

Big Data Analysis for Raw Material Prediction in the Industrial Sector: A Systematic Review

Ethel del Rosario Noe Contreras¹ , Aimar Valderrama Cassani² , Álvaro Mauricio Blancas Castro³ , José Miguel Cornelio Ramos⁴ 

^{1,2,3,4}Universidad Tecnológica del Perú, Perú, U21229613@utp.edu.pe, U21203758@utp.edu.pe, alvaroblancas01@gmail.com, jcornelio@utp.edu.pe

Abstract - In the context of the digital transformation of the industrial sector, this review aimed to analyze the use of Big Data in raw material prediction, with an emphasis on its application in purchase planning and inventory management. A systematic search was conducted in the Scopus database during April 2025, applying the PRISMA protocol and the PIOC strategy. Method: 829 articles were identified, of which 23 met the inclusion criteria related to the implementation of predictive techniques in manufacturing sector organizations. Results: Substantial improvements were observed in forecast accuracy, operational cost reduction, and inventory optimization, especially through the use of models such as LSTM neural networks, hybrid algorithms, and digital twins. These technologies also contributed to more informed and adaptable decision-making in contexts of high variability. Conclusions: The review provides relevant evidence on the benefits of Big Data analysis in the supply chain, although limitations are recognized due to the scarce presence of studies in Latin America and the reliance on research focused on Asia and Europe. It is recommended to conduct empirical studies in local settings in order to validate the applicability of these tools in contexts with lower digital maturity.

Keywords-- Big Data, Inventory management, Demand forecasting, LSTM neural networks, Supply chain

Análisis de Big Data para la predicción de materia prima en el sector industrial: Una revisión sistemática

Ethel del Rosario Noe Contreras¹, Aimar Valderrama Cassani², Álvaro Mauricio Blancas Castro³, José Miguel Cornelio Ramos⁴

^{1,2,3,4}Universidad Tecnológica del Perú, Perú, U21229613@utp.edu.pe, U21203758@utp.edu.pe, alvaroblancas01@gmail.com, jcornelio@utp.edu.pe

Resumen- En el contexto de transformación digital del sector industrial, esta revisión tuvo como objetivo analizar el uso de Big Data en la predicción de materia prima, con énfasis en su aplicación en la planificación de compras y la gestión de inventarios. Se realizó una búsqueda sistemática en la base de datos Scopus durante abril de 2025, aplicando el protocolo PRISMA y la estrategia PIOC. Método: Se identificaron 829 artículos, de los cuales 23 cumplieron con los criterios de inclusión relacionados con la implementación de técnicas predictivas en organizaciones del sector manufacturero. Resultados: Se observaron mejoras sustanciales en la precisión de los pronósticos, la reducción de costos operativos y la optimización del inventario, especialmente mediante el uso de modelos como redes neuronales LSTM, algoritmos híbridos y gemelos digitales. Estas tecnologías también contribuyeron a una toma de decisiones más informada y adaptable frente a contextos de alta variabilidad. Conclusiones: La revisión aporta evidencia relevante sobre los beneficios del análisis de Big Data en la cadena de suministro, aunque se reconocen limitaciones asociadas a la escasa presencia de estudios en Latinoamérica y la dependencia de investigaciones centradas en Asia y Europa. Se recomienda realizar estudios empíricos en entornos locales, a fin de validar la aplicabilidad de estas herramientas en contextos con menor madurez digital.

Palabras clave-- Big Data, Gestión de inventarios, Predicción de demanda, Redes neuronales LSTM, Cadena de suministro

I. INTRODUCCIÓN

Actualmente, el proceso de transformación digital ha repercutido significativamente en la manera en que las organizaciones gestionan sus procesos operativos. La adopción de tecnologías emergentes tiene como propósito principal mejorar la efectividad en las operaciones y fortalecer la capacidad de adaptación frente a las demandas del entorno dinámico. En ese contexto, el análisis de datos se posiciona como un recurso estratégico fundamental, ya que su aplicación facilita la identificación anticipada de tendencias, optimiza los tiempos de respuesta y mejora la fiabilidad de los procesos en el ámbito industrial [1]. De igual manera, estudios recientes destacan que el uso de Big Data no solo potencia la eficiencia operativa, sino que también habilita un seguimiento de los procesos de suministro. Este seguimiento mejora la toma de decisiones, reduce los márgenes de error y permite actuar con mayor agilidad ante cambios o contingencias. Estos beneficios hacen que la gestión sea más inteligente, sostenible y alineada con las nuevas demandas del mercado [2]. Tiene un impacto directo en la continuidad de la producción, en los costos operativos y en la habilidad para adaptarse ante variaciones del entorno comercial. En ese marco, la incorporación de

herramientas de Big Data permite anticipar necesidades de abastecimiento, prevenir interrupciones y generar ventajas competitivas sostenibles [3].

Big Data hace referencia a colecciones de datos caracterizados por su volumen, velocidad y variedad, los cuales requieren tecnologías avanzadas para ser procesados y analizados adecuadamente. Su correcta aplicación permite descubrir patrones ocultos, anticipar comportamientos y respaldar decisiones estratégicas en contextos dinámicos y complejo [4]. Sin embargo, aún persiste un problema importante en el sector industrial: muchas empresas continúan utilizando métodos limitados o poco desarrollados para gestionar el inventario de materias primas. Esto genera problemáticas como paros inesperados en la producción, falta de insumos y aumento en los costos operativos. Estas situaciones afectan la eficiencia general, reducen la rentabilidad e impactan negativamente en la capacidad de adaptación del sistema productivo [5].

Asimismo, se ha identificado que la ausencia de tecnologías predictivas vuelve a las cadenas de suministro más vulnerables, especialmente ante escenarios de incertidumbre o 7 cambios abruptos en la demanda [6]. Ante este panorama, el uso de tecnologías basadas en Big Data surge como una solución estratégica. Su aplicación permite transformar modelos tradicionales en sistemas automatizados y predictivos, capaces de anticipar escenarios y minimizar riesgo [7]. Además, el estudio de macrodatos y la IA permiten supervisar, gestionar, anticipar en tiempo real distintas operaciones industriales. Fortalecen la toma de decisiones estratégicas y permiten optimizar costos de manera significativa [8]. Sin embargo, el ritmo acelerado de la tecnología no siempre va acompañado de una aplicación eficiente. Diversas organizaciones se enfrentan a obstáculos técnicos, económicos o formativos que limitan una integración completa.

Esta investigación se justifica por la necesidad urgente de consolidar el conocimiento actual sobre el uso predictivo del Big Data en el entorno industrial. Dado que muchas organizaciones aún operan con modelos ineficientes, una revisión sistemática permitirá identificar enfoques efectivos, aportar valor práctico al sector y enriquecer el cuerpo teórico disponible en esta área.

