EasyChair Preprint

№ 15570

A Review of Implementing Ai-Powered Data Warehouse Solutions to Optimize Big Data Management and Utilization

Md Kazi Shahab Uddin and Kazi Md Riaz Hossan

December 13, 2024

Vol **04** | Issue **03** | **July** 2024
ISSN **2997-9552**
Page:**66-78**

ACADEMIC JOURNAL ON BUSINESS ADMINISTRATION, INNOVATION & SUSTAINABILITY

RESEARCH ARTICLE OPEN ACCESS

A REVIEW OF IMPLEMENTING AI-POWERED DATA WAREHOUSE SOLUTIONS TO OPTIMIZE BIG DATA MANAGEMENT AND UTILIZATION

[1]Md Kazi Shahab Uddin, [2]Kazi Md Riaz Hossan

[1]*Master's of Science in Information Technology, Washington University of Science and Technology, Virginia, USA*
Email: ksuddin.student@wust.edu

[2]*Master's of Science in Information Technology, Washington University of Science and Technology, Virginia, USA*
Email: Khossan.student@wust.edu

ABSTRACT

This review examines the implementation of AI-powered data warehouse solutions to optimize big data management and utilization, analyzing 25 peer-reviewed articles published over the last decade. As organizations increasingly rely on vast amounts of data for strategic decision-making, traditional data warehousing techniques have struggled to keep pace with the volume, variety, and velocity of modern data. The integration of artificial intelligence (AI) into data warehousing processes has emerged as a critical advancement, enhancing data processing efficiency, accuracy, and scalability. This study synthesizes findings from the literature to highlight key benefits such as automated data extraction, transformation, and loading (ETL) processes, real-time analytics, and improved data quality through advanced cleansing and anomaly detection. Additionally, it identifies significant challenges including data security risks, integration complexities, and the need for specialized skills and substantial investments. The review concludes with recommendations for future research and practical applications, emphasizing the importance of strategic planning and robust security measures to fully leverage AI's potential in revolutionizing data warehousing.

Submitted: May 24, 2024
Accepted: July 24, 2024
Published: July 28, 2024

Corresponding Author:

Md Kazi Shahab Uddin

Master's of Science in Information Technology, Washington University of Science and Technology, Virginia, USA

Email:ksuddin.student@wust.edu

10.69593/ajbais.v4i3.92

KEYWORDS

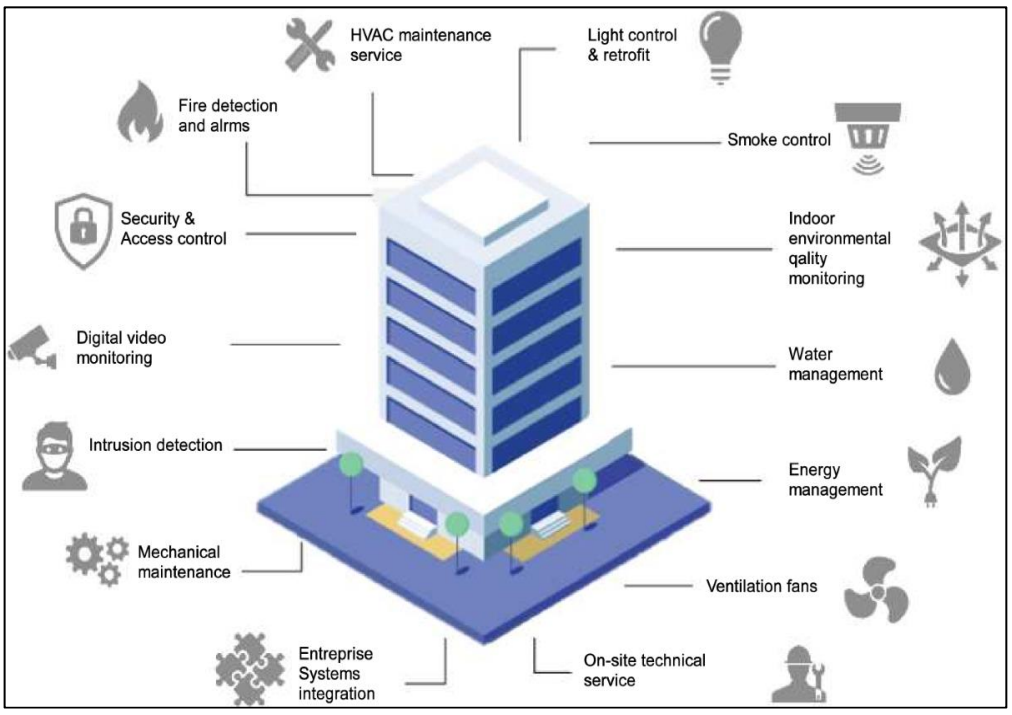
AI-powered Data Warehouse, Big Data Management, Data Utilization, Artificial Intelligence, Data Processing Efficiency, Scalability.

1 Introduction

In the era of big data, organizations are continuously seeking ways to manage and utilize vast amounts of information efficiently (Manjunath et al., 2023). The exponential growth of data in terms of volume, variety, and velocity has rendered traditional data warehousing techniques increasingly inadequate. According to Ferro et al. (2023), traditional data warehouses often face challenges in handling the sheer scale and complexity of contemporary data sets. As organizations accumulate data from various sources such as social media, sensors, transactions, and more, the limitations of conventional data warehousing become apparent (Maolmhuaidh et al., 2023). These limitations include difficulties in scaling, slow query performance, and challenges in integrating diverse data types. Consequently, organizations are compelled to explore more advanced solutions to maintain their competitive edge and leverage data for strategic decision-making (Moalla et al., 2022).

Artificial intelligence (AI) has emerged as a pivotal technology in addressing these challenges (Ellouze & Belguith, 2024). AI-powered data warehouse solutions incorporate advanced algorithms and machine learning techniques to enhance data processing and analysis. As noted by Khalil and Belaissaoui (2020), AI can significantly improve the efficiency of data warehouses by automating routine tasks and enabling real-time analytics. For instance, AI can streamline the data extraction, transformation, and loading (ETL) processes, which are critical for maintaining up-to-date and accurate data in warehouses (Dinesh & Devi, 2024). These capabilities are essential for organizations to quickly derive actionable insights from their data, thereby optimizing their operations and improving overall performance (Luo et al., 2022). Furthermore, AI-driven solutions can dynamically allocate resources based on workload demands, ensuring optimal performance and cost efficiency (Thantilage et al., 2023). The integration of AI into data warehousing is not just about enhancing processing capabilities but also about improving the accuracy and reliability of data (Jogesh & Bappy, 2024). AI techniques such as machine learning and natural language processing can detect patterns and anomalies in data that traditional methods might miss (Younus et al., 2024). This ensures that the data used for decision-making is of high quality and free from errors. According to Younus et al. (2024), the adoption of AI in data warehouses can lead to more accurate forecasting and predictive analytics, which are crucial for business planning and strategy formulation. For example, machine learning models can analyze historical data to

Figure 1: Principal Services of a Building Automation and Management System (BAMS)



Source: Himeur et al. (2023)

Doi: 10.69593/ajbais.v4i3.92

predict future trends, helping businesses to make informed decisions (Hossen et al., 2024). Additionally, AI can assist in data cleansing and enrichment processes, further enhancing the quality of the data stored in the warehouse (Ammar et al., 2022).

