, se evidenció que el uso de big data proveniente de Internet de las Cosas (IoT) puede mejorar la gestión administrativa en plataformas de gobierno electrónico, permitiendo una mejor integración de datos y optimización de procesos en la administración pública a través del análisis de grandes volúmenes de información en tiempo real [25].

Finalmente, se evaluó el rendimiento de las herramientas de análisis de Big Data para desarrollar aprendizaje automático sobre grandes volúmenes de datos y así mejorar el diagnóstico y predicción de resultados, destacando la escalabilidad y eficiencia de cada una [19]. Así, se examinó el impacto del Big Data en la predicción de condiciones de salud y desempeño laboral de empleados mayores en empresas estatales, revelando que el uso de herramientas analíticas permite identificar factores clave en la predicción de productividad y bienestar de estos trabajadores, además de facilitar la toma de decisiones en políticas de salud ocupacional y estrategias de retención de talento [34]. En otro caso, se exploró el uso de un enfoque híbrido basado en aprendizaje profundo para mejorar la detección temprana de enfermedades cardíacas, destacando su precisión y aplicabilidad en entornos clínicos [5].

IV. DISCUSIÓN

De lo revisado, se puede deducir que las herramientas de análisis de Big Data tienen una mayor presencia e implantación en los sistemas de gestión de datos y de apoyo a la toma de decisiones. Esto indica que las organizaciones han reconocido la importancia de la gestión eficiente de datos para la optimización de procesos y la generación de estrategias basadas en información confiable.

De igual manera, entre las herramientas más destacadas con el fin de aportar mejor a la toma de decisiones fueron Power BI, Apache Hadoop y SQL Server teniendo un resultado óptimo a la hora de generar pronósticos de demanda de mercados y para el procesamiento de datos. Entre las herramientas de gestión de datos las que tuvieron mayor relevancia destaca Apache Hadoop, que ofrece como beneficio frente a otras herramientas el poder reducir la complejidad de trabajo hasta un 30% y SQL Server que muestra un alto nivel de eficacia en la detección de errores en los datos al trabajar con bases de datos de gran volumen. Los resultados obtenidos en este estudio evidencian que la adopción de herramientas de Big Data en los sistemas de información empresarial no solo facilita la **automatización y optimización de procesos**, sino que también permite la implementación de modelos predictivos más precisos, impactando directamente en la reducción de costos y la mejora en la eficiencia operativa. Además, la creciente demanda de análisis en tiempo real ha impulsado el uso de tecnologías de visualización como Power BI, lo que refleja la importancia de contar con plataformas intuitivas y accesibles para la interpretación de datos complejos.

V. CONCLUSIONES

Este estudio realizó una revisión sistemática de la literatura para analizar el impacto de las herramientas de Big Data en los sistemas de información empresarial, con un enfoque particular en la mejora de la toma de decisiones. Los resultados indican que las principales herramientas utilizadas (Apache Hadoop, Power BI y SQL Server) no solo mejoran la eficiencia operativa y reducen costos, sino que también facilitan la toma de decisiones basada en datos, lo que es crucial en un entorno empresarial cada vez más competitivo y globalizado.

En particular, se observó que Power BI es altamente efectivo en la visualización y análisis de datos, mientras que Apache Hadoop destaca por su capacidad para manejar grandes volúmenes de datos y reducir la complejidad en el procesamiento. Por su parte, SQL Server demostró ser una herramienta robusta para la gestión de bases de datos y la detección de errores en entornos de alta demanda. Además, se identificó que las herramientas de Big Data son especialmente útiles en sectores como la salud, la manufactura y los servicios de TI, donde la capacidad de procesar y analizar grandes cantidades de datos en tiempo real es fundamental para la toma de decisiones estratégicas.

En conclusión, la integración de herramientas de Big Data en los sistemas de información empresarial no solo optimiza los procesos internos, sino que también contribuye al crecimiento sostenible de las organizaciones al permitir una toma de decisiones más informada y eficiente. Futuras investigaciones deberían centrarse en explorar el impacto de estas herramientas en otros sectores y en desarrollar metodologías que permitan una implementación más efectiva y segura en entornos empresariales complejos.

