Review

Not peer-reviewed version

Evaluating the Impact of Database and Data Warehouse Technologies on Organizational Performance: A Systematic Review

Nyiko G. Maswanganyi , Nyiko M. Fumani , Johannes K. Khoza , Bonginkosi A. Thango * , Lerato Matshaka

Posted Date: 2 October 2024

doi: 10.20944/preprints202410.0059.v1

Keywords: database technologies; data warehousing; organizational performance; innovation; customer satisfaction

Systematic Review

Evaluating the Impact of Database and Data Warehouse Technologies on Organizational Performance: A Systematic Review

Nyiko G. Maswanganyi [1], Nyiko M. Fumani [1], Johannes K. Khoza [1], Bonginkosi A. Thango [1,*] and Lerato Matshaka [2]

[1] Department of Electrical and Electronic Engineering Technology, University of Johannesburg, Johannesburg, South Africa, 2092
[2] Department of Nursing, Medical and Surgical Nursing, University of Johannesburg, Johannesburg, South Africa, 2092
* Correspondence: bonginkosit@uj.ac.za; Tel.: +27(0)11 559 6939

Abstract: In recent years of technological advancements, the digitization of information has become a crucial factor for the growth and sustainability of small and medium-sized enterprises (SMEs), particularly in developing countries. These enterprises face a growing problem in accessing IT resources due to financial and lack of qualified IT personnel in their regions. This led to a growing interest in database and data warehouse technologies, which serve as foundational tools for organizing, storing, and analyzing vast amounts of data. This systematic review proposes to evaluate the impact of database and data warehouse technologies on organizational performance by employing the 2020 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, this review thoroughly synthesizes existing literature to identify the key benefits and challenges associated with the adoption of these technologies in SMEs. The proposed inclusion criteria were bounded by (1) publication date between 2014 and 2024, (2) research written in English, (3) research work focusing on evaluating the impact of database and data warehouse technologies on organizational performance, (4) research with a clear evaluation analysis of database and data warehouse technologies research framework. Following these guidelines, 150 eligible research studies were included. The analysis reveals a wide range of findings concerning the impact of database and data warehouse technologies. Key results include the identification of cost-efficiency trends, improvements in decision-making processes, enhancements in customer relationship management, and gains in operational efficiency. The findings provide actionable guidance for SMEs, policymakers, and IT professionals seeking to leverage these technologies to enhance organizational performance. To the best of our knowledge, this systematic review offers a comprehensive analysis that fills a gap in the existing literature on the impact of these technologies on organizational outcomes.

Keywords: database technologies; data warehousing; organizational performance; innovation; customer satisfaction

1. Introduction

The digital revolution has fundamentally transformed the business landscape, with data emerging as a critical asset for organizations of all sizes. SMEs need database and data warehouse technologies to enhance their organizational performance. The more a business grows, the more their data gets complex, and the more need to have an effective system to organize the information, this will help with informed and effective decision making. Databases enable SMEs to store and retrieve data, ensuring that essential business data is accessible and usable in real time. Data warehouse systems improve these steps by integrating data from various sources, allowing businesses to conduct comprehensive analysis across departments. [1,2]. SMEs in developing face several challenges when adopting and implementing database and data warehouse technologies. The

 doi:10.20944/preprints202410.0059.v1

primary obstacle is financial constraints. These enterprises often operate on a limited capital, making it difficult to invest in hardware and software required for effective data management. Skilled IT personnel in many developing countries are few, further complicating the adoption of these technologies. Outsourcing IT personnel is not an option as well which further strains their limited budget operation. [3–7]. This skills gap makes it harder for SMEs to take full advantage of the potential benefits offered by these technologies.

By automating data collection, storage, and retrieval, database and data warehouse technologies can significantly reduce operational costs and improve efficiency, addressing the financial constraints many faces. Database and data warehouse systems are increasingly designed with user-friendly interfaces, reducing the need for extensive technical knowledge and ultimately the need for expertise. [8–10]. This allows SMEs to manage their data more effectively even with limited IT resources. Some systems offer automated maintenance and support, further alleviating the burden of managing complex IT infrastructure. Databases allow real-time access to data, and data warehouses support cross department adaptation. [11,12]. This supports strategic planning, enabling businesses to identify trends, optimize processes, and improve customer engagement. In turn, these insights lead to better resource allocation, improved productivity, and the ability to make data-driven decisions that foster growth and competitiveness, despite the challenges of smaller workforce and limited financial resources.

1.1. Evolution of Database Technologies

The comprehensive evolution of Database technologies is illustrated in Figure 1. Database evolution began with the first wave (1960-1999), which included early database systems such as network, hierarchical, and inverted list databases, all designed to handle vast amounts of data. By the 1990s, the introduction of Object-Oriented Database Management Systems (OODBMS) enabled data modeling as objects, aligning with object-oriented programming and improving the processing of complex data types like multimedia.

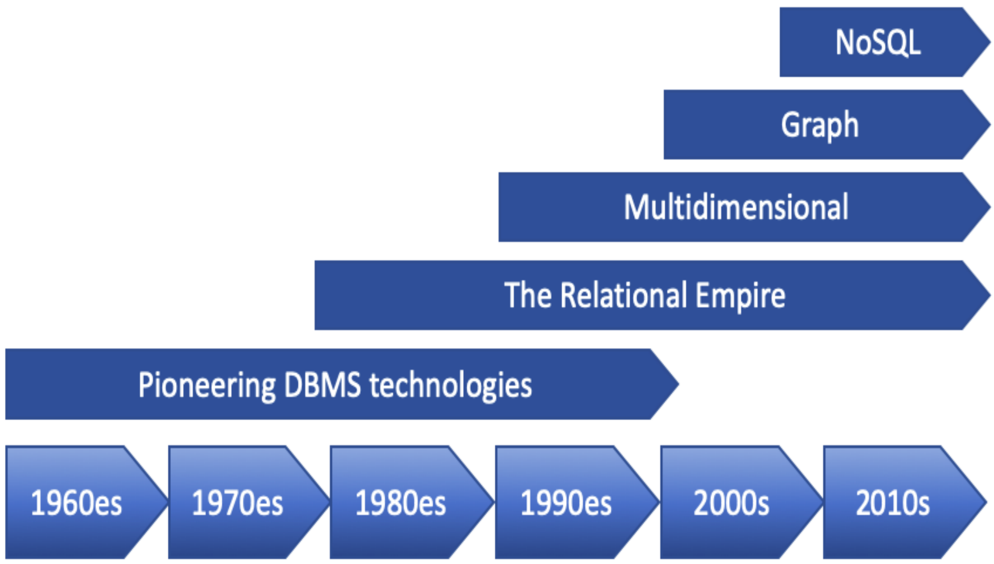


Figure 1. Comprehensive evolution of Databases.

This time established the foundation for current databases. Following this, the relational wave (1990-2008) saw a shift with the introduction of Relational Database Management Systems (RDBMS), which are based on E.F. Codd's relational paradigm. Despite initial problems, the usage of Structured Query Language (SQL) and developments such as Patricia Selinger's cost-based query optimization made relational databases very scalable and dependable, making them the backbone for most corporate applications in data modelling and performance. [13–15]

 doi:10.20944/preprints202410.0059.v1

Simultaneously, the decision support wave (1990-present) began with the introduction of Online Analytical Processing (OLAP) and data warehouses designed for complicated data analysis and corporate intelligence. OLAP systems, with their multidimensional data modeling capabilities, enabled firms to do in-depth historical data analysis. The emergence of graph databases during the graph wave (1999-present) transformed how data relationships were managed, with property graph databases such as Neo4j excelling at social networks and fraud detection. Finally, the NoSQL wave (2008-present) emerged to meet the demand for scalable, flexible databases capable of managing unstructured and semi-structured data produced by big data applications. NoSQL databases, which provide schema flexibility and performance, have become vital for real-time analytics, IoT, and large-scale online platforms, representing a shift from the inflexible structure of relational databases. [16–18]

1.2. Evolution of Data Warehouse Technologies

In Figure 2, evolution of data warehouse technologies is illustrated. The evolution of data warehousing has gone through several stages, driven by the exponential growth of data and advancements in technology. Initially, traditional data warehouses were built on Relational Database Management Systems (RDBMS), like Oracle and SQL Server, and utilized fact and dimension tables to organize data.

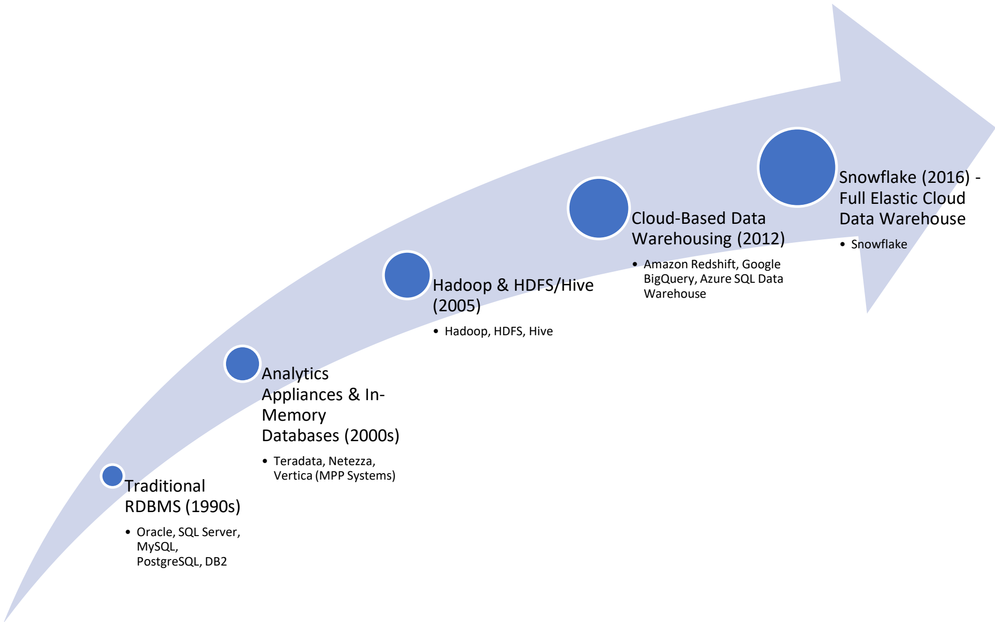


Figure 2. Comprehensive Evolution of Data Warehouse Technologies.