En función del objetivo planteado, esta revisión sistemática ha sido estructurada en cinco secciones principales. La primera presenta el contexto general y la justificación del estudio. Luego, se realiza la metodología siguiendo PRISMA y la estrategia PICOC, incluyendo los criterios de búsqueda,

inclusión y exclusión de estudios. En la tercera sección se exponen los resultados obtenidos, desglosados por indicadores bibliométricos y categorías temáticas. A continuación, la sección de discusión analiza de manera crítica los hallazgos más relevantes, estableciendo conexiones con el estado actual de la literatura. Finalmente, se presentan las conclusiones generales y recomendaciones para futuras investigaciones en el ámbito industrial.

II. METODOLOGÍA

2.1. Estrategia de búsqueda

Se realizó un análisis sistemático de fuentes bibliográficas (SRL) enfocada en examinar el empleo de técnicas predictivas aplicadas a la demanda de insumos en el sector de producción o manufactura. El objetivo principal fue recopilar, evaluar y sintetizar los hallazgos reportados en diversas investigaciones relacionadas con esta temática. Para estructurar el proceso metodológico, se utilizó el protocolo PRISMA, complementado con la estrategia PIOC, que considera los elementos de Problema, Intervención, Comparación, Resultados y Contexto, los cuales se detallan en la Tabla I.

El propósito de esta Revisión Sistemática es reconocer, examinar y resumir las metodologías de Big Data utilizadas para predecir materias primas en contextos industriales, adoptando un enfoque principalmente exploratorio y descriptivo. Por ello, no se considera la comparación entre diferentes intervenciones, como sucede en investigaciones experimentales o análisis cuantitativos, sino que se busca comprender su alcance, resultados e impactos observados.

TABLA I DESCRIPCIÓN DE LOS COMPONENTES PIOC Y PALABRAS CLAVE				
			Palabras clave	Keywords
P	Problema	Carencia de soluciones logísticas que impacta de manera negativa en el rendimiento de la cadena de suministros.	Problemas logísticos, cadena de suministro, pronóstico de demanda, costos logísticos, gestión de suministros, ineficiencia operativa, interrupciones, retrasos	Logistics issues, supply chain, demand forecasting, logistics costs, supply management, operational inefficiency, disruptions
I	Intervención	Implementación de tecnologías de análisis de datos	Análisis predictivo, Big Data, machine Learning, IoT, Inventarios con IA, Minería de datos, abastecimiento industrial, Industria 4.0, Business	Predictive analysis, Big Data, Machine Learning, IoT, AI Inventories, Data Mining, Industrial Supply, Industry 4.0, Business Intelligence, Purchasing Models, Unstructured

			Intelligence, modelos de compras, datos no estructurados, volumen de datos	Data, Data Volume
C	Comparación	-	-	-
O	Resultados	Predicción precisa de materia prima, optimización del tiempo de abastecimiento, reducción de costos operativos, eficiencia en la rotación de inventarios y nivel de inventario óptimo	Eficiencia, continuidad productiva, errores de modelos predictivos, automatización, optimización de inventarios, Big Data, reducción de costos, proyección, compras, materia prima, abastecimiento, ERI, ERU, VMI, SCM, WMS, SKU	Efficiency, production continuity, predictive model errors, automation, inventory optimization, Big Data, cost reduction, projection, purchasing, raw materials, supply, ERI, ERU, VMI, SCM, WMS, SKU
C	Contexto	Empresas de manufactura y producción	Calidad de datos, sector industrial, barreras tecnológicas, transformación digital industrial, capacidad de almacenamiento, almacén, ruptura de stock, gestión de almacenes, stock de seguridad	Data quality, industrial sector, technological barriers, industrial digital transformation, stock break, warehouse management, safety stock

^aEn este estudio no se realizaron comparaciones

A partir de los elementos definidos en la estrategia PIOC, se formuló la pregunta de investigación que orientó el desarrollo del enfoque metodológico, la cual fue la siguiente: ¿Qué problemáticas afectan actualmente a las industrias en el ámbito logístico y la cadena de suministro? Del mismo modo, se formularon interrogantes adicionales para profundizar el análisis: ¿Qué tecnologías de Big Data son más efectivas para optimizar la gestión de inventarios?, ¿Qué efectos genera el uso de Big Data en la planificación de materias primas?, y ¿Cómo impacta la aplicación de Big Data en la cadena de suministro del sector manufacturero?

Del mismo modo, la estrategia PIOC permitió definir las palabras clave empleadas en el proceso de búsqueda. Con el objetivo de alcanzar una indagación más amplia y precisa, a estas palabras clave se integraron operadores booleanos como “OR”, “AND” y “NOT”, para estructurar la consulta de

búsqueda de la siguiente forma: (TITLE ABS-KEY ("Logistics issues" OR "logistics problems" OR "supply chain challenges" OR "supply chain" OR "supply network" OR "logistics network" OR "supply system" OR "demand forecasting" OR "demand planning" OR "logistics costs" OR "supply chain costs" OR "supply management" OR "supply chain management" OR "inventory management" OR "procurement management" OR "operational inefficiency" OR "performance inefficiencies" OR "business inefficiencies" OR "disruptions" OR "interruptions" OR "delays") AND TITLE ABS-KEY ("Predictive analysis" OR "predictive analytics" OR "forecasting models" OR "predictive modeling" OR "data-driven predictions" OR "Big Data" OR "large-scale data" OR "massive data" OR "big data analytics" OR "Machine Learning" OR "ML algorithms" OR "learning algorithms" OR "IoT" OR "Internet of Things" OR "connected devices" OR "AI Inventories" OR "Artificial Intelligence in inventory" OR "smart inventory systems" OR "automated stock control" OR "Data Mining" OR "knowledge discovery" OR "data extraction" OR "pattern mining" OR "Industrial Supply" OR "industrial provisioning" OR "manufacturing supply" OR "production supply" OR "Industry 4.0" OR "Fourth Industrial Revolution" OR "smart industry" OR "digital industry" OR "Business Intelligence" OR "BI tools" OR "data-driven decision making" OR "analytics solutions" OR "Purchasing Models" OR "procurement models" OR "acquisition strategies" OR "buying models" OR "Unstructured Data" OR "non-tabular data" OR "unorganized data" OR "raw data" OR "Data Volume" OR "large datasets" OR "big data size" OR "data magnitude") AND TITLE ABS-KEY ("Efficiency" OR "Productivity" OR "Operational efficiency" OR "Production continuity" OR "Operational continuity" OR "Manufacturing continuity" OR "Predictive model errors" OR "Forecasting inaccuracies" OR "Prediction errors" OR "Automation" OR "Process automation" OR "Industrial automation" OR "Inventory optimization" OR "Stock optimization" OR "Inventory management" OR "Limitations" OR "Constraints" OR "Barriers" OR "Challenges" OR "Obstacles" OR "Difficulties" OR "Big Data" OR "Massive data analytics" OR "Large-scale data analysis" OR "Cost reduction" OR "Expense minimization" OR "Operational cost savings" OR "Projection" OR "Forecasting" OR "Trend projection" OR "Purchasing" OR "Procurement" OR "Sourcing" OR "Raw materials" OR "Primary materials" OR "Base materials" OR "Supply" OR "Supply chain" OR "Supply network" OR "ERI" OR "Efficient Replenishment Inventory" OR "Inventory Replenishment Strategy" OR "ERU" OR "Efficient Replenishment Unit" OR "Inventory Replenishment Unit" OR "VMI" OR "Vendor Managed Inventory" OR "Supplier Managed Inventory" OR "Vendor-Controlled Inventory" OR "SCM" OR "Supply Chain Management" OR "Logistics Chain Management" OR "Supply Network Management" OR "WMS" OR "Warehouse Management System" OR "Inventory Control Software" OR "Warehouse Software" OR "SKU" OR "Stock Keeping Unit" OR "Product Code" OR "Inventory Item Code") AND TITLE