REFERENCIAS

- [1] C. T. Gonçalves, M. J. A. Gonçalves, y M. I. Campante, "Developing integrated performance dashboards visualisations using power BI as a platform", *Information (Basel)*, vol. 14, núm. 11, p. 614, 2023.
- [2] A. H. Khan, A. Mustafa, A. Yusuf, y R. Khan, "Software abstractions for large-scale deep learning models in big data analytics", *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, núm. 4, 2019.
- [3] L. Lamrini, M. C. Abounaima, y M. Talibi Alaoui, "New distributed-topsis approach for multi-criteria decision-making problems in a big data context", *J. Big Data*, vol. 10, núm. 1, 2023.
- [4] Y. Jie, L. Hao, S. Yan, y X. Zhang, "Optimization method of high-speed train composite material workshop planning and scheduling", *Adv. Mater. Sci. Eng.*, vol. 2022, pp. 1–11, 2022.
- [5] G. M. Rao, D. Ramesh, V. Sharma, A. Sinha, M. M. Hassan, y A. H. Gandomi, "AttGRU-HMSI: enhancing heart disease diagnosis using hybrid deep learning approach", *Sci. Rep.*, vol. 14, núm. 1, 2024.
- [6] Assistant Professor in Department of Computer Science, Shaheed Rajguru College of Applied Sciences for women (University of Delhi). *et al.*, "Identifying requirements for Big Data Analytics and mapping to Hadoop tools", *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, núm. 3, pp. 4384–4392, 2019.
- [7] D. H. Nabil, M. H. Rahman, A. H. Chowdhury, y B. C. Menezes, "Managing supply chain performance using a real time Microsoft Power BI dashboard by action design research (ADR) method", *Cogent Eng.*, vol. 10, núm. 2, 2023.
- [8] H. Moghaddasi, R. Rahimi, A. Kazemi, K. Arjmandi Rafsanjani, G. Bahoush, y F. Rahimi, "A clinical decision support system for increasing compliance with protocols in chemotherapy of children with acute lymphoblastic leukemia", *Cancer Inform.*, vol. 21, p. 117693512210848, 2022.
- [9] Y. Zhao y W. Du, "Construction and optimization of distributed electronic archives system merging control messages", *Math. Probl. Eng.*, vol. 2022, pp. 1–12, 2022.
- [10] N. Wikamulia y S. M. Isa, "Predictive business intelligence dashboard for food and beverage business", *Bull. Electr. Eng. Inform.*, vol. 12, núm. 5, pp. 3016–3026, 2023.
- [11] Y. Liu, Y. Jin, y Q. Liu, "Big data security and international settlement system of electronic economy based on blockchain", *Mob. Inf. Syst.*, vol. 2021, pp. 1–11, 2021.
- [12] C. E. Atay y G. Garani, "Building a lung and ovarian cancer data warehouse", *Healthc. Inform. Res.*, vol. 26, núm. 4, pp. 303–310, 2020.
- [13] G. Araque González, A. Suárez Hernández, M. Gómez Vásquez, J. Vélez Uribe, y A. Bernal Avellaneda, "Sustainable manufacturing in the fourth industrial revolution: A big data application proposal in the textile industry", *J. Ind. Eng. Manag.*, vol. 15, núm. 4, p. 614, 2022.
- [14] A. Valdez, G. Cortes, S. Castaneda, L. Vazquez, A. Zarate, Y. Salas, y G. H. Atondo, "Big data strategy", *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 4, pp. 285–290, 2019.
- [15] J. Lee, B. Kim, y J.-M. Chung, "Time Estimation and Resource Minimization Scheme for Apache Spark and Hadoop Big Data Systems with Failures", *IEEE Access*, vol. 7, pp. 9658–9666, 2019.
- [16] X.-D. Zhang y Y.-M. Cheng, "Automatic three-dimensional sorting system based on internet and database", *Int. J. Reason.-Based Intell. Syst.*, vol. 11, no. 1, pp. 90–100, 2019.