These early systems could handle modest amounts of data but struggled as the volume and complexity of customer and sensor data grew in the early 2000s. To address the challenges of performance degradation and data processing limitations, appliance-based solutions like Teradata and Netezza, which utilized Massively Parallel Processing (MPP), emerged. These solutions offered faster data processing capabilities but were costly and lacked scalability. With the rise of big data, Hadoop and its ecosystem (HDFS, Hive) became prominent in 2005, introducing open-source solutions for massive data scaling. However, Hadoop's complexity and steep learning curve limited its adoption by businesses with traditional IT infrastructures. In 2012, cloud computing brought the next significant advancement, with services like Amazon Redshift, Google Bigquery, and Microsoft Azure Data Warehouse enabling businesses to leverage cloud scalability. While these solutions offered greater flexibility, they were still rooted in older technologies and had limitations in elasticity. The most recent phase in data warehousing evolution is represented by fully cloud-native platforms like Snowflake (introduced in 2016), which were designed specifically for the cloud. Snowflake's architecture separates compute and storage, allowing companies to scale resources independently and pay only for what they use. Additionally, it offers unique features such as cloning data for testing without duplicating physical storage, and its metadata-driven architecture supports real-time elasticity. This evolution has drastically lowered the Total Cost of Ownership (TCO), making it easier for businesses of all sizes to access, analyze, and derive value from their data while minimizing maintenance and infrastructure costs. [19–21]

The proposed systematic review facilitates the consolidation of existing research on the impact of database and data warehouse technologies on organizational performance. By focusing on various types of databases and data warehouses, the review seeks to provide a comprehensive understanding of how these technologies can be leveraged to enhance organizational efficiency. It will assess different technologies, their cost implications, effects on decision-making, customer relationship management, and operational efficiency. By identifying gaps in the current literature and highlighting areas that require further investigation as shown in Table 1 this review is expected to drive innovation and development in the field. Ultimately, the systematic review will serve as a foundational document, offering researchers a valuable resource to build upon for advancing the understanding and application of database and data warehouse technologies in improving organizational performance.

Table 1. Comparative analysis of the existing review works and proposed systematic review on database and data warehousing impact on organizational performance.

Ref.	Cites	Year	Contribution	Pros	Cons
[22]	174	2016	Review of LCA databases for construction materials	Structured selection approach	Geographical mismatch and incomplete data
[23]	3	2020	Systematic review on project management in data warehouse (DWH) implementations	Comprehensive reference for project managers	Scarcity of literature on PM in DWH
[24]	136	2017	Comprehensive review of Business Intelligence (BI) literature	Structured overview of BI literature.	Review period ends in 2015, missing recent developments.
[25]	1	2024	Review on AI's impact on organizational justice and project performance	Thorough analysis of AI in the context	Limited by a small number of reviewed articles
[26]	13	2023	Systematic review on AI's impact on student performance	Empirical evidence of AI's positive effects in STEM	Challenges of AI in education need further exploration
[28]	27	2022	Review on factors influencing AI adoption in healthcare	Highlights key psychosocial factors	Limited focus on patient demographics
[29]	69	2017	Two-cloud secure database architecture for privacy in SQL queries.	Robust privacy-preserving mechanism	Complexities in managing two clouds.
[30]	11	2021	Review of multi-tenancy scheduling in cloud platforms.	Comprehensive overview and challenges	Broad focus may dilute specific insights.
[31]	179	2020	Review of deep learning techniques for 3D point clouds.	Comprehensive overview and structure for learning	May exclude foundational methods.
[32]	2	2021	Review on challenges of traditional storage systems and NoSQL databases.	Comprehensive analysis of NoSQL technologies	Lacks in-depth case studies or practical examples

[33]	7	2022	Review on digitalization trends in patent information databases.	Comprehensive overview of trends	Lacks practical case studies
[34]	18	2022	Review on cloud-based knowledge management in SMEs	Highlights benefits of cloud computing for SMEs.	Limited to specific databases.
Proposed systematic review			Evaluates the impact of business intelligence on SME performance, highlighting benefits such as improved decision-making, competitive advantages, and operational efficiency.	Provides a comprehensive understanding of factors influencing BI adoption in SMEs. Identifies critical research gaps.	Limited focus on industry-specific applications and geographic limitations.

1.3. Research Questions

In this research, we evaluate the impact of Database and Data Warehouse Technologies on organizational performance of small and medium enterprises (SME)s. This study inquires intensively on how Database and Data Warehouse Technologies can improve SMEs' efficiency and performance, highlighting both the advantages and challenges faced in its employment. The following research questions aim to clarify these aspects to comprehend the effective use of Database and Data Warehouse Technologies in SMEs' organizational performance.

- How does the utilization of Database and Data Warehouse technologies impact SMEs' performance?
- What are the challenges SMEs face when adopting Database and Data Warehouse technologies?
- What specific aspects of organizational performance are looked at in relation to Database and Data Warehouse technologies?
- What are the crucial factors influence successful adoption of Database and Data Warehouse technologies?
- What are the consequences of not utilizing Database and Data Warehouse technologies?

1.4. Rationale

The fast evolution of data digitization has presented an opportunity for businesses to stay competitive, relevant, and in touch with their customers in real time. These volumes of data are unusable unless there are effective tools that will identify, store, and analyze this information. This is a huge obstacle for SMEs, which are an important part of the economy but lack the resources, awareness, and expertise needed to successfully integrate Database and Data Warehouse Technologies. Research highlights how the utilization of DB and DWT has the necessary infrastructure to boost decision-making processes, competitiveness, and operational efficiency. However, it identifies several challenges faced by SMEs, such as lack of awareness, no information management personnel, and lack of capital that hinders the investment and adoption of DB and DWT.

Present research concentrates on one side which is often the adoption of DB and DWT, disregarding their comprehensive influence on crucial organizational metrics such as financial stability, operational effectiveness, and creativity. By evaluating the impact of DB and DWT on organizational performance, we aim to provide an understanding of how SMEs can utilize DB and DWT to improve efficiency and performance. This paper will add to how decision makers can invest in DB and DWT and boost SME success.

1.5. Research Contribution

This work presents a comprehensive systematic review of the impact of database and data warehouse technologies on organizational performance. The study addresses various unresolved issues and research gaps in this field. The key contributions of this research are as follows:

 doi:10.20944/preprints202410.0059.v1

- We provide a comprehensive analysis of the influence that database and data warehouse technologies have on organizational operations. This evaluation explores the effects of these technologies on decision-making processes, operational efficiency, and the overall performance of the organization. It offers valuable insights for companies seeking to leverage data management solutions to enhance productivity.
- We gather and analyze the existing knowledge concerning the impacts of data warehouse and database technologies. By highlighting deficiencies in the current literature, particularly in terms of their effectiveness and practical use across different organizational contexts, we pinpoint areas that require further investigation. This approach aims to stimulate advancements in the field and direct future research initiatives.
- We present a range of advanced analytical models to assess the impact of database and data warehouse technologies on organizational performance. These models aim to establish a foundation for future empirical research in this area and to provide a deeper insight into the diverse factors that influence organizational success.

1.6. Research Novelty

This paper presents a unique perspective to the field, according to our thorough investigation, there is currently little to no study that delves on the influence of both DB and DWT on SME performance with specific focus on factors enabling or barricading its adoption and possible long-term impacts. Our systematic review tackles gaps by identifying the key contributing factors and drawbacks to the implication of DB and DWT in SMEs. We explore the often-overlooked areas such as lack of expertise, lack of capital, and the technological readiness of SMEs, which hinder their ability to harness DB and DWT and their full potential. Through this work, we provide specific insights and practical recommendations designed for SMEs, propelling them to integrate Database and Data Warehouse Technologies to improve operational performance and efficiency.

1.7. Research Organization

The paper is structured to provide a clear and systematic examination of the impact of Database and Data Warehouse Technologies on the performance of SMEs. Following this introductory section, Section 2 outlines the materials and methods employed in the research, detailing the systematic review process, the criteria for study inclusion, and the data collection methods used. This section also explains the search strategy applied to gather relevant literature, ensuring a comprehensive review of the subject. In Section 3, the results of the review are presented, synthesizing the key findings from the studies analyzed. This section focuses on the core factors influencing the adoption of Database and Data Warehouse Technologies by SMEs, highlighting both the enablers and barriers to implementation, and providing insights into the impact of these technologies on operational performance. The Discussion follows in Section 4, where the implications of the findings are explored. This section interprets the results in the context of SME performance, drawing comparisons with existing literature and discussing the practical challenges and opportunities that SMEs face in adopting these technologies. Additionally, it offers an analysis of how the findings contribute to the broader understanding of technology adoption in small and medium enterprises.