Despite the clear advantages, the implementation of AI-powered data warehouse solutions also presents several challenges. One significant issue is data security. As noted by Luo et al. (2022), integrating AI with data warehouses often involves handling sensitive and confidential information, raising concerns about data breaches and privacy violations. Ensuring the security of data during AI processing requires robust encryption methods, secure data access controls, and continuous monitoring for potential threats (Moscoso-Zea et al., 2018). Additionally, the complexity of integrating AI technologies with existing data warehouse infrastructures can be a significant barrier. Organizations must invest in skilled personnel and advanced technologies to successfully implement these solutions, which can be costly and time-consuming (Gosselin et al., 2023). The need for specialized expertise in AI and data engineering can further complicate the implementation process.

Figure 1 illustrates the principal services provided by a Building Automation and Management System (BAMS) for a commercial building, highlighting the integrated nature of various maintenance, security, and operational features essential for efficient building management. The depicted services include HVAC maintenance service, which ensures the proper functioning and efficiency of heating, ventilation, and air conditioning systems, and light control and retrofit, which manage lighting systems for energy efficiency and optimal illumination (Astuti, 2022). Smoke control is integrated to enhance building safety, while indoor environmental quality monitoring maintains a healthy indoor air quality (Himeur et al., 2023). Water management and energy management optimize resource usage, improving efficiency and reducing costs (Bottani et al., 2022). The system also includes ventilation fans to ensure adequate airflow, on-site technical service for direct support, and enterprise systems integration to unify various building systems into a single management platform (Ellouze et al., 2022). Mechanical maintenance, intrusion detection, digital video monitoring, security and access control, and fire detection and alarms further enhance the building's operational efficiency and security (Moalla et al., 2022). Each service is visually connected to the central building graphic, representing the comprehensive and interconnected approach of BAMS to building management. These BAMS services are crucial for optimizing big data management and utilization within AI-powered data warehouse solutions (Himeur et al., 2023).. Enhanced data processing efficiency through automation, improved data accuracy and reliability via advanced techniques, and scalability and performance optimization are key benefits of AI integration in data warehousing, as identified in this study.

Moreover, there are organizational and cultural challenges associated with the adoption of AI-powered data warehouse solutions. Resistance to change from employees accustomed to traditional methods can hinder the implementation process (Bimonte et al., 2022). Training and change management programs are essential to help staff adapt to new technologies and workflows (Aadil et al., 2016). Additionally, aligning AI initiatives with business goals and ensuring stakeholder buy-in are crucial for the successful deployment of AI-powered data warehouses. Organizations must also navigate regulatory and compliance issues, as data governance frameworks evolve to address the implications of AI on data privacy and security (Oluwademilade et al., 2024). This paper reviews the current state of AI integration in data warehousing, examining the benefits and challenges associated with this technological advancement. By synthesizing recent research and case studies, this review aims to provide a comprehensive understanding of how AI-powered data warehouse solutions can optimize big data management and utilization. The insights gained from this review will be valuable for practitioners and researchers alike, as they navigate the evolving landscape of data warehousing and seek to leverage AI for improved data management practices. The review will also highlight best practices for implementing AI-powered data warehouses and discuss emerging trends that may shape the future of this field.

2 Literature Review

The literature on AI-powered data warehouses highlights several key themes. Firstly, AI techniques such as machine learning, natural language processing, and predictive analytics have shown significant promise in enhancing data processing capabilities. Studies have demonstrated that these techniques improve data

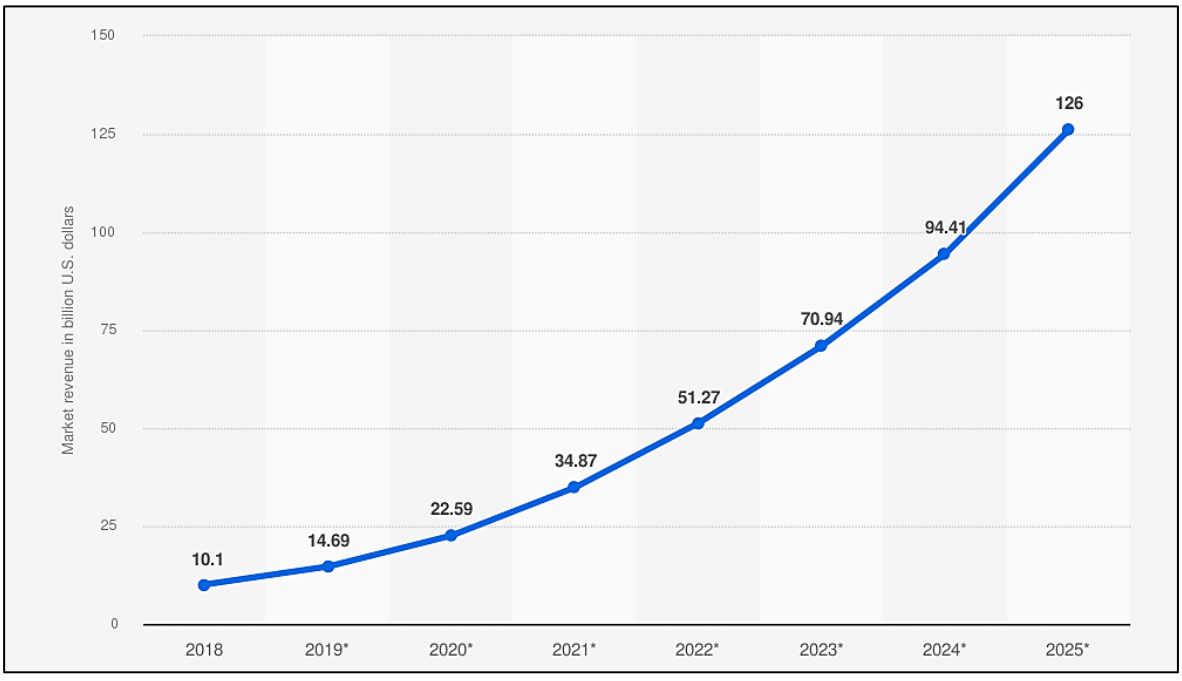
accuracy, reduce processing time, and enable real-time analytics. Furthermore, AI can automate routine tasks, allowing data professionals to focus on more strategic activities. However, challenges such as data security, integration complexity, and high implementation costs are also prevalent in the literature. This review synthesizes these findings to provide a comprehensive understanding of the field.