ABS-KEY ("Data quality" OR "information quality" OR "data accuracy" OR "data reliability" OR "data consistency" OR "industrial sector" OR "manufacturing sector" OR "production industry" OR "industrial domain" OR "technological barriers" OR "technology limitations" OR "tech challenges" OR "technological constraints" OR "industrial digital transformation" OR "digitalization in industry" OR "smart manufacturing" OR "industrial innovation" OR "digital industry transition" OR "stock break" OR "stockout" OR "inventory shortage" OR "out of stock" OR "warehouse management" OR "inventory management" OR "storage management" OR "warehouse operations" OR "logistics warehouse management" OR "safety stock" OR "buffer stock" OR "reserve inventory" OR "contingency stock")). Se obtuvo un resultado de 829 artículos mediante “SCOPUS”.

La búsqueda se llevó a cabo a través del repositorio SCOPUS en el mes de abril de 2025, con el objetivo de identificar artículos científicos que aportaran de manera significativa al desarrollo de la presente investigación.

2.2. Criterios de inclusión y exclusión

Los artículos seleccionados permitieron establecer los criterios de elegibilidad, los cuales se fundamentaron en los lineamientos de la declaración PRISMA. La Tabla II presenta una descripción de estos criterios.

TABLA II
CRITERIOS DE ELEGIBILIDAD

Inclusión	Exclusión
<ul style="list-style-type: none">• CII: Serán considerados todos los estudios en los que se haya implementado alguna técnica o enfoque para la predicción del abastecimiento o el consumo de materias primas.	<ul style="list-style-type: none">• CE1: Se excluirán todos aquellos artículos cuya fecha de publicación sea anterior al periodo comprendido entre los años 2020 y 2025.
<ul style="list-style-type: none">• CII2: Artículos que mencionen la aplicación de herramientas de Big Data para la aplicación de materia prima o demanda.	<ul style="list-style-type: none">• CE2: Quedarán descartados los estudios que no estén vinculados al ámbito de ingeniería industrial.
<ul style="list-style-type: none">• CII3: Se requiere que los artículos incluyan evidencias claras sobre la implementación de herramientas empleadas para la proyección o previsión de compras.	<ul style="list-style-type: none">• CE3: No se tomarán en cuenta documentos que no correspondan al formato de artículo.
<ul style="list-style-type: none">• CII4: Solo se aceptarán estudios realizados en organizaciones pertenecientes al sector productivo o manufacturero.	<ul style="list-style-type: none">• CE4: Serán excluidos aquellos trabajos en los que no se reporten resultados verificables o evidencias de aplicación.
	<ul style="list-style-type: none">• CE5: No se considerarán artículos redactados en idiomas distintos al español o al inglés

2.3. Proceso de selección de estudios

La Figura 1 muestra la etapa de filtrado y elección de publicaciones mediante el esquema gráfico PRISMA. En primer lugar, se realizó la identificación de registros, donde se descartaron artículos duplicados y aquellos que no contaban con autoría, eliminándose un total de 27 registros. El segundo

filtro de cribado excluyó los artículos que no cumplían con 11 los siguientes criterios: publicaciones fuera del período 2020–2025, investigaciones desarrolladas en sectores distintos al de la ingeniería industrial, documentos que no correspondían a artículos científicos, estudios que no presentaban resultados, y aquellos escritos en lenguas diferentes al español o al inglés. En esta etapa se descartaron 675 artículos. En el tercer proceso de cribado, se excluyeron los artículos que no contaban con acceso libre, en esta fase se eliminó 68 artículos. Finalmente, el cuarto proceso excluyó los artículos que no se ajustaban a las condiciones establecidas, los cuales se basan en la lógica de la declaración PRISMA. En relación con el CI1 (“Serán considerados todos los estudios en los que se haya implementado alguna técnica o enfoque para la predicción del abastecimiento o el consumo de materias primas”), se descartaron 28 artículos. Asimismo, respecto al CI4 (“Solo se aceptarán estudios realizados en organizaciones pertenecientes al sector productivo o manufacturero”), se eliminaron 8 artículos. En total, se descartaron 36 artículos en esta etapa. Como resultado final, se seleccionaron 23 artículos que fueron incluidos para el análisis en esta revisión sistemática.

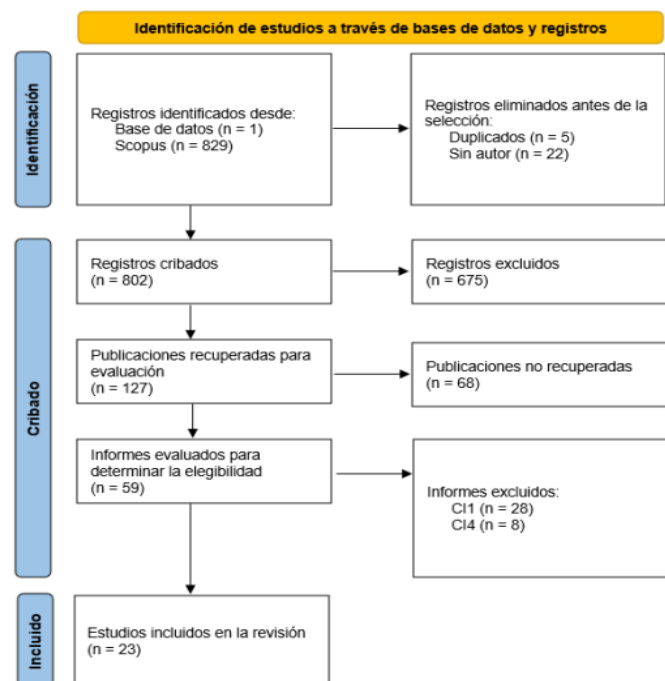


Fig. 1 Diagrama de flujo PRISMA

Nota. Adaptado de <https://www.prisma-statement.org/prisma-2020>

2.4. Artículos seleccionados

Una vez aplicado la estrategia PRISMA, se concluyó lo siguiente, 23 de los 829 artículos inicialmente identificados cumplían con los criterios de elegibilidad establecidos, por lo que fueron seleccionados para la evaluación definitiva.

III. RESULTADOS

Los resultados fueron estructurados en tres apartados. En primer lugar, se detallaron los indicadores cuantitativos que se obtienen de las publicaciones consultadas. Posteriormente, se llevó a cabo un estudio de las principales características de las investigaciones analizadas. Finalmente, se desarrolló un análisis para identificar los asuntos centrales y las líneas prioritarias abordadas en la literatura revisada.