2. Materials and Methods

This section outlines the procedure followed to carry out the review of the evaluation of the impact of Database and Data Warehouse Technologies on Organizational Performance of SMEs. The review investigates studies within the past decade (2014 – 2024). As shown in Figure 3, the methodology highlights guidelines for study selection, data collection process and origin, and the method used to evaluate the collected literature laying the foundation for a detailed investigation of each aspect in the following sections.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

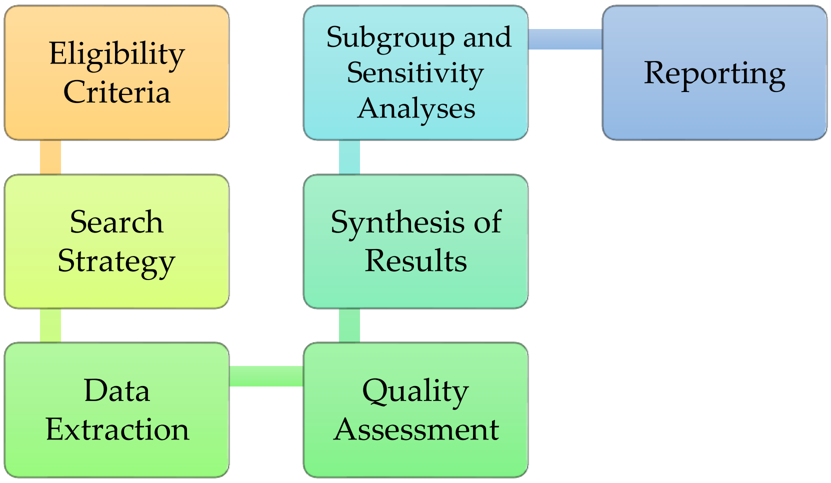


Figure 3. SLR Procedure Block Diagram.

2.1. Eligibility Criteria

A systematic study of all peer-reviewed and published research works appropriate to the effects of database and data warehouse technologies on organizational performance was the focus of a thorough systematic evaluation that was carried out. The review covered studies that were written up in English and released between 2014 and 2024. To ensure that pertinent studies were included and to weed out those that did not directly address the evaluation of the impact of database and data warehouse technologies on organizational performance, a definitive set of inclusion and exclusion criteria was developed. As such, the review considered only peer-reviewed research that were exclusively focused on this subject. Table 2 contains specifics about the study's inclusion and exclusion criteria.

Table 2. Proposed inclusion and exclusion criteria.

Criteria	Inclusion	Exclusion
Topic	Research papers focusing on Evaluating the Impact of Database and Data Warehouse Technologies on Organizational Performance: A Systematic Review	Research papers not focusing on Evaluating the Impact of Database and Data Warehouse Technologies on Organizational Performance: A Systematic Review
Research framework	The articles must include a research framework or methodology for database and data warehouse technologies on organizational performance.	Articles lacking a clear research framework or methodology for database and data warehouse technologies on organizational performance.
Language	Papers written in English	Papers not in English
Period	Publications within 2010 and 2024	Publications outside 2010 and 2024

The assessment of studies for inclusion in the systematic review followed a structured, multi-stage process. In the first stage, titles and abstracts were screened by two independent reviewers to determine if they met the inclusion criteria. Any discrepancies in judgment were discussed between the reviewers to ensure consistency. In the second stage, full-text versions of the remaining studies were thoroughly reviewed to assess their adherence to the inclusion criteria, particularly focusing on the presence of a clear research framework or methodology related to the impact of database and data warehouse technologies on organizational performance. Studies that lacked sufficient methodological detail or relevance were excluded at this stage. Finally, in the third stage, any

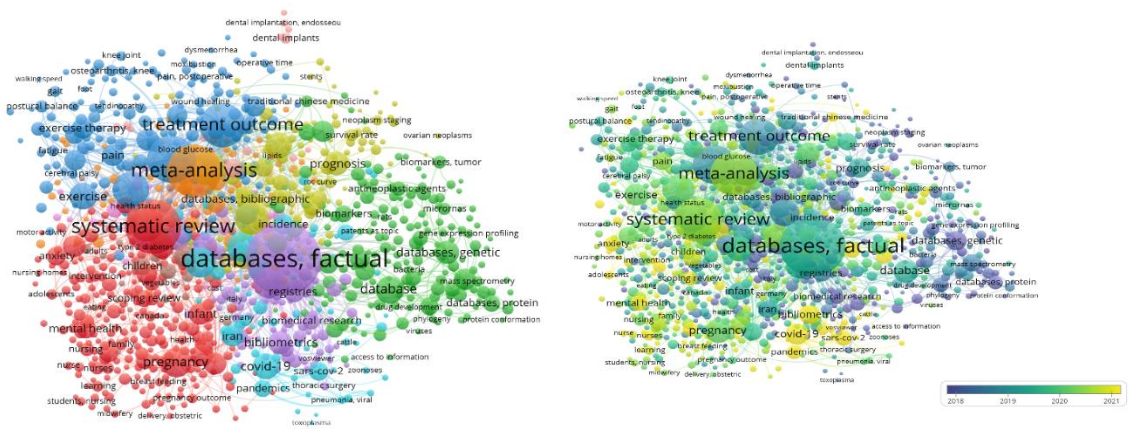
Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

borderline cases were brought to a consensus discussion involving all reviewers. If disagreements remained unresolved, a third reviewer was consulted, and decisions were made based on a majority vote. This rigorous assessment process ensured that only studies of high relevance and methodological rigor were included in the review.

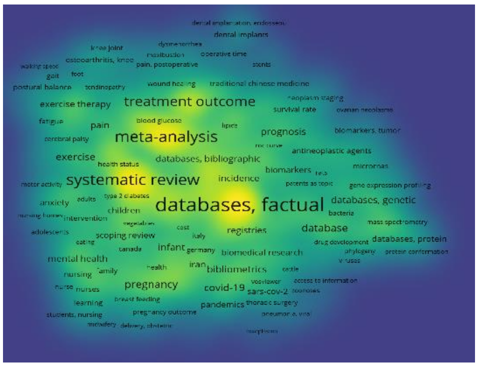
2.2. Data collection Process

The searches were conducted across multiple academic databases to ensure a broad and representative collection of studies. These databases included Scopus, Web of Science, and Google Scholar, which were chosen for their robust coverage of scientific, technical, and business literature. The choice of these databases was based on their relevance to the fields of technology, business, and organizational studies, ensuring access to high-quality and peer-reviewed research. The selected synonyms and keyword combinations were applied universally across all databases. However, slight adjustments were made in cases where the database's search engine required specific formatting or syntax. For instance, in SCOPUS, the keyword “NoSQL databases” was expanded with the phrase “OR non-relational databases” to accommodate specific database architecture discussions.

The search keywords used to conduct this systematic review were formulated to pinpoint studies that presented attributes of Database and Data Warehouse Technologies on Organizational Performance. To select index key-terms, experimental searches in a reiterative method were carried out. Keywords that did not yield research works aligned to the inclusion criteria were excluded. To ensure the inclusion of analogous terms for the fundamental keywords, synonymous terms were established for each key concept. For database technologies, synonyms included: "database technologies," "database systems," "database management systems," "DBMS," "relational databases," "NoSQL databases," and "SQL databases. Keywords related to data warehouse technologies included: "data warehousing," "data warehouse systems," "data marts," "enterprise data warehouses," "DW," and "OLAP systems." The term organizational performance was described using synonyms such as: "business performance," "company performance," "enterprise performance," "operational efficiency," "organizational efficiency," "organizational effectiveness," "business success," and "firm performance." During the search procedure, logic operators "AND" and "OR" were utilized to refine the search for relevant scholarly papers. The "AND" operator ensured that all chosen keywords were included, while the "OR" operator captured papers containing any of the selected keywords. Additionally, the wildcard asterisk (*) was employed to capture plurals and other suffix variations. This methodical search strategy ensured comprehensive coverage of the relevant literature for evaluating the impact of data-base and data warehouse technologies on organizational performance. The bibliometric analysis of study search Keywords i.e., the network visualization, overlay visualization and density visualization is illustrated in Figure 4.



(**a**) (**b**)



(c)

Figure 4. Bibliometric Analysis of Study search Keywords: (a) Network Visualization. (b) Overlay Visualization. (c) Density Visualization.

Table 3 shows sample keyword combinations and their respective results in the different databases.

Table 3. Search Keyword combination result summary.

Keyword Combination	Database	Results
"Database technologies" AND "organizational performance"	SCOPUS	87
"Data warehousing" OR "data marts" AND "company performance"	Web of Science	293
"NoSQL databases" AND "business success"	Google Scholar	576
"DBMS" OR "relational databases" AND "operational efficiency"	Google Scholar	623
"SQL databases" AND "organizational efficiency"	SCOPUS	210
"Enterprise data warehouses" AND "firm performance"	Web of Science	350
"OLAP systems" AND "business performance"	Google Scholar	913
"Database management systems" AND "organizational effectiveness	Google Scholar	1469

Table 3 provides transparency in the search process, illustrating how the various synonyms and keyword combinations were applied and the number of relevant studies retrieved from each search query.

2.3. Conceptualization of Data Warehouse

The concept of data warehousing emerged in the early 1990s as organizations began accumulating large volumes of data from various sources. A data warehouse is a centralized repository designed to store historical and current data from multiple disparate systems, enabling organizations to perform advanced analytics and derive insights for strategic decision-making. Data warehouses are structured to allow efficient querying and reporting, often integrating with business intelligence (BI) tools. Over time, data warehousing technology has evolved to accommodate the growing need for scalability, flexibility, and real-time processing capabilities, particularly as businesses have moved towards more dynamic data environments.

2.3.1. Types of Data Warehouses.

Data warehouses can be classified into several types, each catering to specific organizational needs as follows.

a) Enterprise Data Warehouse (EDW)

As shown in Figure 4, an EDW is a comprehensive, centralized data warehouse designed to store and manage data from across an entire organization. It provides a unified view of the organization's data, facilitating enterprise-wide analytics and reporting. EDWs are typically used in large

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

organizations with complex data needs, and they are optimized for complex queries, batch processing, and reporting.