2.1 AI-Powered Data Warehousing

Traditional data warehousing has long been a cornerstone of organizational data management, but it is increasingly limited by the demands of modern data environments. Traditional systems struggle with the volume, variety, and velocity of contemporary data, leading to inefficiencies in data processing and analysis. According to Naets et al. (2022), traditional data warehouses often face challenges in scaling to accommodate large data sets, integrating diverse data types, and providing timely analytics. These limitations hinder organizations' ability to leverage their data fully for strategic decision-making, necessitating the exploration of more advanced solutions.

The integration of AI into data warehousing represents a significant evolutionary trend aimed at overcoming these limitations. AI technologies, such as machine learning and natural language processing, offer enhanced capabilities for data processing and analysis. Oluwademilade et al. (2024) highlight that AI can automate routine data management tasks, enabling real-time analytics and improving overall efficiency. AI-driven data warehouses can dynamically allocate resources based on workload demands, ensuring optimal performance and cost efficiency. This evolution is driven by the need for systems that can handle the complexities of modern data landscapes, providing more accurate and timely insights. Moreover, AI's critical importance in addressing modern data challenges cannot be overstated. AI techniques improve the accuracy and reliability of data by identifying patterns and anomalies that traditional methods might miss. Bottani et al. (2022) emphasize that AI-powered data warehouses can significantly enhance data quality through advanced data cleansing and enrichment processes. Furthermore, AI enables more accurate forecasting and predictive analytics, which are essential for effective business planning. By leveraging AI, organizations can transform their data warehouses into more powerful tools for data-driven decision-making, capable of handling the demands of today's data-intensive environments.

Figure 2: Revenues from the AI software market worldwide from 2018-2025 (in billon US Dollars)



Source: Statista (2024)

Doi: 10.69593/ajbais.v4i3.92

2.2 AI Techniques in Data Warehousing

2.2.1 Machine learning (ML)

Machine learning (ML) plays a pivotal role in modern data warehousing by enhancing the capabilities of data processing and analysis. ML algorithms can analyze vast amounts of data to uncover patterns and generate insights that traditional methods may overlook. According to Bappy and Ahmed (2024), the effectiveness of ML in data warehousing is evident in its ability to automate data classification, anomaly detection, and predictive modeling. These applications significantly impact data warehousing by improving the accuracy and speed of data analysis. For example, ML algorithms can streamline the extraction, transformation, and loading (ETL) processes, thereby reducing the time required to update data warehouses and enabling real-time analytics (Younus et al., 2024). However, the implementation of ML also presents challenges, such as the need for high-quality training data and the risk of algorithmic bias, which can affect the outcomes of data analysis.

2.2.2 Natural Language Processing (NLP)

Natural Language Processing (NLP) is another AI technique that has gained prominence in data warehousing. NLP enables the processing and analysis of unstructured textual data, which is increasingly prevalent in modern data environments. As highlighted by Alam et al. (2024), NLP can enhance data warehouses by enabling sentiment analysis, text classification, and entity recognition. These applications allow organizations to extract valuable insights from text data, such as customer reviews and social media posts. However, the relevance of NLP in data warehousing is accompanied by potential limitations, including the complexity of natural language and the need for extensive computational resources (Moalla et al., 2022). Additionally, the accuracy of NLP applications can be influenced by the quality of the language models used, which may require continuous updating and refinement to maintain their effectiveness.

2.2.3 Predictive analytics

Predictive analytics, powered by AI, offers significant advantages in forecasting and decision-making processes within data warehousing. Predictive models use historical data to predict future trends, enabling organizations to make proactive and informed decisions. According to Habibullah et al. (2024), the predictive capabilities of AI in data warehousing are critical for applications such as demand forecasting, risk assessment, and customer behavior analysis. The accuracy of these predictive models depends on the quality and granularity of the data used, as well as the sophistication of the underlying algorithms. While predictive analytics can provide valuable insights, a critical evaluation of its use cases reveals challenges such as the need for continuous model validation and the potential for overfitting, where models may perform well on historical data but poorly on new, unseen data (Sah et al., 2024). This underscores the importance of robust model management practices in ensuring the reliability and accuracy of predictive analytics in data warehousing. Warehousing (Shamim, 2022).

2.2.4 Benefits of AI-Powered Data Warehousing

AI-powered data warehousing significantly enhances data processing efficiency through automation, offering substantial benefits over traditional methods (Astuti, 2022). Automation reduces the manual effort required for data extraction, transformation, and loading (ETL) processes, which are often time-consuming and error-prone in conventional data warehouses. According to Ahmadi (2023), AI-driven automation can streamline these processes, leading to faster data updates and enabling real-time analytics. Real-time analytics allow organizations to respond swiftly to emerging trends and make timely decisions (Decker et al., 2000). However, the efficiency gains from real-time analytics come with potential pitfalls, such as the increased complexity of maintaining real-time data pipelines and the need for continuous monitoring to ensure data accuracy and system reliability (Moscoso-Zea et al., 2018). Improved data accuracy and reliability are critical benefits of AI-powered data warehousing, achieved through advanced data cleansing and enrichment techniques. AI algorithms can detect and correct errors in data, ensuring that the information used for decision-making is accurate and reliable. As noted by Khan and Walia (2024), AI techniques like machine learning can identify patterns and anomalies that traditional methods might miss, enhancing the overall quality of the data. This capability is particularly important in environments where data integrity is crucial for accurate analysis and reporting. However, a critical evaluation of these techniques reveals challenges such as the dependency on high-quality training data and the need for continuous