3.1. Características bibliométricas de los artículos seleccionados

La revisión de la cantidad de investigaciones publicadas por año, representado en la Figura 2, evidencia una variabilidad significativa a lo largo del periodo 2020–2025. En el año 2021 se observa un repunte importante con 5 publicaciones, a comparación del año anterior y se observa un aumento sustancial en la actividad investigativa; sin embargo, el año 2024 destaca significativamente con el mayor volumen de publicaciones, alcanzando un total de 8, lo que representa el 35% del total de artículos dentro del periodo analizado. Estas variaciones podrían deberse a factores como el interés fluctuante en la digitalización de procesos productivos y los cambios en la disponibilidad de datos e infraestructura tecnológica para investigación en el ámbito industrial.

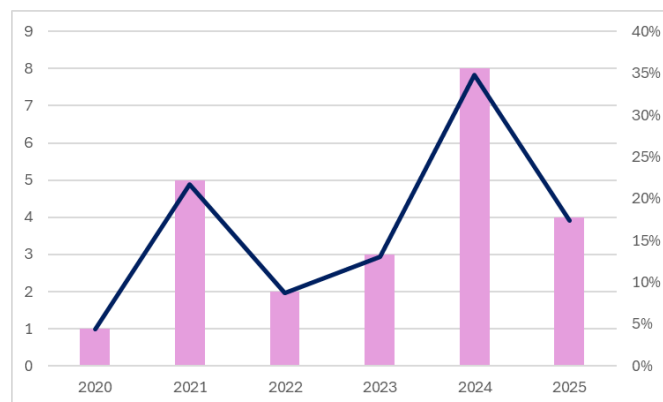


Fig. 2 Número de publicaciones por año

La Figura 3 demuestra que IEEE Access destaca por su mayor número de artículos, lo que evidencia su relevancia en la difusión de investigaciones tecnológicas aplicadas a contextos industriales. Le siguen revistas como International Journal of Production Research y EAI que también albergan estudios orientados a la optimización de Endorsed Transactions on Internet of Things, procesos mediante herramientas digitales. La concentración en estas revistas sugiere una preferencia por medios especializados en innovación, producción y tecnologías emergentes, altamente alineadas con el enfoque de esta investigación.

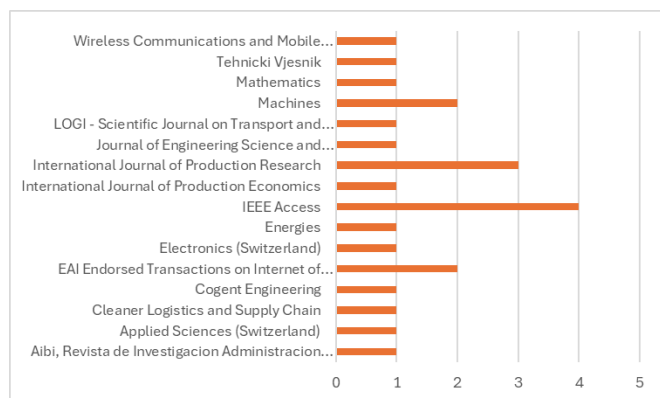


Fig. 3 Número de publicaciones por revista

Tal como se aprecia en la Figura 4, dentro de los 23 artículos seleccionados se identificaron las publicaciones con mayor número de citas dentro del área de investigación. El top 5 elegido está comprendido por la cantidad de 84, 63, 61, 55 y 27. Esta diferencia en el número de citas puede relacionarse tanto con la relevancia del enfoque propuesto como el prestigio de las revistas en las que fueron publicados, se identifica que la tres primeras publicaciones pertenecen a revistas indexadas en el cuartil 1, la investigación de [14] está publicada en una revista del Q2 y la última pertenece al Q3, lo cual refuerza la asociación entre el impacto académico y la calidad percibida del medio de publicación. Los cinco estudios seleccionados presentan enfoques metodológicos consistentes y desarrollo tecnológico aplicados al pronóstico de materias primas mediante Big Data. Su presencia en revistas de alto impacto evidencia que han contribuido significativamente a consolidar las bases teóricas y prácticas del campo, siendo referente clave para investigaciones futuras. Asimismo, su variedad temática que incluye desde redes neuronales hasta simulación avanzada y gemelos digitales refleja de manera integral las tendencias más influyentes en esta línea de investigación.

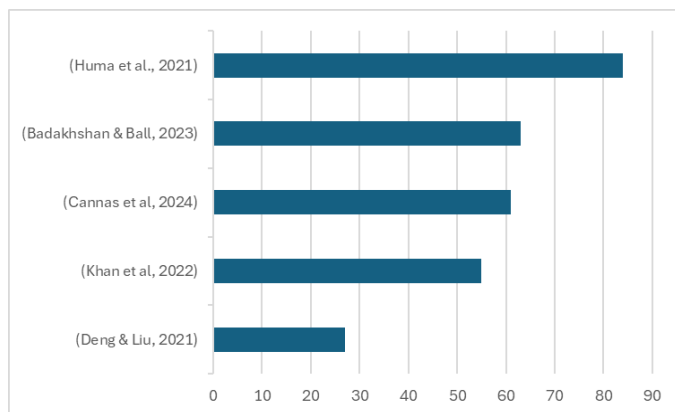


Fig. 4 Top 5 de publicaciones con más citas

La Figura 5 integra datos cuantitativos y geospaciales para mostrar el origen de los artículos revisados, permitiendo visualizar la frecuencia de publicaciones y su dispersión

geográfica. El análisis evidencia un claro predominio de Asia, que concentra el 57 % de los estudios, con China como el principal contribuyente (4 artículos), seguido por India, Bangladesh y Pakistán con dos cada uno. Europa representa el 30 % de los estudios, incluyendo contribuciones de Reino Unido, Italia, Irlanda, Bulgaria y Croacia. Europa aporta el 30 %, destacando países como Reino Unido e Italia, mientras que América (Colombia) y África (Sudáfrica) registran menores participaciones con 9% y 4% respectivamente. Esta concentración asiática se asocia con la digitalización industrial y el interés por la Industria 5.0. La superposición de ambas figuras no solo permite cuantificar, sino también interpretar la distribución del conocimiento científico en torno a la predicción de materias, reflejando patrones regionales relevantes para futuras investigaciones.