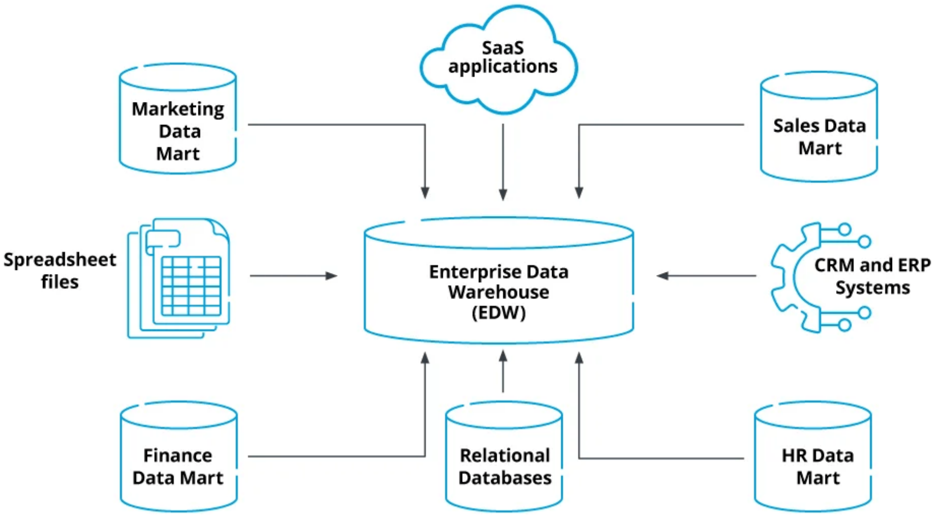


Figure 4. Enterprise Data Warehouse (EDW), [Available at: LINK].

b) Operational Data Store (ODS)

as shown in Figure 5, an ODS is a type of data repository that focuses on storing real-time or near-real-time operational data. Unlike EDWs, which typically deals with historical data, ODS is designed to support daily operational activities and decision-making by providing up-to-date information. An ODS is often a precursor to a full-fledged data warehouse, serving as a temporary storage area for data before it is loaded into the data warehouse.

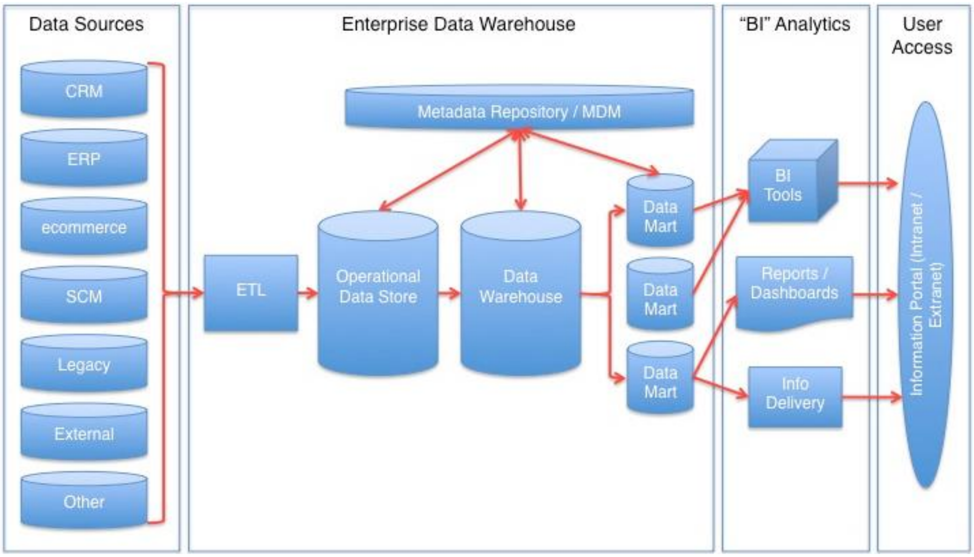


Figure 5. Operational Data Store (ODS), [Available at: LINK].

c) Data Mart

A data mart is a subset of a data warehouse, designed to focus on specific business areas or departments such as finance, marketing, or sales. As shown in Figure 6, data marts are smaller and more specialized than EDWs, allowing faster access to relevant information for specific decision-makers. They can either be dependent (created from an EDW) or independent (created from external sources).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

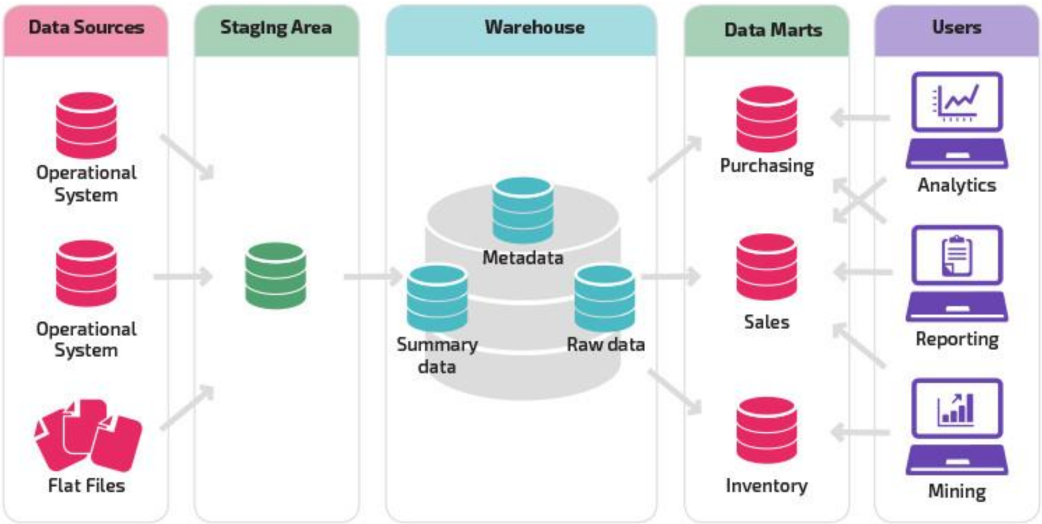


Figure 6. Data Mart, [Available at: LINK].

d) Data Lake

A modern approach to data storage, data lakes are often compared to data warehouses but differ in that they store both structured and unstructured data in raw form. While a data warehouse typically imposes a schema on data at the time of storage, a data lake applies a schema on read. Data lakes, as shown in Figure 7, offer more flexibility in terms of storage and processing but require more advanced data management tools to maintain.

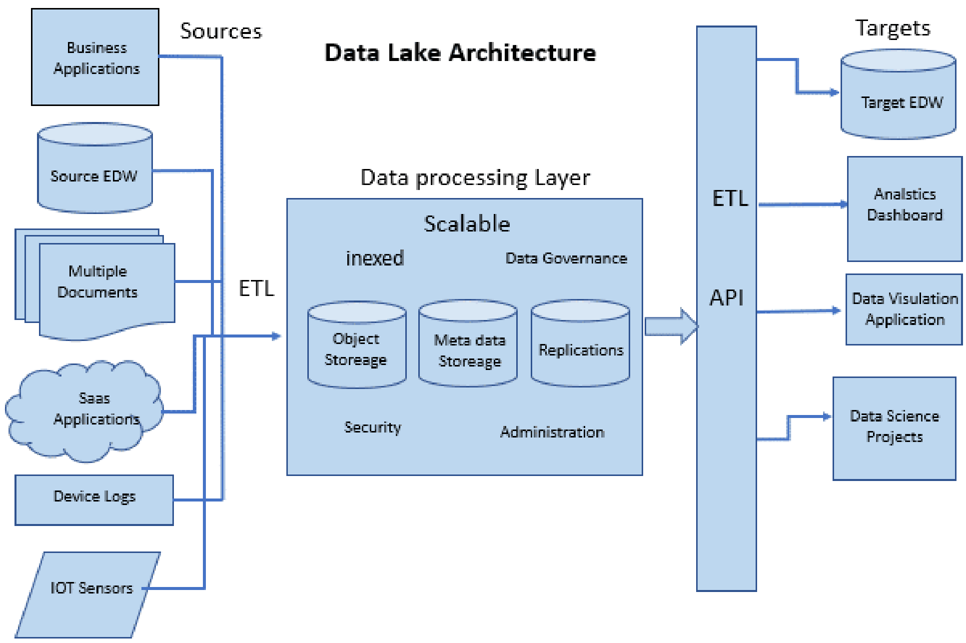


Figure 7. Data Mart, [Available at: LINK].

2.3.3. Modern Data Warehouse Architecture

The architecture of data warehouses has evolved significantly in recent years, driven by advances in cloud computing, distributed storage, and big data technologies. Modern data warehouses are designed to handle larger volumes of data, offer greater flexibility in querying, and support advanced analytics such as machine learning. With the growth of cloud computing, many organizations are moving their data warehouses to the cloud. Cloud-based data warehouses, such as Amazon Redshift, Google BigQuery, and Snowflake, provide scalability, flexibility, and cost-

 doi:10.20944/preprints202410.0059.v1

efficiency. These platforms can scale up or down based on demand and provide advanced querying capabilities, enabling real-time insights and analytics. Modern data warehouses often use distributed architecture, which spreads data across multiple nodes or servers to improve performance and reliability. This architecture allows data warehouses to handle massive amounts of data and provides fault tolerance by distributing processing tasks across different systems. A newer approach to data warehousing involves data virtualization, where data from various sources can be accessed in real-time without needing to move or replicate it. This reduces data redundancy and enables faster query responses, especially when dealing with dynamic and disparate data sources.

2.4. Big Data Frameworks

Big data frameworks have become an essential part of modern data analytics, enabling organizations to process, analyze, and derive insights from vast amounts of structured and unstructured data. These frameworks, as shown in Table 4, provide the necessary tools to handle complex data operations and integrate data from diverse sources at scale. The rise of big data frameworks has been closely tied to the growth of cloud computing, distributed systems, and the demand for real-time analytics.