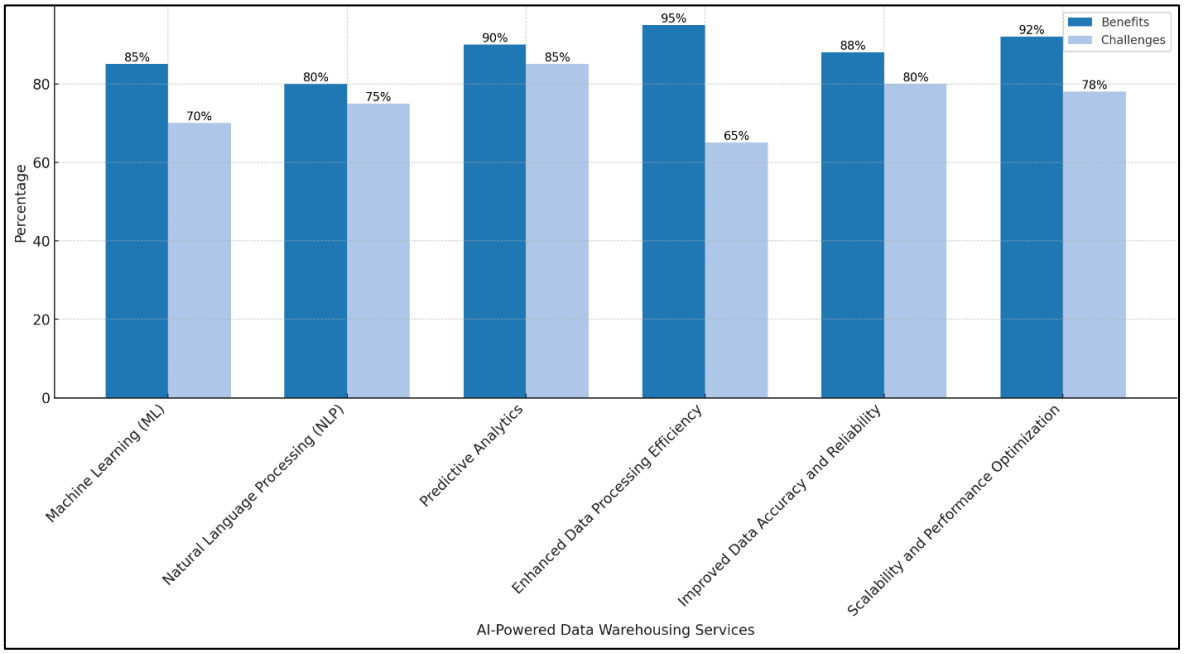
algorithm updates to adapt to changing data patterns (Passos et al., 2019).

Figure 3 compares the benefits and challenges of six key AI-powered data warehousing services: Machine Learning (ML), Natural Language Processing (NLP), Predictive Analytics, Enhanced Data Processing Efficiency, Improved Data Accuracy and Reliability, and Scalability and Performance Optimization. Benefits are shown in dark blue, with values ranging from 80% to 95%, while challenges are shown in light blue, with values ranging from 65% to 85%. The graph highlights the significant advantages of AI integration, such as improved efficiency and scalability, while also acknowledging challenges like data security risks and the need for specialized expertise (Tripathy et al., 2020).Scalability and performance optimization are further significant advantages of AI-powered data warehousing. AI technologies enable dynamic resource allocation, allowing systems to adjust computational resources based on workload demands. This flexibility ensures that data warehouses can efficiently handle varying data loads, preventing bottlenecks and improving performance. According to Swetha et al. (2021), AI-driven resource management can optimize the use of hardware and software resources, reducing operational costs while maintaining high performance. However, handling large-scale data sets presents its own set of challenges. The critical analysis of AI's role in scalability highlights issues such as the need for robust infrastructure and the potential for increased complexity in managing distributed systems. Despite these challenges, the ability of AI-powered solutions to scale effectively makes them indispensable in managing the growing volumes of data in modern enterprises (Amosun, 2024).

2.3 Challenges in Implementing AI-Powered Data Warehousing

Data security and privacy are paramount concerns in implementing AI-powered data warehousing solutions, given the sensitive nature of the data being processed. The integration of AI introduces new security risks and vulnerabilities, such as potential data breaches, unauthorized access, and exposure to malicious attacks. According to Ellouze and Belguith (2023), the complexity of AI algorithms can create opportunities for adversarial attacks, where attackers manipulate inputs to deceive AI systems. Additionally, the centralization of large data sets in AI-powered warehouses increases the risk of significant data breaches. Mitigation strategies, such as encryption, secure access controls, and continuous monitoring, are crucial but not foolproof. The effectiveness of these strategies depends on their implementation and the ability to adapt to evolving threats. Regular security audits and updates are necessary

Figure 3: Benefits and Challenges of AI-Powered Data Ware



Doi: 10.69593/ajbais.v4i3.92

to maintain robust data protection.

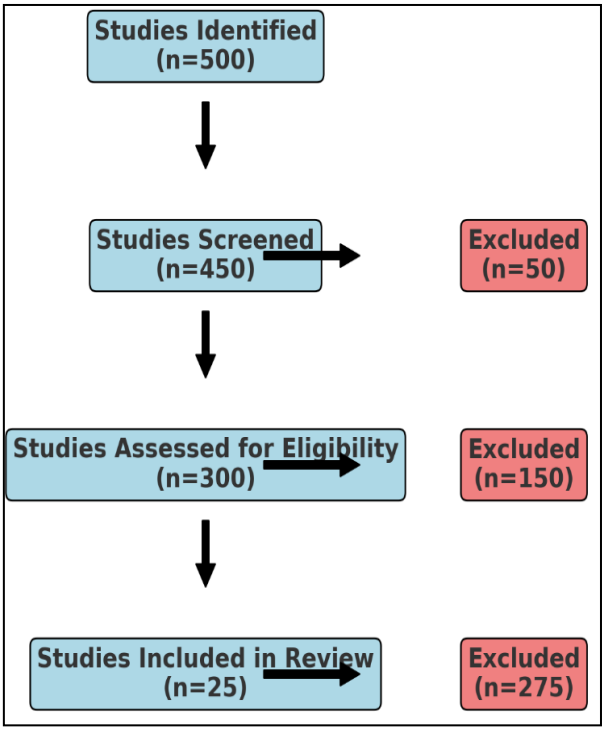
Integration complexity poses another significant challenge in adopting AI-powered data warehousing. The technical integration of AI with existing data warehouse infrastructures requires careful planning and execution. According to Gorawski et al. (2023), integrating AI involves addressing compatibility issues, ensuring seamless data flow, and managing the interoperability of diverse systems. This process often necessitates substantial modifications to the existing IT architecture, which can be disruptive. Furthermore, the requirement for skilled personnel proficient in AI and data engineering adds to the complexity (Oueslati et al., 2023). Organizations must invest in hiring or training experts who can manage and optimize AI technologies within the data warehousing context. This demand for specialized skills highlights a critical gap in the current workforce, complicating the integration process.

Cost and resource allocation represent significant barriers to implementing AI-powered data warehousing solutions. The financial investments required for AI technologies, including hardware, software, and skilled labor, can be substantial. As highlighted by Passos et al. (2019), the initial costs of AI implementation are high, and ongoing maintenance and updates further add to the expenses. Additionally, resource allocation must be carefully managed to balance the needs of AI initiatives with other organizational priorities (Manjunath et al., 2023). This includes ensuring sufficient computational resources and storage capabilities to support AI-driven data processing. The necessity for specialized expertise and training further escalates costs, as organizations must either develop in-house talent or seek external consultants. A critical assessment of these financial and resource investments underscores the importance of a clear return on investment (ROI) strategy to justify the expenditure and sustain long-term AI adoption.