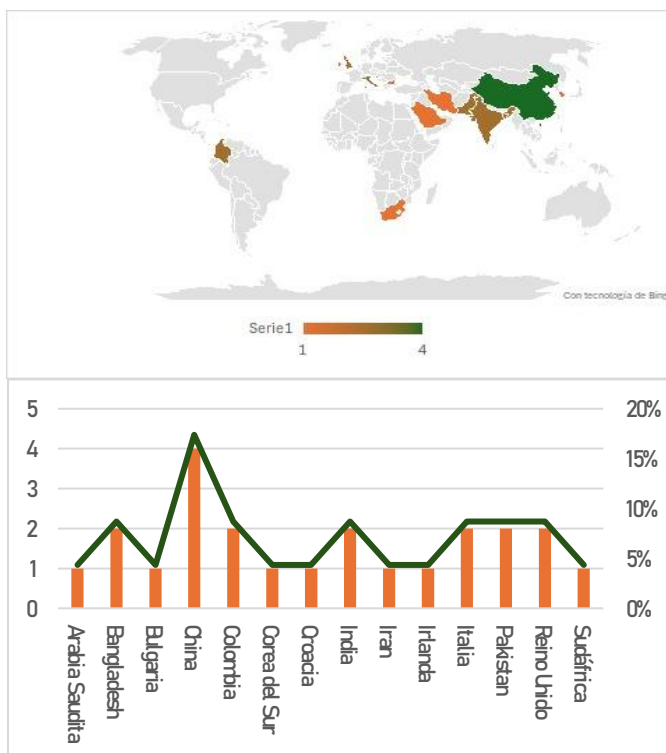


Fig. 5 Número de publicaciones por país

La Figura 6, elaborada mediante el software VOSviewer, presenta una red de coocurrencia basada en 27 palabras clave extraídas de los 23 artículos analizados, identificando dos clústeres temáticos predominantes. El primero se orienta a supply chain management vinculando términos como inventory management, internet of things y decision making, mientras que el segundo se centra en machine learning, asociado con forecasting, deep learning y Big Data analytics. Además, el gradiente de color refleja una evolución terminológica: entre 2020 y 2021 destacan conceptos como Big Data y metadata, mientras que hacia 2023–2024 emergen con mayor fuerza artificial intelligence, quality control y automation. Estos hallazgos evidencian una transición progresiva desde enfoques

exploratorios hacia aplicaciones predictivas y automatizadas, alineadas con los principios de la Industria 4.0 y 5.0.

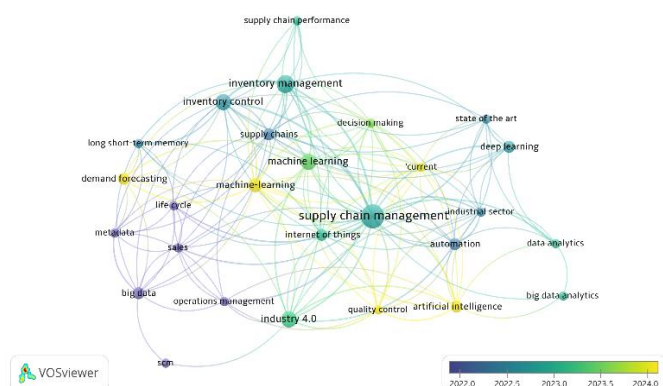


Fig. 6 Red de coocurrencia: Incidencia de palabras clave

IV. RESULTADOS DE INGENIERÍA

En primer lugar, se llevó a cabo la recopilación de datos a partir de los 23 documentos seleccionados para esta RSL, siguiendo los criterios establecidos por las preguntas PIOC.

P: ¿Qué problemáticas afectan actualmente a las industrias en el ámbito logístico y la cadena de suministro?

Las industrias enfrentan desafíos importantes relacionados con la administración logística y de suministros, muchos conectados a los cambios constantes de la demanda. La personalización de productos y la participación del cliente han ampliado los ciclos de suministro y generado mayor incertidumbre en la planificación [4]. Esta situación complica la aplicación de modelos tradicionales de pronóstico, como la suavización exponencial, que resultan insuficientes frente a la variabilidad del mercado actual [22].

Una problemática común es la falta de información clara del proceso y de integración de datos en tiempo real. Muchas empresas aún operan con sistemas que no están conectados entre sí, lo que impide una respuesta ágil y efectiva. Esta desconexión genera errores en inventario, pérdida de trazabilidad y decisiones reactivas [10]. Además, la ausencia de interoperabilidad entre plataformas internas y externas impide una buena coordinación entre las partes del proceso [12], [24]. Como se observa en la Figura 7, estos obstáculos se agrupan dentro de la categoría “Integración e interoperabilidad de sistemas”, la cual representa el 27% de las problemáticas identificadas, y se ubica como el eje con mayor presencia en los estudios revisados, evidenciando su impacto estructural sobre la eficiencia logística.

En términos estratégicos, persiste una tensión entre eficiencia y resiliencia. Las políticas centradas en la

optimización de costos, como el just-in-time, han expuesto a las organizaciones a problemas internacionales, obligándolas a reconsiderar sus modelos logísticos [13]. Esta complejidad se intensifica globalmente, donde factores como regulaciones, conflictos o fenómenos naturales dificultan la coordinación efectiva de cadenas extendidas [23].

Otros desafíos relevantes incluyen el cambio en las expectativas del consumidor, que exige entregas más rápidas y personalizadas, lo cual presiona los sistemas tradicionales [21]; la desarticulación entre áreas funcionales, que genera redundancia y baja adaptabilidad [3]; y la dificultad para implementar prácticas sostenibles, por la carencia de mecanismos que permitan evaluar las consecuencias ambientales derivadas de las decisiones logísticas [20].



Fig. 7 Distribución de problemáticas en logística y cadena de suministro industrial

I: ¿Qué tecnologías de Big Data son más efectivas para optimizar la gestión de inventarios?

El uso de Big Data combinado con algoritmos de aprendizaje automático ha demostrado buenos resultados para optimizar la gestión de inventarios. Tecnologías como Python, SQL y R permiten construir modelos predictivos sobre datos heterogéneos [9], mientras que herramientas de inteligencia empresarial como Power BI facilitan la toma de decisiones colaborativas mediante dashboards que integran múltiples fuentes de información [20].

Entre las técnicas más efectivas se encuentran las redes neuronales LSTM, valoradas por su capacidad para predecir series temporales con alta variabilidad, superando a modelos como ARIMA o Holt-Winters en contextos complejos [4], [17]. Asimismo, el uso de algoritmos de clustering como K-means y modelos multivariados de machine learning permite segmentar productos y anticipar puntos de reorden más precisos, mejorando el control del inventario [11], [22].

El desarrollo de gemelos digitales y simulaciones por eventos discretos ha facilitado la evaluación de escenarios inusuales, contribuyendo a una planificación más adaptativa [13], [1]. También se han propuesto enfoques avanzados como algoritmos de optimización por enjambre de partículas (PSO) y aprendizaje por refuerzo (DRL), los cuales permiten mejorar la planificación ante múltiples variables o entornos cambiantes [16], [3]. Además, el uso de datos generados por inteligencia artificial ha surgido como alternativa en contextos con información limitada o sensible [19], [14].

Finalmente, tecnologías como data lakes, edge computing e infraestructuras en la nube han permitido un procesamiento con mayor rendimiento de datos operativos [12], [24]. La integración con sensores IoT, RFID, blockchain y protocolos como MQTT ha mejorado la visibilidad y el control del inventario en tiempo real [21], [24]. Asimismo, plataformas como Hadoop, Spark y TensorFlow, junto con redes como GCN y LSTM, han fortalecido la colaboración entre áreas y mejorado la precisión en la planificación logística [23], [26]. Asimismo, plataformas como Hadoop, Spark y TensorFlow, junto con redes como GCN y LSTM, han fortalecido la colaboración entre áreas y mejorado la precisión en la planificación logística [23], [26]. La Figura 8, evidencia que las estrategias de análisis y modelado predictivo representan el 38% de los enfoques identificados, consolidándose como la tendencia principal, mientras que los sistemas inteligentes y de conectividad avanzada como IoT, blockchain y computación en la nube concentran el 26%, mostrando su importancia en la gestión de inventarios más dinámicos y eficientes.