Table 4. Big Data Frameworks.

	Hive	Presto	Spark	BigQuery
Perfomance	Slow	Fast	Fast	Ultra Fast (using many disks)
Intermediate Storage	HDFS	None	Memory/ Disk	Colossus (?)
Data Transfer	HTTP	HTTP	HTTP	?
Query Execution	Stage-wize MapReduce	Run all stages at once (pipelining)	Stage-wise	?
Fault Tolerance	Yes	None (but, TD will retry the query)	Yes, but limited	?
Multiple Job Support	Good Can handle many jobs	Limited (- 5 concurrent queries per account in TD)	Require another resource manager (e.g. YARN, mesos)	Limited (Query queue)

2.4.1. Popular Big Data Frameworks

Figure 8 shows how Big Data frameworks have evolved to address the complex demands of managing and processing large-scale datasets through various specialized systems. Big data frameworks are essential for managing and analyzing vast datasets, each offering unique capabilities depending on the specific needs of the organization. Several widely used frameworks have emerged to address different aspects of big data processing, from batch processing to real-time analytics. Apache Hadoop stands as one of the foundational frameworks, offering a distributed storage and processing system with its core components, the Hadoop Distributed File System (HDFS) and the MapReduce programming model. This framework excels in handling vast amounts of data across multiple machines, providing scalability and fault tolerance essential for batch processing.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

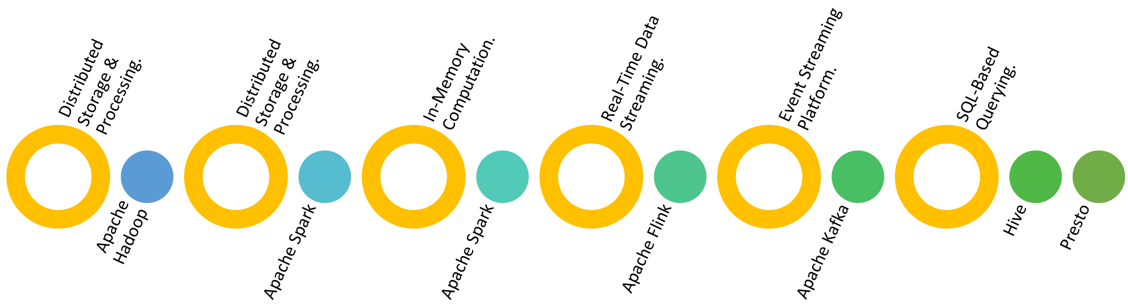


Figure 8. Types of Big Data Frameworks.

Apache Spark builds upon this by introducing in-memory computation, significantly enhancing processing speed for iterative algorithms. It supports a broad range of workloads, including batch processing, real-time streaming, and machine learning, making it a versatile choice for big data analytics. Apache Flink extends this versatility into real-time data streaming, with a focus on delivering low-latency, real-time analytics, which is crucial for applications such as fraud detection and Internet of Things (IoT) data analysis. Complementing these frameworks is Apache Kafka, a distributed event streaming platform designed to handle high-throughput, low-latency data streams, and facilitate real-time data processing across various systems. For SQL-based querying, Hive provides a SQL-like interface on top of Hadoop, making it easier to work with large datasets using familiar syntax, while Presto offers a distributed SQL query engine that enables fast, interactive querying of big data. Together, these frameworks represent a spectrum of solutions tailored to different aspects of big data processing, from storage and batch processing to real-time analytics and querying, addressing the diverse needs of modern data management.

2.2. Taxonomy of Data

The taxonomy of big data refers to the classification of big data into various categories based on its characteristics and how it is managed. Understanding this taxonomy is crucial for developing effective big data strategies and choosing the right tools for data processing and analysis.

2.5.1. The 4 Vs of Big Data

Figure 9 offers a graphical representation of the 4 Vs of Big Data that provide a comprehensive framework for understanding the key challenges and opportunities associated with big data. Volume refers to the massive amounts of data generated daily, as organizations are inundated with data from multiple sources, including transactional data, social media interactions, IoT devices, and more. Managing the scale and storage of this data is a significant challenge for modern enterprises. The velocity of data generation and processing is critical.

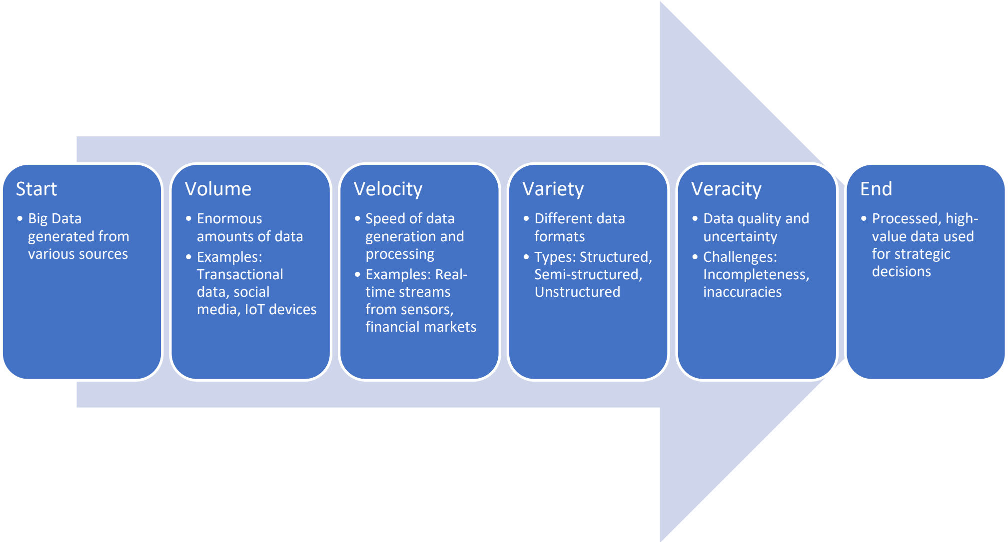


Figure 9. The 4 Vs process.

With real-time data streams flowing from sources like sensors, financial markets, and social media, organizations need systems that can handle this fast-paced data influx and generate real-time insights to stay competitive. Another key element is variety, which refers to the diverse formats of big data. Data can be structured, such as traditional relational databases, semi-structured, like XML and JSON files, or unstructured, such as multimedia content, including images, videos, and text. Managing and integrating these different data formats within a single framework is a complex task but necessary for comprehensive data analysis. Along with the variety comes veracity, which deals with the accuracy and quality of the data. Big data is often incomplete, inconsistent, or incorrect, requiring organizations to adopt data cleansing and validation processes to ensure the data's reliability and utility. The 4 Vs framework encapsulates the multifaceted challenges of handling big data while emphasizing the potential value that can be unlocked with the right tools and strategies.

2.5.2. Big Data Management Models.

Traditionally, big data processing has relied on batch processing, where data is collected over a set period and then processed all at once. This method is particularly efficient for handling large datasets, as it allows organizations to process vast amounts of information in a single operation. However, the drawback of batch processing is that it does not provide real-time insights, which can be a limitation for industries or applications that need immediate data responses. In contrast, stream processing has emerged as a solution for real-time data analysis, enabling organizations to process data continuously as it is generated. This approach is especially critical for applications requiring instantaneous decision-making, such as in financial markets, where milliseconds can impact transactions, or in Internet of Things (IoT) systems that rely on immediate feedback to adjust operations. Modern big data architectures often adopt hybrid models, combining both batch and stream processing to leverage the strengths of each approach. This allows businesses to gain real-time insights from streaming data while also efficiently processing historical data in batches, ensuring both immediate responsiveness and comprehensive data analysis.

2.2. Data Items

In this paper, the focus is on gathering information that intensively investigates the effects of Database and Data Warehouse Technologies on small and medium-sized enterprises (SMEs). The outcomes are grouped in accordance with the research questions to avoid straying from topic. This systematic review evaluates the impact of database and data warehouse technologies on organizational performance by examining a variety of outcome dimensions, including operational, financial, innovation, collaboration, employee, and customer outcomes. A comprehensive search

 doi:10.20944/preprints202410.0059.v1

strategy is employed to identify relevant studies, including both qualitative and quantitative research. Data extraction is performed independently by two reviewers, with quality assessed using the Newcastle-Ottawa Scale and GRADE framework. A narrative synthesis and, where applicable, meta-analysis are conducted to summarize findings. Subgroup and sensitivity analyses are performed to explore variations across factors such as industry and company size. The review follows the PRISMA framework for transparent reporting.

2.6.1. Variables

This section outlines the general characteristics of the studies and contributors covered in the evaluation. Table 5 describes essential elements such as the publication year, the databases used, and the types of studies included in the analysis. The goal is to thoroughly define and evaluate key variables, including research design, techniques, and other relevant study aspects. The studies examined provide insight into how database and data warehouse technologies influence organizational outcomes and performance metrics. In studies like "The Impact of Cloud-Based Database Systems on SME Performance: An Exploratory Case Study" (2019) and "The Role of Data Warehousing in Enhancing Business Decision-Making in Developing Economies" (2023) used a combination of qualitative and quantitative methods, such as case studies and surveys, to explore the relationship between data warehousing technologies and business outcomes like operational efficiency and revenue growth. Another study, "Adoption of Hybrid Cloud Data Warehousing for Competitive Advantage in Startups" (2022), employed mixed-method approaches, integrating statistical analysis and thematic analysis, to evaluate the benefits of cloud-based and on-premises technologies in small businesses. Some research focused on specific technological implementations, such as the study "ETL Tools and their Impact on Data Quality and Performance Metrics in Large Enterprises" (2021), which employed experimental designs to test the efficacy of various data processing tools. This study used a combination of interviews and document analysis to evaluate outcomes related to data accuracy, query performance, and scalability.