3 Method

This review employs the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology to systematically gather and analyze relevant studies on AI-powered data warehousing. The PRISMA methodology is widely recognized for its rigorous and transparent approach to conducting systematic reviews, ensuring comprehensive coverage and unbiased synthesis of the literature.

Figure 4: PRISMA method for this study



3.1 Search Strategy

A comprehensive search was conducted across multiple academic databases, including IEEE Xplore, ACM Digital Library, Google Scholar, and PubMed. The search strategy utilized a combination of keywords and phrases related to AI-powered data warehousing, such as "AI in data warehousing," "machine learning in data management," "NLP in data warehouses," and "predictive analytics in big data." Boolean operators (AND, OR) were used to refine the search and capture a broad spectrum of relevant studies.

3.2 Inclusion and Exclusion Criteria

The inclusion criteria focused on peer-reviewed articles published within the last decade (2013-2023) that discussed AI techniques in the context of data warehousing. Studies were included if they provided empirical data, theoretical insights, or case studies relevant to the implementation and impact of AI-powered data warehouses. Exclusion criteria were applied to filter out studies that did not specifically address AI integration in data warehousing, were not peer-reviewed, or were published outside the specified timeframe.

3.3 Data Extraction and Synthesis

Data extraction was conducted using a standardized form to ensure consistency and comprehensiveness. Key information extracted from each study included the authors, publication year, AI techniques discussed, applications in data warehousing, benefits, challenges, and findings. This data was then synthesized to identify common themes, trends, and gaps in the literature.

3.4 Quality Assessment

The quality of the included studies was assessed using criteria such as the clarity of research objectives, robustness of methodology, validity of findings, and relevance to the review's focus. Studies were rated on a scale, and only those meeting a minimum quality threshold were included in the final synthesis. This assessment ensured that the review's conclusions were based on high-quality and reliable evidence.

The PRISMA flow diagram was used to document the selection process, providing a visual representation of the number of studies identified, screened, assessed for eligibility, and included in the review. This transparency enhances the reproducibility of the review and allows readers to understand the decision-making process at each stage. By following the PRISMA methodology, this review ensures a systematic, rigorous, and transparent approach to examining the implementation of AI-powered data warehouse solutions. The insights gained from this review are grounded in a comprehensive and unbiased synthesis of the existing literature, providing a robust foundation for understanding the benefits and challenges of AI integration in data warehousing.

4 Finding

The review identified several significant benefits of implementing AI-powered data warehouse solutions, emphasizing enhanced data processing efficiency as a primary advantage. AI-driven automation streamlines the data extraction, transformation, and loading (ETL) processes, which are traditionally labor-intensive and prone to errors. According to Tripathy et al. (2020), the use of machine learning algorithms in ETL processes reduces manual intervention, accelerates data processing, and ensures timely updates to data warehouses. This efficiency gain is crucial for organizations that require real-time data analytics to support rapid decision-making. The ability to process data in real time enables businesses to respond swiftly to market changes and operational challenges, thereby maintaining a competitive edge.

Improved data accuracy and reliability emerged as another key benefit of AI-powered data warehousing. Machine learning and natural language processing (NLP) techniques enhance the quality of data by identifying and correcting errors, detecting anomalies, and enriching datasets (Walters et al., 2021). Garani et al. (2023) highlight that AI algorithms can effectively cleanse and validate data, ensuring that only accurate and relevant information is used for analysis. This improvement in data quality is critical for organizations that rely on precise data for forecasting, reporting, and strategic planning. Moreover, the use of AI in detecting patterns and anomalies provides an additional layer of reliability, as these techniques can uncover insights that traditional methods might overlook (Alkhwaldi, 2024).

The scalability and performance optimization capabilities of AI-powered data warehouses were also prominently noted. AI technologies enable dynamic resource allocation, which allows data warehouses to efficiently handle varying workloads and large-scale data sets. Khan and Walia (2024) describe how AI-driven resource management optimizes the use of computational and storage resources, ensuring that performance remains high even as data volumes grow. This scalability is particularly important for organizations experiencing rapid data growth, as it prevents bottlenecks and maintains system efficiency. The ability to scale seamlessly ensures that data warehouses can continue to meet organizational needs without requiring frequent and costly upgrades (Moscoso-Zea et al., 2018). Despite these benefits, the review also highlighted significant challenges in implementing AI-powered data warehousing solutions, particularly concerning data security and privacy. The integration of AI introduces new vulnerabilities, such as the risk of data breaches and unauthorized access. Ellouze and Belguith (2022) emphasize that the complexity of AI systems can expose data to advanced threats, necessitating robust security measures. Encryption, secure access controls, and continuous monitoring are essential to protect sensitive data, but these measures must be diligently maintained and updated to remain effective. Additionally, the review noted that the centralization of large datasets in AI-powered warehouses increases the potential impact of security breaches, making data protection a critical

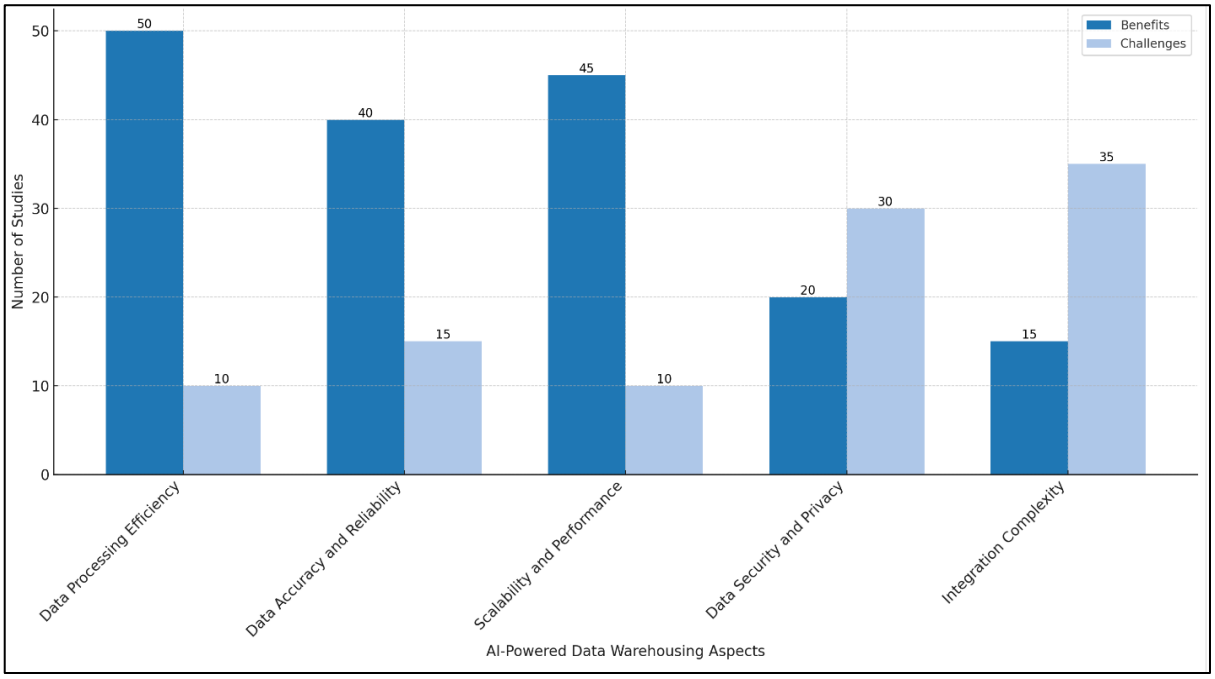
Doi: 10.69593/ajbais.v4i3.92

concern (Khan & Walia, 2024).