Fig. 8 Panorama de Big Data en la mejora de inventarios

O: ¿Qué efectos generó el uso de Big Data en la planificación de materias primas?

El uso de Big Data en la planificación de materias primas ha generado mejoras significativas, especialmente en la precisión de los pronósticos de demanda. Modelos como LSTM han logrado altos niveles de exactitud, lo que permite reducir tanto los excesos como los faltantes de inventario, optimizando costos y mejorando la eficiencia operativa [4], [17], [3]. Además, técnicas como el clustering o el uso de machine learning para demandas intermitentes han facilitado una planificación más flexible, adaptada a productos con comportamientos de consumo más variables [11], [22]. Se puede evidenciar en la Figura 9, que el impacto más reportado es la mejora de la precisión y la eficiencia, representando el 39% de los efectos identificados.

Otro efecto importante ha sido la capacidad para anticipar interrupciones y responder firmemente ante los cambios. El análisis predictivo con Big Data permite detectar fallas antes de que ocurran y simular distintos escenarios para optimizar decisiones estratégicas y financieras [12], [13], [1]. Asimismo, su integración con modelos avanzados como DRL y algoritmos de optimización ha permitido afrontar contextos poco predecibles y variaciones de consumo durante el año. [16], [3].

También se han reportado beneficios en la coordinación entre áreas y actores de la cadena de suministro. Big Data ha facilitado la toma de decisiones colaborativas, mejorado la sincronización entre proveedores y fabricantes, y contribuido a

una mayor transparencia operativa [15], [21]. En el plano tecnológico, herramientas como dashboards y sensores IoT han fortalecido el seguimiento en tiempo real, elevando los niveles de eficiencia en sectores como el textil y de consumo masivo [20], [24].

Sin embargo, se ha advertido que la calidad de los datos y avance en el uso de tecnologías son factores determinantes para que estos beneficios se materialicen. En contextos con baja capacidad tecnológica, como el caso de Bulgaria, los efectos positivos son más limitados [23], [25]. Como respuesta, se proponen soluciones como el uso de datos sintéticos generados por inteligencia artificial para mejorar la calidad del entrenamiento de los modelos en entornos con escasez de datos [19]. En conjunto, Big Data ha permitido a las industrias una planificación más precisa, dinámica y eficiente de sus materias primas [14], [5]. Los resultados reflejan que el principal aporte de Big Data en la gestión de abastecimiento no solo se concentra en la implementación tecnológica, sino en su impacto directo sobre la eficiencia operativa y la precisión en la planificación.

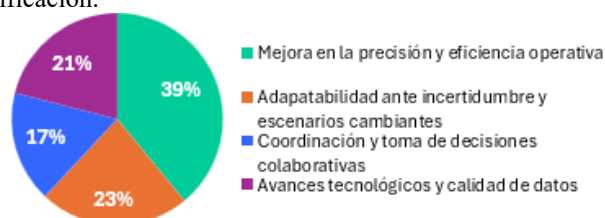


Fig. 9 Impacto de Big Data en la gestión de abastecimiento

C: ¿Cómo afecta la aplicación de Big Data en la cadena de suministros para el sector de manufactura?

La adopción de tecnologías Big Data en el sector manufacturero ha mejorado significativamente la agilidad y capacidad de adaptación de las redes de suministro. El análisis en tiempo real de datos históricos permite pronosticar la demanda con mayor precisión, optimizar inventarios y reducir tiempos de entrega, lo que favorece una toma de decisiones más eficiente y un mejor servicio al cliente [4], [11], [12].

Tecnologías como los gemelos digitales y la simulación por eventos discretos han permitido a las organizaciones analizar escenarios difíciles y prever posibles interrupciones en los procesos. Estas herramientas, al integrarse con modelos de optimización, fortalecen la adaptabilidad operativa frente a entornos inciertos [13], [1]. Asimismo, el uso de IoT, blockchain y plataformas en la nube ha promovido una mayor trazabilidad e integración en la cadena, facilitando la colaboración entre distintos actores [15], [5], [21].

No obstante, la adopción de estas tecnologías enfrenta limitaciones. Las empresas con baja madurez digital suelen no contar con una infraestructura adecuada, personal especializado y sistemas interoperables. Esta situación impide el uso óptimo de las herramientas analíticas basadas en datos, como lo muestran casos como el de Bulgaria, donde la fragmentación tecnológica obstaculiza la integración efectiva [14], [6].

Finalmente, Big Data permite superar los enfoques tradicionales de planificación, especialmente en cadenas con demanda irregular. Modelos como algoritmos PSO ofrecen predicciones más adaptadas a la realidad operativa [16], [3]. A nivel estratégico, la integración de datos favorece un entorno colaborativo alineado con los principios de la Industria 4.0 y 5.0 [23], [25], [26]. Sin embargo, su éxito depende también de factores organizacionales, como la calidad de los datos y la disposición al cambio cultural [18], [24]. La Figura 10 muestra que las transformaciones orientadas a la digitalización industrial representan el 27 % de las contribuciones identificadas, posicionándose como el eje más recurrente de los estudios seleccionados. Además, el 21 % de las investigaciones se centran en entornos con alta complejidad y volumen de datos, donde Big Data resulta esencial para optimizar la escalabilidad, resiliencia y eficiencia de los procesos. Estos hallazgos evidencian que el aprovechamiento de estas tecnologías depende no solo de su adopción, sino también del nivel de madurez organizacional y de la capacidad para gestionar entornos dinámicos e interconectados.

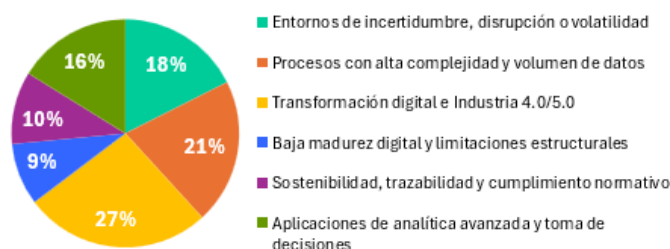


Fig. 10 Contribuciones y desafíos del Big Data en supply chain de manufactura

IV. DISCUSION

En esta revisión sistemática, se identificaron cambios relevantes derivados de la aplicación de Big Data en la predicción y gestión de materia prima dentro del sector industrial. La planificación de compras y la gestión de inventarios de materia prima se ven afectadas por limitaciones estructurales y tecnológicas. De acuerdo con [10], una de las más comunes es la falta de integración entre sistemas, lo que obstaculiza la disponibilidad de datos en tiempo real y genera respuestas reactivas ante cambios en la demanda. Además, [22] señala que la variabilidad en los patrones de consumo ha vuelto obsoletos muchos modelos de pronóstico tradicionales, dificultando la anticipación precisa de abastecimiento. A esto se suma [13], la presión por reducir costos con esquemas ajustados como el *just-in-time*, poco adaptables frente a disrupciones o escenarios inciertos, afectando la continuidad operativa y la eficiencia general de las cadenas de suministro.