Table 5. Important Terms.

Heading	Explanation	Selection
Title	The title of the study or paper	None
Year	The year the study was published	2014, 2015, 2016…2024
Online database	Database where the study is available	Google Scholar, SCOPUS, Web of Science
Journal name	The name of the journal or conference where the study was published.	None
Research type	The type of research publication	Journal, Article, Conference Paper, Book Chapter, Thesis, Dissertation
Discipline or subject area	The field of study related to the research.	Database Technologies, Data Warehouse, SME Performance, Business
Industry context	The industry focus of the study	SME, Startup, Small Businesses
Geographic location	The country or region the study was conducted.	America, Europe, Asia, Africa
Economic context	The economic setting of the study.	Developed Country, Developing Country
Types of database technologies	The types of database technologies used in the study.	Cloud Databases, In-memory Databases, Relational Databases, NoSQL Databases, Distributed Databases
Types of data warehouse technologies	The specific data warehouse technologies mentioned in the study.	ETL Tools, OLAP, Data Marts, Data Lakes
Technology providers	The companies providing the technologies in the study.	Oracle, Microsoft SQL Server, IBM, AWS Redshift, Google Big Query, SAP
Technology implementation model	The deployment model for the technology.	On-premises, Cloud-based, Hybrid
Research design	The methodological approach of the study.	Experimental, Quasi-experimental, Case Study, Survey, Mixed Methods
Type of study	The study approach.	Quantitative, Qualitative, Mixed Methods
Sample size	Number of participants or data points in the study.	None

Sample characteristics	The key traits of the study participants.	SME, Data Managers, IT Professionals, Business Owners
Data collection methods	The methods used to gather data.	Interviews, Surveys, Observations, Document Analysis
Data analysis technique	The techniques used to analyze the collected data.	Quantitative Analysis, Thematic Analysis, Statistical Analysis
IT Performance metrics	The metrics used to evaluate IT performance.	Query Performance, Data Accuracy, Scalability, Latency, Uptime
Business performance metrics	The metrics used to measure business outcomes.	Operational Efficiency, Revenue Growth, Cost Savings, Time to Market
Organizational outcomes	The broader organizational outcomes observed in the study.	Employee Satisfaction, Customer Satisfaction, Innovation, Collaboration
Long term impacts	The long-term impacts or benefits of the technologies on the organization.	Competitive Advantage, Market Share Growth, Long-term Cost Reduction

The characteristics outlined in table 5 provide a comprehensive overview of the studies included in the review, detailing the factors that influence the effectiveness of database and data warehouse technologies in driving performance outcomes. This structured representation allows for a focused analysis of key variables across various studies.

2.2. Study Risk of Bias Assessment

The quality appraisal aimed to enhance the veracity of the selected research papers and establish the suitability and comprehensiveness of their findings. A total of 150 studies included in this review were evaluated for quality using a scoring technique based on the Newcastle-Ottawa Scale (NOS), which assesses studies across three dimensions: Selection, Comparability, and Outcome/Exposure. To ensure transparency in the assessment process, a flowchart process was developed to visually illustrate the steps taken during the bias assessment. Figure 10 outlines how studies were selected, evaluated, and scored, providing an easily understandable overview. The combined approach of multiple independent reviewers, a structured scoring system, and visual representation helped ensure that the risk of bias assessment was comprehensive and objective, thus supporting the reliability and validity of the conclusions drawn from the reviewed studies.

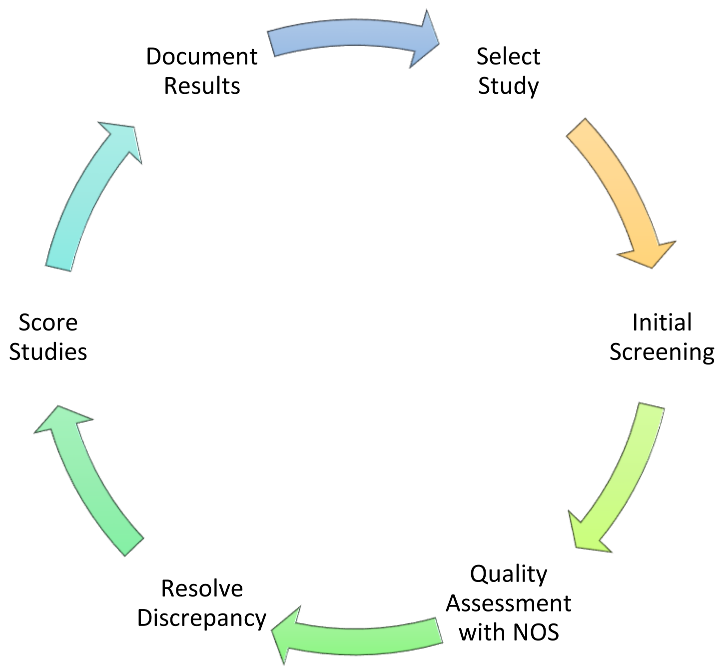


Figure 10. Study Risk of Bias Assessment flow diagram.

This systematic approach provided a consistent measure of the studies' reliability and robustness. The NOS allows for a maximum of nine stars across these three dimensions, with the sum

 doi:10.20944/preprints202410.0059.v1

of stars determining the overall quality rating of each study as High (7-9 stars), Moderate (4-6 stars), or Low Quality (0-3 stars). By applying the NOS, the assessment was designed to ensure objectivity and minimize bias, thus reinforcing the reliability of the findings. Two independent reviewers conducted assessments for all studies, facilitating inter-rater reliability. In cases of discrepancies, a third impartial reviewer mediated to resolve differences, ensuring that subjective judgments were collectively discussed and agreed upon. This collaborative approach further enhanced the credibility of the bias assessment.

The Selection dimension of the Newcastle-Ottawa Scale, contributing up to four stars, evaluates how well studies select participants, focusing on whether cases and controls in case-control studies are drawn from the same population to minimize selection bias. In cohort studies, it examines the accurate representation of exposed and unexposed groups. The Comparability dimension, which can contribute up to two stars, assesses how effectively studies control for confounding factors, with higher scores awarded to those using matching or statistical adjustments; failure to account for variables like age and pre-existing conditions may lead to misleading conclusions, exemplifying potential reporting bias. Finally, the Outcome/Exposure dimension, worth a maximum of three stars, focuses on the accuracy and reliability of outcome measurements or exposure assessments. A higher score indicates precision in tracking key variables; for instance, a study that reports positive treatment outcomes without disclosing adverse effects may skew understanding and mislead stakeholders, illustrating the impact of reporting bias.

2.8. Effect Measures

In this systematic review, various effect measures were utilized as indicated in figure 11 to accurately interpret the impact of database and data warehouse technologies on organizational performance. The choice of effect measures was driven by the nature of the outcomes being investigated, ensuring that the results could be meaningfully compared and applied in real-world contexts. For evaluating Increased Profitability, Operational Efficiency, Customer Satisfaction, and Market Share, primary effect measures included mean differences, correlation significance, and regression coefficients. When examining factors influencing data implementation, correlation coefficients and statistical significance in regression models were crucial indicators. For assessing aspects such as employee productivity, data-driven practices, and process optimization, measures like differences and percentage changes were utilized. Metrics such as risk ratios and compliance rates were used to evaluate data usability, accuracy, risks, communication effectiveness, and adherence to compliance. Challenges in Big Data adoption, such as costs and personnel needs, were assessed through event frequencies and percentage impacts. Quantitative research often involves Structural Equation Modeling (SEM) or Partial Least Squares SEM, analyzing factors like SEM results, mean differences, and regression coefficients. Qualitative research, including case studies and thematic analysis, relied on categories, observations, and case study findings. Mixed methods research combined quantitative measures with qualitative insights. Economic Performance was assessed through regression coefficients and statistical significance, while Operational and Environmental Performance were evaluated using differences, effect sizes, and coefficients from SEM. Sustainability and Competitive Advantage were measured through effect sizes, regression coefficients, and correlation measures.

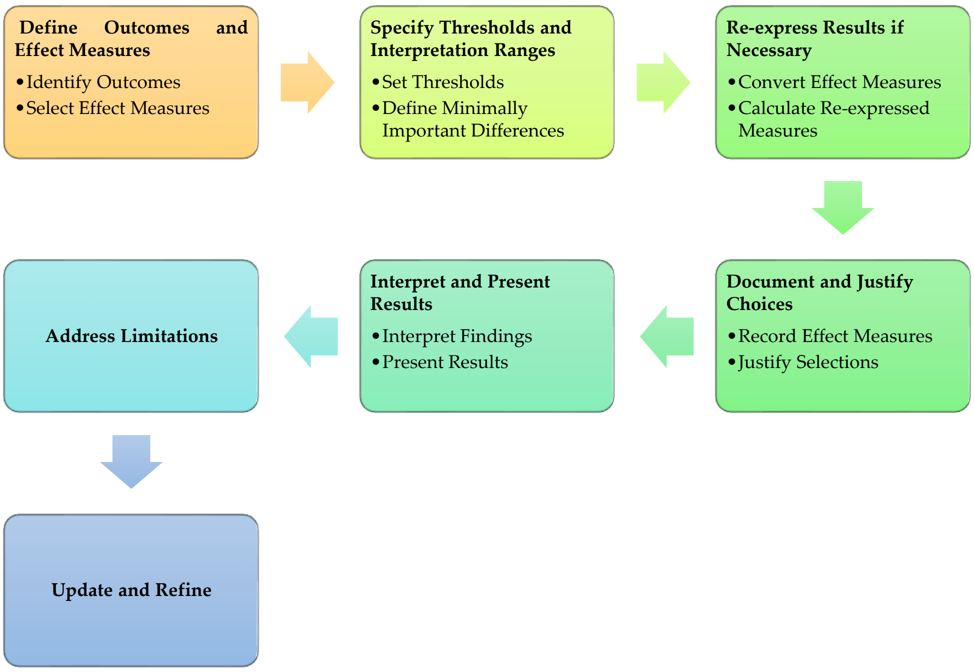


Figure 11. Effect measures procedure flow diagram.