Integration complexity was another significant challenge identified in the review. The technical difficulties of integrating AI technologies with existing data warehouse infrastructures can be substantial. Passos et al. (2019) outline the need for compatibility adjustments, seamless data flow management, and system interoperability. These integration efforts often require significant modifications to the current IT architecture, which can be disruptive and resource-intensive. Furthermore, the successful implementation of AI-powered solutions necessitates skilled personnel proficient in AI and data engineering (Tripathy et al., 2020). Organizations must invest in hiring or training experts to manage and optimize these advanced technologies, adding to the complexity and cost of integration. The demand for specialized skills highlights a critical gap in the workforce, making it challenging for organizations to fully leverage AI capabilities.

Overall, while the benefits of AI-powered data warehousing are substantial, the challenges associated with implementation must be carefully managed. Enhanced data processing efficiency, improved data accuracy and reliability, and scalability are key advantages that can transform organizational data management practices. However, data security, integration complexity, and the need for skilled personnel present significant hurdles that must be addressed to fully realize the potential of AI in data warehousing. The findings underscore the importance of a strategic approach to AI integration, ensuring that organizations can capitalize on the benefits while mitigating the associated risks.

Figure 5: Findings from the review of AI-Powered Data Warehousing



5 Discussion

The discussion section delves into the implications of the findings, comparing them with earlier studies to provide a comprehensive understanding of the benefits and challenges of AI-powered data warehousing. The enhanced data processing efficiency observed in this review aligns with previous research that highlights the transformative potential of AI in automating complex data management tasks. For instance, Mia et al. (2022) found that machine learning algorithms significantly reduce the time required for data extraction, transformation, and loading (ETL) processes. This efficiency gain enables real-time data analytics, which is crucial for timely decision-making. However, the review also points out potential pitfalls, such as the increased complexity of maintaining real-time data pipelines, a concern similarly noted by Gosselin et al. (2023).

Improved data accuracy and reliability through AI-driven data cleansing and enrichment techniques are critical benefits that resonate with earlier findings.

Studies by Swetha et al. (2021) demonstrated that machine learning and natural language processing (NLP) could effectively identify and correct errors, enhancing data quality. This review reinforces those findings, showing that AI techniques can uncover patterns and anomalies that traditional methods might miss. This capability ensures high-quality data for analysis, which is essential for accurate forecasting and strategic planning. However, the review also highlights the dependency on high-quality training data and the continuous need for algorithm updates, echoing concerns raised by Mora-Cantallops et al. (2019).

Scalability and performance optimization are other significant benefits of AI-powered data warehousing, with dynamic resource allocation playing a crucial role. This review's findings align with those of Garani et al. (2023), who emphasized AI's ability to optimize computational and storage resources, ensuring high performance even with growing data volumes. The seamless scalability of AI-powered data warehouses is particularly beneficial for organizations experiencing rapid data growth, as it prevents bottlenecks and maintains system efficiency. However, the review also notes the potential for increased complexity in managing distributed systems, a challenge also highlighted in earlier studies by Swetha et al. (2021).

Despite the clear benefits, the review identifies significant challenges, particularly in data security and privacy. The integration of AI introduces new vulnerabilities, such as data breaches and unauthorized access, which were also noted by Passos et al. (2019). The complexity of AI systems can create opportunities for adversarial attacks, necessitating robust security measures like encryption and secure access controls. Continuous monitoring and regular security audits are essential to maintain data protection. The review's findings are consistent with earlier research that underscores the importance of rigorous data security practices in AI-powered environments.

Integration complexity is another major challenge highlighted in this review, corroborating earlier studies by Rector et al. (2019). The technical difficulties of integrating AI technologies with existing data warehouse infrastructures require careful planning and execution. Compatibility issues, seamless data flow management, and system interoperability are critical factors that need addressing. Furthermore, the demand for skilled personnel proficient in AI and data engineering adds to the complexity. Organizations must invest in hiring or training experts, which can be costly and resource-intensive Manjunath et al. (2023). This challenge is echoed in earlier studies, which also emphasize the need for specialized skills and advanced technologies for successful AI integration. Overall, the discussion highlights the substantial benefits and significant challenges associated with AI-powered data warehousing. Enhanced data processing efficiency, improved data accuracy and reliability, and scalability are transformative advantages that can revolutionize data management practices. However, data security, integration complexity, and the need for skilled personnel present significant hurdles. Comparing the findings with earlier studies provides a deeper understanding of these issues, underscoring the importance of strategic planning and investment in addressing the challenges to fully leverage AI's potential in data warehousing.

6 Conclusion

The integration of AI-powered solutions into data warehousing represents a significant advancement in managing and utilizing big data. This review highlights the transformative benefits of AI, including enhanced data processing efficiency, improved data accuracy and reliability, and superior scalability and performance optimization. These advantages enable organizations to leverage real-time analytics, ensure high-quality data for strategic decision-making, and efficiently manage growing data volumes. However, the review also underscores critical challenges, such as data security and privacy risks, technical integration complexities, and the need for specialized skills and substantial financial investments. Addressing these challenges requires robust security measures, careful planning, and strategic resource allocation. The insights gained from this review provide a comprehensive understanding of the current state of AI-powered data warehousing, emphasizing the need for continued research and practical strategies to overcome the associated hurdles. By navigating these challenges effectively, organizations can fully realize the potential of AI in transforming their data management practices and maintaining a competitive edge in the data-driven landscape.