Se resalta la contribución adicional que aporta este estudio, centrada en evidenciar como la implementación de herramientas basadas en Big Data ha generado mejoras notables en la predicción de demanda y la eficiencia operativa en

distintos contextos industriales. Por ejemplo, modelos como DIM, basados en redes neuronales LSTM, alcanzaron una precisión del 82.33% en la predicción de la demanda, lo cual permitió una reducción del 25% en el costo total del inventario [10]. Esta tendencia de mejora se refuerza con otros enfoques como el uso de algoritmos ARIMA, que lograron una precisión del 98.41% en la planificación de demanda en cadenas de suministro automotrices, facilitando una gestión más precisa del stock [21]. Del mismo modo, se observó que la aplicación de algoritmos híbridos como BiGRU+CNN superó ampliamente a modelos tradicionales, alcanzando una precisión cercana al 94% frente al 86.57% de los árboles de decisión [17]. Esta evidencia resalta la ventaja de enfoques basados en aprendizaje profundo para datos complejos y entornos logísticos dinámicos. En términos de gestión de inventarios, se reportaron mejoras significativas como la reducción del 30% en el stock de materias primas [1] y una disminución del 20% en los faltantes de inventario, acompañadas de una mayor rotación de productos y reducción del 18% en los costos de mantenimiento [18]. Enfocándonos en el ámbito del control financiero, el uso de árboles de decisión permitió optimizar el ciclo de conversión de efectivo (CCC) con una precisión del 85%, incluso en escenarios de disrupción [16]. Asimismo, tecnologías como sistemas inteligentes de almacén basados en IoT elevaron la eficiencia del sistema hasta en un 88%, mejorando la gestión en tiempo real [14]. Finalmente, se evidenció que la incorporación de técnicas como la ingeniería de características en modelos Naïve Bayes mejoró la precisión en contextos de demanda intermitente, alcanzando un 92% de efectividad frente al 77% inicial [20]. En general, estos resultados confirman que las tecnologías de Big Data no solo optimizan la predicción, sino que también fortalecen el desempeño logístico y financiero de las organizaciones industriales.

La presente revisión sistemática enfrentó ciertas limitaciones que deben ser consideradas al interpretar sus hallazgos. La búsqueda se centró únicamente en la base de datos Scopus, lo que garantiza la calidad de las fuentes, pero pudo restringir el acceso a estudios relevantes en otras plataformas. Además, la mayoría de investigaciones analizadas provienen de contextos europeos y asiáticos, lo que dificulta extrapolar los resultados al entorno latinoamericano, donde las condiciones tecnológicas y organizacionales pueden variar considerablemente. Parte de la literatura consultada se basa en enfoques emergentes como la Industria 5.0, aún en proceso de consolidación, y varios estudios reconocen que el aprovechamiento de Big Data en entornos industriales reales puede verse limitado por las capacidades actuales de hardware y software. A pesar de los resultados prometedores, la aplicabilidad de las soluciones identificadas dependerá del tipo de industria y su nivel de madurez tecnológica. En este sentido, se recomienda impulsar investigaciones aplicadas en empresas manufactureras latinoamericanas que validen el impacto de herramientas como las redes neuronales LSTM, con alta precisión en la predicción de series temporales complejas [4], o

los gemelos digitales, capaces de simular escenarios logísticos disruptivos y optimizar la rotación de inventarios [16]. Para lograr una implementación eficaz, será clave fortalecer la capacitación técnica del personal y consolidar una cultura organizacional orientada a decisiones basadas en datos.

Se destaca la efectividad de herramientas como las redes LSTM, los modelos híbridos de aprendizaje automático, la analítica predictiva basada en Big Data y los gemelos digitales. Estas tecnologías son particularmente útiles en la predicción de materia prima en el sector industrial. Estas tecnologías han demostrado ser superiores en capturar la dinámica de la demanda, anticipar necesidades y ajustar de manera inteligente los niveles de inventario. Además, enfoques como el deep learning, el análisis por clustering y el monitoreo en tiempo real mediante IoT refuerzan la precisión y flexibilidad de las decisiones logísticas.

En conjunto, los resultados de esta revisión sistemática confirman que las herramientas basadas en Big Data permiten superar muchas de las limitaciones tradicionales en la gestión de inventarios y la planificación de materias primas, especialmente en entornos industriales complejos. Sin embargo, su implementación efectiva dependerá no solo de la infraestructura tecnológica disponible, sino también de factores organizacionales como la capacitación técnica del personal, la calidad de los datos y la apertura a una cultura de decisiones basadas en evidencia. Esta doble dimensión tecnológica y organizacional constituye el principal desafío para trasladar los beneficios documentados a contextos menos desarrollados, como muchas regiones de Latinoamérica, donde el potencial transformador del Big Data sigue siendo una oportunidad pendiente.

V. CONCLUSIÓN

Esta revisión sistemática permitió identificar, analizar y sintetizar evidencias relevantes sobre el uso de Big Data en la predicción de materia prima en el sector industrial. Se constató que la implementación de modelos predictivos avanzados, como redes neuronales LSTM, algoritmos ARIMA y enfoques híbridos que combinan aprendizaje profundo, ha contribuido significativamente a mejorar la precisión de los pronósticos, optimizar los inventarios, reducir costos logísticos y fortalecer la eficiencia operativa, incluso en contextos de alta variabilidad y demanda intermitente.

No obstante, se reconocen limitaciones importantes: la concentración de estudios en Asia y Europa, el enfoque emergente en Industria 5.0 y el uso exclusivo de la base Scopus, lo cual restringe la generalización de los hallazgos a entornos con menor madurez digital. Estos factores sugieren la necesidad de validar los resultados en otros contextos.

Se recomienda promover investigaciones empíricas en empresas latinoamericanas, especialmente manufactureras, que permitan evaluar la efectividad de estas tecnologías en condiciones reales. Asimismo, será fundamental fomentar procesos de capacitación técnica, inversión en infraestructura digital y una cultura organizacional orientada a la toma de decisiones basada en datos.

En suma, el uso estratégico del Big Data se configura como una herramienta clave para transformar la gestión de suministros en la industria, ofreciendo una respuesta eficiente, inteligente y sostenible a los desafíos del entorno productivo contemporáneo.