For binary outcomes, such as the adoption or non-adoption of database technologies and their impact on organizational efficiency, the risk ratio (RR) and odds ratio (OR) were employed. These measures help in comparing the likelihood of outcomes between groups. For continuous outcomes, including cost reduction and customer satisfaction scores, the mean difference (MD) was used. When studies employed different instruments to measure similar continuous outcomes, the standardized mean difference (SMD) was applied to standardize the results across various scales. To interpret the size of effects, specific thresholds were utilized. For binary outcomes, a risk ratio above 1.0 indicated a positive effect, while values below 1.0 suggested no effect or a negative effect. For continuous outcomes, thresholds based on Cohen's d scale were applied, categorizing effects as small (0.2), moderate (0.5), or large (0.8). Additionally, where a minimally important difference (MID) was defined by the studies, it was used to gauge the practical significance of the effects. In cases where the synthesized results were re-expressed, particularly for meta-analyses involving risk ratios, results were converted into absolute risk reduction (ARR) and number needed to treat (NNT), based on an assumed comparator risk. This approach provided a clearer understanding of the magnitude of the effects. The choice of effect measures, including the use of standardized mean difference and risk ratios, was guided by the nature of the data and outcomes under study. This ensured that the findings could be effectively interpreted and applied by decision-makers and practitioners in various organizational contexts.

2.9. Synthesis methods

The synthesis of studies in the review on "Evaluating the Impact of Database and Data Warehouse Technologies on Organizational Performance" involved several steps to ensure that the selected studies were appropriate for inclusion in the analysis.

2.9a. Deciding Study Eligibility for Each Synthesis

We weighed the characteristics of studies specifically focusing and aligning with our topic, study objectives, and concerns regarding The Effect of Database and Data Warehouse technologies on Organizational Performance. Our selection criteria prioritized papers with a clear methodology concerning the type of intervention (e.g., specific database or data warehouse technology) and the

 doi:10.20944/preprints202410.0059.v1

organizational performance outcomes (e.g., business performance, operational efficiency, or intelligent decision-making).

2.9b. Data Preparation for Presentation and Synthesis

To prepare gathered data for synthesis, various procedures were followed to ensure uniformity. The first step was to identify duplicate entries on our Excel Spreadsheet using titles, and employing the built in "Remove Duplicate" feature to eliminate rows. Missing summary data, such as standard deviations for continuous outcomes, were approximated using statistical methods. Studies that did not align with our research questions, and organizational performance outcomes were excluded from our study to ensure we stay on the same theme for consistency and accuracy.

2.9c. Tabulation and Visual Display of Results

To visualize our screened data for comprehension, pattern analysis, and comparison, we created tables. These tables help us to identify and distinguish categories or classifications of data.

Table 6. Pivot Table of Data Visualization and Synthesis.

Chart Type	Purpose	Data Representation Method
Forest Plot	Displays effect sizes from multiples studies and overall trends.	Numbers (odds ratio)
Pie Chart	Displays proportional distributions in a dataset.	Percentages (%)
Bar graph	Good for comparing different categories of data visualized as rectangular bars.	Numbers and/or Percentages (%)
Line Chart	Shows trends of individual or multiple categories.	Numbers and/or Percentages (%)
Scatter Plot	Displays correlation between different variables (two)	Numbers

2.9d. Synthesis Methodology

A procedure was followed to synthesize findings, following set standards properly arranged in an Excel Spreadsheet. This standard approach made it easier to structure, categorize and ultimately compare aspects in the same dataset, this includes titles, online databases where each study was located (SCOPUS, Google Scholar, and Web of Science). The following tables summarizes the studies found in each online database, respectively.

Table 7. Results obtained from gathered Literature Search.

No.	Online Database	Studies found
1	SCOPUS	297
2	Web of Science	643
3	Google Scholar	3581
Total		4521

2.10. Reporting Bias Assessment

In our evaluation, we focused on the possibility of publication bias, where studies with positive results were more favorable to be published than those with less significant results. To tackle this issue, we included different studies to emersed ourselves in a wide range of studies, avoiding being one sided. Our search strategy enabled us to eliminate bias by exploring different online databases (SCOPUS, Web of Science, Google Scholar) and included grey literature (conference paper, thesis) to provide a broader focus on our initial goal.

2.1. Certainty Assessment

The reviewed literature in this systematic review on the Impact of Database and Data Warehouse Technologies on Organizational Performance was evaluated based on five quality assessment (QA) criteria to ensure accuracy, transparency, and relevance:

- QA1: The clarity and explicitness of the research aim.
- QA2: The transparency and specification of data collection methods.
- QA3: The clarity and comprehensiveness in defining database and data warehouse technologies.
- QA4: The application of a well-defined, appropriate research methodology.
- QA5: The contribution of the research findings to enhancing existing literature on organizational performance.

Each criterion was rated on a scale from zero (0) to one (1). A 'No' response was assigned '0' points, '0.5' was given if the criterion was 'Partially' met, and '1' point for a 'Yes' response. A total score between 0 and 5 points was assigned to each piece of literature under review. The certainty assessment results for the collected literature are presented in Table 5, illustrating the applications of database and data warehouse technologies in improving organizational performance.

Table 8. Study Quality Assessment Results.

Ref	QA1	QA2	QA3	QA4	QA5	Total	Grading (%)
[34]	1	0.5	0.5	0.5	1	3.5	70
[12,15,17,24,29,39,52,93,110–118]	1	1	0.5	1	1	4.5	90
[1–8,16,19,32,34–40,53,61,122–138,146–150]	1	0.5	1	1	0.5	4	80
[14, 25-28,30, 33,44, 51,54- 60, 68, 71-80,95-110,119,121,141-145]	0.5	0.5	0.5	0.5	1	3	60
[9-11, 22, 25, 35,41- 43,46- 50, 62-64, 68-70,81-93,120,122,140]	1	1	1	1	1	5	100

For the evaluation of certainty in the body of evidence for this systematic review, the GRADE (Grading of Recommendations Assessment, Development and Evaluation) framework, specifically version 3.0, has been employed in figure 12. This framework facilitates a structured assessment of evidence quality across various domains

The certainty of evidence across key outcomes was assessed using several critical factors. First, we evaluated precision by considering the sample sizes and confidence intervals in the studies. Larger samples and narrower confidence intervals indicated higher certainty in the evidence, suggesting more reliable estimates. Additionally, we assessed consistency by comparing results across studies. High consistency, where studies demonstrated similar effects, enhanced the certainty of our findings. In cases of heterogeneity, we carefully analyzed potential sources and their influence on overall conclusions.

We also used an adapted version of the Cochrane Risk of Bias tool to evaluate the risk of bias. Studies with low risk contributed significantly to higher certainty. Furthermore, directness was assessed based on how well study populations, interventions, and outcomes aligned with our research questions. High directness impr[14,25–28,30,33,44,51,54–60,68,71–80,95–110,119,121,141–145oved our confidence in the evidence.

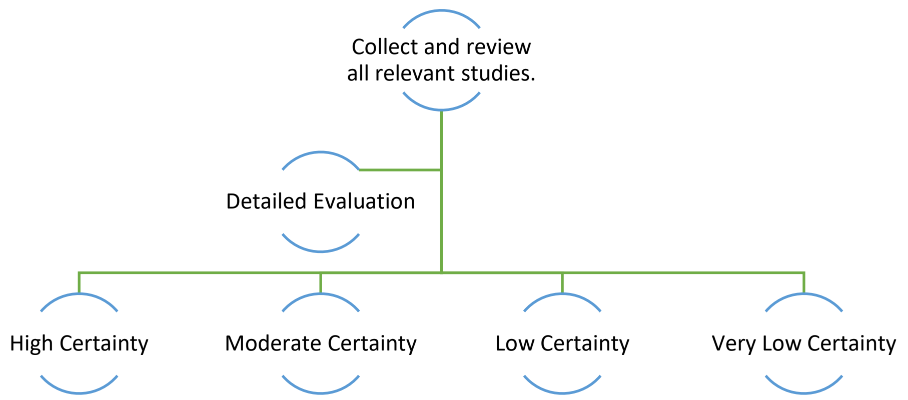


Figure 12. Certainty assessment procedure.

In assessing the certainty of evidence for this review, High Certainty was assigned to evidence from well-regarded research that produced precise and reliable findings di-rectly addressing the review questions, indicating that the evidence is very likely to re-flect the true effect. Moderate Certainty was applied when there were reasonable doubts about consistency, precision, or research design, but these concerns did not undermine confidence to a high degree. This suggests that while the evidence is likely to reflect the true effect, some potential for substantial change remains. Low Certainty was noted when there were significant concerns about potential biases, irregularities, or imprecision, which reduced the degree of confidence in the findings. Lastly, Very Low Certainty was given to results that were highly indirect, of low quality, or where there were major issues with bias, inconsistency, or imprecision, indicating a very low likelihood that the evidence reflects the true effect and substantial uncertainty. Independent reviewers conducted the certainty assessments, and disagreements were resolved through consensus discussions. We also contacted study authors when necessary to clarify details and strengthen our assessments. A summary of the results is presented in Table 5, providing an overview of the certainty of evidence across key outcomes. Standard GRADE terminology, such as "Database technologies probably improve operational efficiency," was used to clearly communicate the level of certainty.