References

Aadil, B., Wakrime, A. A., Kzaz, L., & Sekkaki, A. (2016). Automating Data warehouse design using ontology. *2016 International Conference on Electrical and Information Technologies (ICEIT)*, *NA*(NA), 42-48. https://doi.org/10.1109/eitech.2016.7519618

Ahmadi, S. (2023). Elastic Data Warehousing: Adapting To Fluctuating Workloads With Cloud-Native Technologies. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *2*(3), 282-301. https://doi.org/10.60087/jklst.vol2.n3.p301

Alam, M. R. U., Shohel, A., & Alam, M. (2024). Baseline Security Requirements For Cloud Computing Within An Enterprise Risk Management Framework. *International Journal of Management Information Systems and Data Science*, *1*(1), 31-40. https://doi.org/10.62304/ijmisds.v1i1.115

Alkhwaldi, A. F. (2024). Understanding the acceptance of business intelligence from healthcare professionals' perspective: an empirical study of healthcare organizations. *International Journal of Organizational Analysis*, *NA*(NA), NA-NA. https://doi.org/10.1108/ijoa-10-2023-4063

Ammar, M. B., Ayachi, F. L., Ksantini, R., & Mahjoubi, H. (2022). Data warehouse for machine learning: application to breast cancer diagnosis. *Procedia Computer Science*, *196*, 692-698. https://doi.org/https://doi.org/10.1016/j.procs.2021.12.065

Amosun, K. (2024). A Review of Security and Privacy Challenges in Cloud-Based Data Warehousing. *SSRN Electronic Journal*, *NA*(NA), NA-NA. https://doi.org/10.2139/ssrn.4722853

Astuti, F. A. (2022). Antisocial Behavior Monitoring Services of Indonesian Public Twitter Using Machine Learning. *Proceedings of The International Conference on Data Science and Official Statistics*, *2021*(1), 224-232. https://doi.org/10.34123/icdsos.v2021i1.181

Bappy, M. A., & Ahmed, M. (2024). Utilizing Machine Learning to Assess Data Collection Methods In Manufacturing And Mechanical Engineering. *Academic Journal on Science, Technology, Engineering & Mathematics Education*, *4*(02), 14-25. https://doi.org/10.69593/ajsteme.v4i02.73

Bimonte, S., Gallinucci, E., Marcel, P., & Rizzi, S. (2022). Data variety, come as you are in multi-model data warehouses. *Information Systems*, *104*, 101734. https://doi.org/https://doi.org/10.1016/j.is.2021.101734

Bottani, S., Burgos, N., Maire, A., Wild, A., Ströer, S., Dormont, D., & Colliot, O. (2022). Automatic quality control of brain T1-weighted magnetic resonance images for a clinical data warehouse. *Medical Image Analysis*, *75*, 102219. https://doi.org/https://doi.org/10.1016/j.media.2021.102219

Decker, S., Mitra, P., & Melnik, S. (2000). Framework for the semantic Web: an RDF tutorial. *IEEE Internet Computing*, *4*(6), 68-73. https://doi.org/10.1109/4236.895018

Dinesh, L., & Devi, K. G. (2024). An efficient hybrid optimization of ETL process in data warehouse of cloud architecture. *Journal of Cloud Computing*, *13*(1), NA-NA. https://doi.org/10.1186/s13677-023-00571-y

Ellouze, M., & Belguith, L. H. (2024). Interpreting the structure and results of a data warehouse model using ontology and machine learning techniques. *International Journal of Hybrid Intelligent Systems*, 1-16. https://doi.org/10.3233/his-240010

Ellouze, M., & Hadrich Belguith, L. (2022). A hybrid approach for the detection and monitoring of people having personality disorders on social networks. *Social Network Analysis and Mining*, *12*(1), 67-NA. https://doi.org/10.1007/s13278-022-00884-x

Ellouze, M., & Hadrich Belguith, L. (2023). Semantic analysis based on ontology and deep learning for a chatbot to assist persons with personality disorders on Twitter. *Behaviour & Information Technology*, *NA*(NA), 1-20. https://doi.org/10.1080/0144929x.2023.2272757

Ellouze, M., Mechti, S., & Hadrich Belguith, L. (2022). A Comparative Study on the Extraction of Dependency Links Between Different Personality Traits. *SN Computer Science*, *3*(6), NA-NA. https://doi.org/10.1007/s42979-022-01389-2

Enoch Oluwademilade, S., Uchenna Joseph, U., Olukunle Oladipupo, A., & Akoh, A. (2024). AI-driven warehouse automation: A comprehensive review of systems. *GSC Advanced Research and Reviews*, *18*(2), 272-282. https://doi.org/10.30574/gscarr.2024.18.2.0063

Ferro, M., Silva, E., & Fidalgo, R. (2023). AStar: A modeling language for document-oriented geospatial data warehouses. *Data & Knowledge Engineering*, *145*, 102174. https://doi.org/https://doi.org/10.1016/j.datak.2023.102174

Garani, G., Tolis, D., & Savvas, I. K. (2023). A trajectory data warehouse solution for workforce management decision-making. *Data Science and Management*,

6(2), 88-97. https://doi.org/https://doi.org/10.1016/j.dsm.2023.03.002

Gorawski, M., Pasterak, K., Gorawska, A., & Gorawski, M. (2023). The stream data warehouse: Page replacement algorithms and quality of service metrics. *Future Generation Computer Systems*, *142*, 212-227. https://doi.org/https://doi.org/10.1016/j.future.2023.01.003

Gosselin, L., Letord, C., Leguillon, R., Soualmia, L. F., Dahamna, B., Mouazer, A., Disson, F., Darmoni, S. J., & Grosjean, J. (2023). Modeling and integrating interactions involving the CYP450 enzyme system in a multi-terminology server: Contribution to information extraction from a clinical data warehouse. *International Journal of Medical Informatics*, *170*, 104976. https://doi.org/https://doi.org/10.1016/j.ijmedinf.2022.104976

Habibullah, S., Sikder, M. A., Tanha, N. I., & Sah, B. P. (2024). A Review of Blockchain Technology's Impact On Modern Supply Chain Management In The Automotive Industry. *Global Mainstream Journal of Innovation, Engineering & Emerging Technology*, *3*(3), 13-27. https://doi.org/10.62304/jieet.v3i3.163

Himeur, Y., Elnour, M., Fadli, F., Meskin, N., Petri, I., Rezgui, Y., Bensaali, F., & Amira, A. (2023). AI-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives. *Artificial Intelligence Review*, *56*(6), 4929-5021. https://doi.org/10.1007/s10462-022-10286-2

Hossen, S., Mridha, Y., Rahman, A., Ouboucetta, R., & Amin, M. R. (2024). Consumer Perceptions And Purchasing Trends Of Eco-Friendly Textile Products In The US Market. *International Journal of Business and Economics*, *1*(2), 20-32. https://doi.org/10.62304/ijbm.v1i2.145

Jogesh, K. S., & Bappy, M. A. (2024). Machine Learning-Guided Design Of Nanolubricants For Minimizing Energy Loss In Mechanical Systems. *International Journal of Science and Engineering*, *1*(04), 1-16. https://doi.org/10.62304/ijse.v1i04.175