REFERENCIAS

- [1] V. G. Cannas, M. P. Ciano, M. Saltalamacchia, and R. Secchi, "Artificial intelligence in supply chain and operations management: a multiple case study research," *Int J Prod Res*, vol. 62, no. 9, pp. 3333–3360, 2024, doi: 10.1080/00207543.2023.2232050.
- [2] S. Zhang, Q. Yu, S. Wan, H. Cao, and Y. Huang, "Digital supply chain: literature review of seven related technologies," *Manuf Rev (Les Ulis)*, vol. 11, p. 8, Apr. 2024, doi: 10.1051/mfreview/2024006.
- [3] R. Rakholia, A. L. Suarez-Cetulo, M. Singh, and R. Simon Carbajo, "Advancing Manufacturing Through Artificial Intelligence: Current Landscape, Perspectives, Best Practices, Challenges, and Future Direction," *IEEE Access*, vol. 12, pp. 131621–131637, 2024, doi: 10.1109/ACCESS.2024.3458830.
- [4] M. Valencia-Cárdenas, J. A. Restrepo-Morales, and F. J. D. Serna, "Big Data Analytics in the Agribusiness Supply Chain Management," *Aibi, Revista de Investigacion Administracion e Ingenierias*, vol. 9, no. 3, pp. 32–42, 2021, doi: 10.15649/2346030X.2583.
- [5] D. H. Nabil, M. H. Rahman, A. H. Chowdhury, and 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, no. 2, 2023, doi: 10.1080/23311916.2023.2257924.
- [6] F. Stranieri, E. Fadda, and F. Stella, "Combining deep reinforcement learning and multi-stage stochastic programming to address the supply chain inventory management problem," *Int J Prod Econ*, vol. 268, 2024, doi: 10.1016/j.ijpe.2023.109099.
- [7] S. Siripurapu, N. K. Darimireddy, A. Chehri, B. Sridhar, and A. V. Paramkusam, "Technological Advancements and Elucidation Gadgets for Healthcare Applications: An Exhaustive Methodological Review-Part-I (AI, Big Data, Block Chain, Open-Source Technologies, and Cloud Computing)," *Electronics (Basel)*, vol. 12, no. 3, p. 750, Feb. 2023, doi: 10.3390/electronics12030750.
- [8] M. I. Khan *et al.*, "Integrating industry 4.0 for enhanced sustainability: Pathways and prospects," *Sustain Prod Consum*, vol. 54, pp. 149–189, Mar. 2025, doi: 10.1016/j.spc.2024.12.012.
- [9] K. Tomić-Pupek, I. Srpač, L. Havaš, and D. Srpač, "Algorithm for customizing the material selection process for application in power engineering," *Energies (Basel)*, vol. 13, no. 23, 2020, doi: 10.3390/en13236458.
- [10] C. Deng and Y. Liu, "A Deep Learning-Based Inventory Management and Demand Prediction Optimization Method for Anomaly Detection," *Wirel Commun Mob Comput*, vol. 2021, 2021, doi: 10.1155/2021/9969357.
- [11] D. García-Barrios, K. Palomino, E. García-Solano, and A. Cuello-Quiroz, "A Machine Learning based Method for Managing Multiple Impulse Purchase Products: An Inventory Management Approach," *Journal of Engineering Science and Technology Review*, vol. 14, no. 1, pp. 25–37, 2021, doi: 10.25103/jestr.141.02.
- [12] N. Dragomirov and L. Boyanov, "Supply Chain Management and Logistics Big Data Challenges in Bulgaria," *LOGI - Scientific Journal on Transport and Logistics*, vol. 12, no. 1, pp. 171–181, 2021, doi: 10.2478/logi-2021-0016.
- [13] Z. E. Huma *et al.*, "A Hybrid Deep Random Neural Network for Cyberattack Detection in the Industrial Internet of Things," *IEEE Access*, vol. 9, pp. 55595–55605, 2021, doi: 10.1109/ACCESS.2021.3071766.
- [14] M. G. Khan, N. Ul Huda, and U. K. Uz Zaman, "Smart Warehouse Management System: Architecture, Real-Time Implementation and Prototype Design," *Machines*, vol. 10, no. 2, 2022, doi: 10.3390/machines10020150.

- [15] A. E. Matenga and K. Mpofu, "Blockchain-Based Cloud Manufacturing SCM System for Collaborative Enterprise Manufacturing: A Case Study of Transport Manufacturing," *Applied Sciences (Switzerland)*, vol. 12, no. 17, 2022, doi: 10.3390/app12178664.
- [16] E. Badakhshan and P. Ball, "Applying digital twins for inventory and cash management in supply chains under physical and financial disruptions," *Int J Prod Res*, vol. 61, no. 15, pp. 5094–5116, 2023, doi: 10.1080/00207543.2022.2093682.
- [17] A. Alzahrani and M. Z. Asghar, "Intelligent Risk Prediction System in IoT-Based Supply Chain Management in Logistics Sector," *Electronics (Switzerland)*, vol. 12, no. 13, 2023, doi: 10.3390/electronics12132760.
- [18] X. U. Jie and B. O. Lixing, "Enhancing Supply Chain Efficiency Resilience using Predictive Analytics and Computational Intelligence Techniques," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3507161.
- [19] G. Rui and M. Li, "Utilizing Internet Big Data and Machine Learning for Product Demand Forecasting and Analysis of Its Economic Benefits," *Tehnicki Vjesnik*, vol. 31, no. 4, pp. 1385–1394, 2024, doi: 10.17559/TV-20240318001408.
- [20] A. K. Singh, J. B. Simha, and R. Agarwal, "Prediction of Intermittent Demand Occurrence using Machine Learning," *EAI Endorsed Transactions on Internet of Things*, vol. 10, 2024, doi: 10.4108/eetiot.5381.
- [21] A. Limbare and R. Agarwal, "Demand Forecasting and Budget Planning for Automotive Supply Chain," *EAI Endorsed Transactions on Internet of Things*, vol. 10, 2024, doi: 10.4108/eetiot.4514.
- [22] T. Mohammadi, S. M. Sajadi, S. E. Najafi, and M. Taghizadeh-Yazdi, "Multi Objective and Multi-Product Perishable Supply Chain with Vendor-Managed Inventory and IoT-Related Technologies," *Mathematics*, vol. 12, no. 5, 2024, doi: 10.3390/math12050679.
- [23] Y. Long, S. Kroeger, M. F. Zaeh, and A. Brintrup, "Leveraging synthetic data to tackle machine learning challenges in supply chains: challenges, methods, applications, and research opportunities," *Int J Prod Res*, 2025, doi: 10.1080/00207543.2024.2447927.
- [24] F. Li and J. Xu, "Revolutionizing AI-enabled Information Systems Using Integrated Big Data Analytics and Multi-modal Data Fusion," *IEEE Access*, 2025, doi: 10.1109/ACCESS.2025.3552039.
- [25] M. A. Amin, A. Chakraborty, and R. Baldacci, "Industry 5.0 and green supply chain management synergy for sustainable development in Bangladeshi RMG industries," *Cleaner Logistics and Supply Chain*, vol. 14, 2025, doi: 10.1016/j.clscn.2025.100208.
- [26] D.-H. Kim, G.-Y. Kim, and S. D. Noh, "Digital Twin-Based Prediction and Optimization for Dynamic Supply Chain Management," *Machines*, vol. 13, no. 2, 2025, doi: 10.3390/machines13020109.