- [17] E. Nazari, M. H. Shahriari, y H. Tabesh, "Big Data Analysis in Healthcare: Apache Hadoop, Apache Spark and Apache Flink", *Front. Health Inform.*, vol. 8, 2019.
- [18] S. Shokouhyar, S. Seifhashemi, H. Siadat, y M. M. Ahmadi, "Implementing a fuzzy expert system for ensuring information technology supply chain", *Expert Syst.*, vol. 36, no. 1, 2019.
- [19] J. Santosh Kumar, B. K. Raghavendra, S. Raghavendra, y Meenakshi, "Performance evaluation of Map-reduce jar pig hive and spark with machine learning using big data", *Int. J. Electr. Comput. Eng.*, vol. 10, no. 4, pp. 3811–3818, 2020.
- [20] U. Bharti, D. Bajaj, A. Goel, y S. C. Gupta, "Identifying requirements for big data analytics and mapping to hadoop tools", *Int. J. Recent Technol. Eng.*, vol. 8, no. 3, pp. 4384–4392, 2019.
- [21] S. Han et al., "A health management system for large vertical mill", *Int. J. Distrib. Sensor Netw.*, vol. 16, no. 3, 2020.
- [22] S. H. M. S. Andrade, "Data Analytics to Increase Performance in the Human Resources Area", *J. Auton. Intell.*, vol. 4, no. 1, pp. 59–64, 2021.
- [23] Zh. D. Mamykova, M. Bolatkhan, O. L. Kopnova, M. R. Zubairova, y Sh. Zh. Rabat, "Development of the information and analytical system of the university", *KazNU Bull. Math. Mech. Comput. Sci. Ser.*, vol. 112, no. 4, pp. 148–161, 2021.
- [24] H. Atsatryan, A. Kocharyan, D. Hagimont, y A. Lalayan, "Performance optimization system for Hadoop and Spark frameworks", *Cybern. Inf. Technol.*, vol. 20, no. 6.
- [25] M. Qi y J. Wang, "Using the Internet of Things E-Government Platform to Optimize the Administrative Management Mode", *Wirel. Commun. Mob. Comput.*, vol. 2021, 2021.
- [26] E. Lee, H. Oh, y D. Park, "Big Data Processing on Single Board Computer Clusters: Exploring Challenges and Possibilities", *IEEE Access*, vol. 9, pp. 142551–142565, 2021.
- [27] J. D. Wark, "Power Up: Combining Behavior Monitoring Software with Business Intelligence Tools to Enhance Proactive Animal Welfare Reporting", *Animals*, vol. 12, no. 13, 2022.
- [28] J. Antas, R. R. Silva, y J. Bernardino, "Assessment of SQL and NoSQL Systems to Store and Mine COVID-19 Data", *Computers*, vol. 11, no. 2, 2022.
- [29] F. Rodrigues, A. D. Alves, y R. Matos, "Construction Management Supported by BIM and a Business Intelligence Tool", *Energies*, vol. 15, no. 9, 2022.
- [30] X. Liu et al., "Surface roughness prediction method of titanium alloy milling based on CDH platform", *Int. J. Adv. Manuf. Technol.*, vol. 119, no. 11–12, pp. 7145–7157, 2022.
- [31] H. I. Tillekaratne et al., "Situation report (SITREP) visualization for effective management of disaster incidents in Sri Lanka", *J. Infrastruct. Policy Dev.*, vol. 7, no. 3, 2023.
- [32] G. M. M. Castro, H. G. W. Aitken, y A. A. Calvanapon, "Business Intelligence Tools for a Digital Services Company in Peru, 2022", *Int. J. Bus. Intell. Res.*, vol. 14, no. 1, 2023.
- [33] A. Fultz, R. Lewis, L. Kelly, y J. Garbarino, "Behavioral Welfare Research for the Management of Sanctuary Chimpanzees (*Pan troglodytes*)", *Animals*, vol. 13, no. 16, 2023.
- [34] Q. Zhang, "Big health data for elderly employees job performance of SOEs: visionary and enticing challenges", *Multimed. Tools Appl.*, vol. 83, no. 2, pp. 4409–4442, 2024.