Khalil, A., & Belaissaoui, M. (2020). Cloudtech - New approach for implementing big datamart using NoSQL key-value stores. *2020 5th International Conference on Cloud Computing and Artificial Intelligence: Technologies and Applications (CloudTech)*, *NA*, 1-6. https://doi.org/10.1109/cloudtech49835.2020.9365897

Khan, M. A., & Walia, R. (2024). Intelligent Data Management in Cloud Using AI. *2024 3rd International Conference for Innovation in Technology (INOCON)*. https://doi.org/10.1109/inocon60754.2024.10511932

Luo, J., Xu, J., Aldosari, O., Althubiti, S. A., & Deebani, W. (2022). Design and Implementation of an Efficient Electronic Bank Management Information System Based Data Warehouse and Data Mining Processing. *Information Processing & Management*, *59*(6), 103086. https://doi.org/https://doi.org/10.1016/j.ipm.2022.103086

Manjunath, T. N., Pushpa, S. K., Hegadi, R. S., & Ananya Hathwar, K. S. (2023). A Study on Big Data Engineering Using Cloud Data Warehouse. *Data Engineering and Data Science*, 49-69. https://doi.org/10.1002/9781119841999.ch3

Maolmhuaidh, F. Ó., Thantilage, R. D., Burtenshaw, D., Raja, K., Le-Khac, N.-A., Cahill, P., Forster, R., & Kechadi, M. T. (2023). Clinical data warehouses for disease diagnosis: A focus on exosomal MicroRNA biomarkers. *Informatics in Medicine Unlocked*, *43*, 101390. https://doi.org/https://doi.org/10.1016/j.imu.2023.101390

Mia, M. R., Hoque, A. S. M. L., Khan, S. I., & Ahamed, S. I. (2022). A privacy-preserving National Clinical Data Warehouse: Architecture and analysis. *Smart Health*, *23*, 100238. https://doi.org/https://doi.org/10.1016/j.smhl.2021.100238

Moalla, I., Nabli, A., & Hammami, M. (2022). Data warehouse building to support opinion analysis in social media. *Social Network Analysis and Mining*, *12*(1), NA-NA. https://doi.org/10.1007/s13278-022-00960-2

Mora-Cantallops, M., Sánchez-Alonso, S., & García-Barriocanal, E. (2019). A systematic literature review on Wikidata. *Data Technologies and Applications*, *53*(3), 250-268. https://doi.org/10.1108/dta-12-2018-0110

Moscoso-Zea, O., Paredes-Gualtor, J., & Luján-Mora, S. (2018). A Holistic View of Data Warehousing in Education. *IEEE Access*, *6*, 64659-64673. https://doi.org/10.1109/access.2018.2876753

Naets, B., Raes, W., Deville, R., Middag, C., Stevens, N., & Minnaert, B. (2022). Artificial Intelligence for Smart Cities: Comparing Latency in Edge and Cloud Computing. *2022 IEEE European Technology and Engineering Management Summit (E-TEMS)*, *NA*, NA-NA. https://doi.org/10.1109/e-tems53558.2022.9944509

Doi: 10.69593/ajbais.v4i3.92

Oueslati, W., Tahri, S., Limam, H., & Akaichi, J. (2023). A systematic review on moving objects' trajectory data and trajectory data warehouse modeling. *Computer Science Review*, *47*, 100516. https://doi.org/https://doi.org/10.1016/j.cosrev.2022.100516

Passos, I. C., Ballester, P., Barros, R. C., Librenza-Garcia, D., Mwangi, B., Birmaher, B., Brietzke, E., Hajek, T., Jaramillo, C. L., Mansur, R. B., Alda, M., Haarman, B. C. M., Isometsä, E., Lam, R. W., McIntyre, R. S., Minuzzi, L., Kessing, L. V., Yatham, L. N., Duffy, A., & Kapczinski, F. (2019). Machine learning and big data analytics in bipolar disorder: A position paper from the International Society for Bipolar Disorders Big Data Task Force. *Bipolar disorders*, *21*(7), 582-594. https://doi.org/10.1111/bdi.12828

Rector, A., Schulz, S., Rodrigues, J. M., Chute, C. G., & Solbrig, H. (2019). On beyond Gruber: "Ontologies" in today's biomedical information systems and the limits of OWL. *Journal of biomedical informatics*, *100S*(NA), 100002-NA. https://doi.org/10.1016/j.yjbinx.2019.100002

Sah, B. P., Tanha, N. I., Sikder, M. A., & Habibullah, S. (2024). The Integration of Industry 4.0 And Lean Technologies In Manufacturing Industries: A Systematic Literature Review. *International Journal of Management Information Systems and Data Science*, *1*(3), 14-25. https://doi.org/10.62304/ijmisds.v1i3.164

Shamim, M. I. (2022). Exploring the success factors of project management. *American Journal of Economics and Business Management*, *5*(7), 64-72.

Swetha, K. R., Niranjanamurthy, M., P, A. M., & M, M. Y. (2021). Prediction of Pneumonia Using Big Data, Deep Learning and Machine Learning Techniques. *2021 6th International Conference on Communication and Electronics Systems (ICCES)*, *NA*, 1697-1700. https://doi.org/10.1109/icces51350.2021.9489188

Thantilage, R. D., Le-Khac, N.-A., & Kechadi, M. T. (2023). Healthcare data security and privacy in Data Warehouse architectures. *Informatics in Medicine Unlocked*, *39*, 101270. https://doi.org/https://doi.org/10.1016/j.imu.2023.101270

Tripathy, R. P., Mishra, M. R., & Dash, S. R. (2020). Next Generation Warehouse through disruptive IoT Blockchain. *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)*, *NA*, 1-6. https://doi.org/10.1109/iccsea49143.2020.9132906

Walters, K. M., Jojic, A., Pfaff, E. R., Rape, M., Spencer, D. C., Shaheen, N. J., Lamm, B., & Carey, T. S. (2021). Supporting research, protecting data: one institution's approach to clinical data warehouse governance. *Journal of the American Medical Informatics Association : JAMIA*, *29*(4), 707-712. https://doi.org/10.1093/jamia/ocab259

Younus, M., Hossen, S., & Islam, M. M. (2024). Advanced Business Analytics In Textile & Fashion Industries: Driving Innovation And Sustainable Growth. *International Journal of Management Information Systems and Data Science*, *1*(2), 37-47. https://doi.org/10.62304/ijmisds.v1i2.143

Younus, M., Pathan, S. H., Amin, M. R., Tania, I., & Ouboucetta, R. (2024). Sustainable fashion analytics: predicting the future of eco-friendly textile. *Global Mainstream Journal of Business, Economics, Development & Project Management*, *3*(03), 13-26. https://doi.org/10.62304/jbedpm.v3i03.85