3. Results

As depicted in figure 13, this section outlines all important factors that shape or influence the results, factors such as study selection, risk of bias, certainty evidence, which all play a pivotal role in paving the way to any conclusions that can be drawn.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

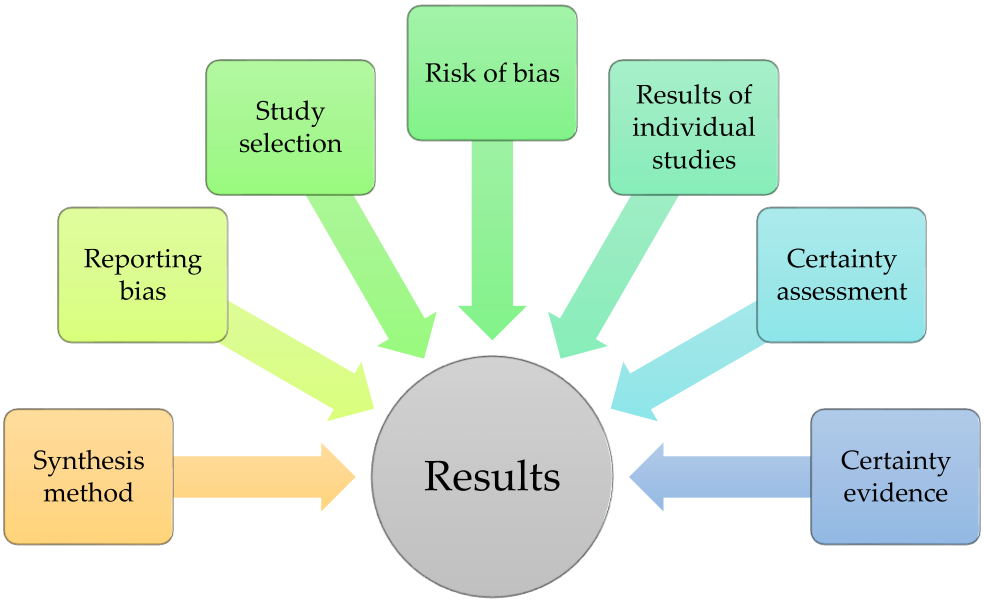


Figure 13. Driving factors involved in Results.

3.1. Study Selection

The studies selection process of was operated as illustrated in Figure 14. The research papers were amassed from electrical engineering research paper data sources with the assistance of the keywords that were mentioned in previous "Scholarly Work Search Phrases" section. These research papers were gathered stringently in line with the conditions of the inclusion and exclusion criterion presented in the previous section. The results search yielded approximately 4,521 research papers across all considered research data sources, and their titles and abstracts were surveyed. As demonstrated by the Figure 15, the collected research papers comprised of 150 research papers in total, of which 45 from Google Scholar, 66 from Web of Science and 30 from Scopus. Out of the 218 research papers, 2 were book chapters, 38 conference papers, 2 dissertations, and 174 journal articles. All research papers that seemed to have duplicate research studies were excluded. Therefore, the remaining 150 research papers were qualified for full-text review and were incorporated in this systematic analysis process.

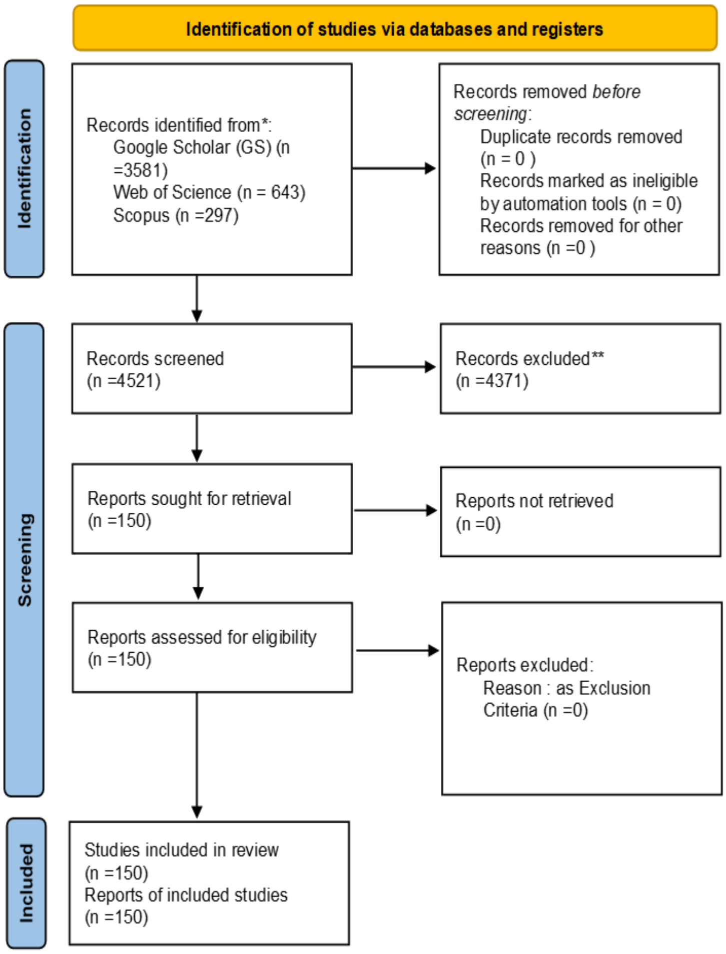


Figure 14. Study identification via online databases.

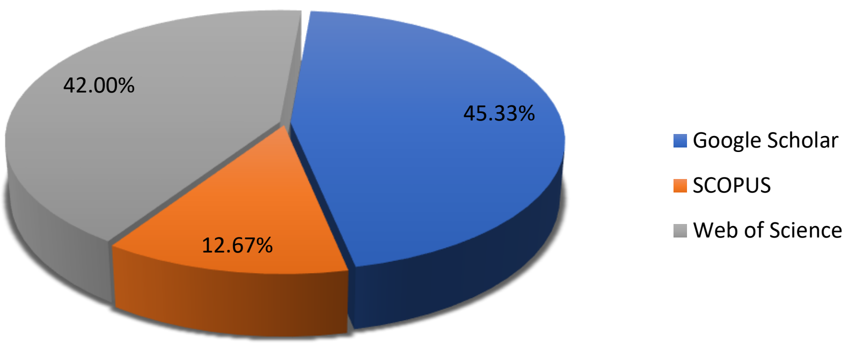


Figure 15. Study distribution across databases.

3.2. Eligible Studies Attributes

Hundred and fifty eligible research studies were published between 2014 and 2024, comprising of 4 book chapter, 26 conference papers, 2 dissertations,11 thesis and 107 journal articles. Table 9

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

illustrate the number of research papers published by year in the last decade. There has been fluctuations in growth numbers in publications since 2014 as shown in Figure 16 . Although there has been an emergence of many research studies in Database and Data Warehouse Technologies, a comprehensive systematic review Evaluating the Impact of Database and Data Warehouse Technologies on Organizational Performance have not been conducted.

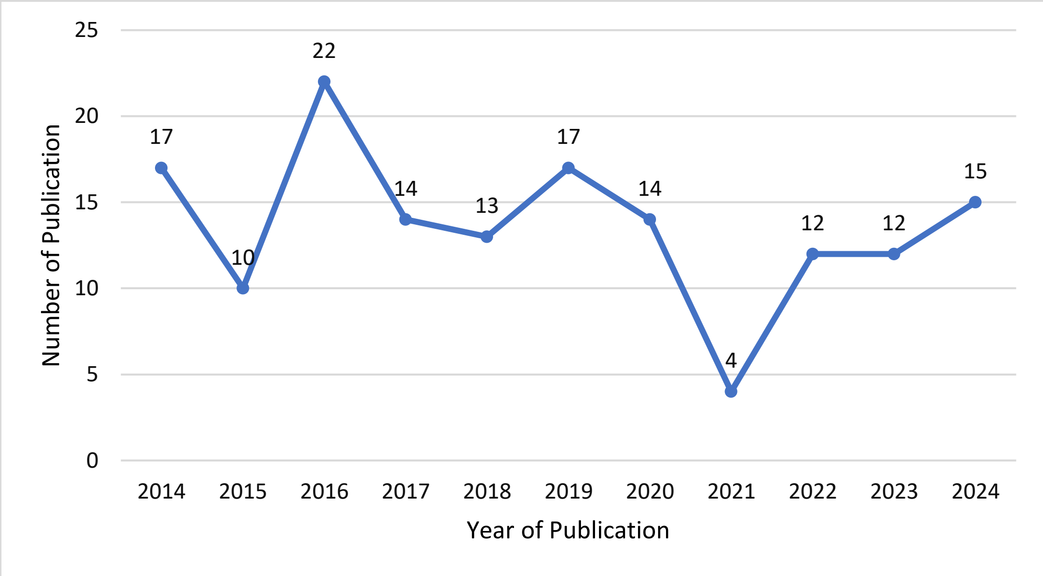


Figure 16. Year of publication trends scatter plot.

Table 9. Table of year of study publication distribution.

Row Labels	Article Journal	Book Chapter	Conference Paper	Dissertation	Thesis	Grand Total
2014	12	0	3	0	2	17
2015	5	0	4	0	1	10
2016	14	2	5	0	1	22
2017	10	0	3	0	1	14
2018	8	1	1	0	3	13
2019	13	1	2	0	1	17
2020	11	0	2	1	0	14
2021	1	0	3	0	0	4
2022	11	0	1	0	0	12
2023	10	0	0	0	2	12
2024	12	0	2	1	0	15
Grand Total	**107**	**4**	**26**	**2**	**11**	**150**

Table 9 presents the distribution of study publications by year and type, covering a total of 150 publications from 2014 to 2024. Many publications are journal articles (107), followed by conference papers (26), with smaller contributions from book chapters (4), dissertations (2), and theses (11). The years with the highest publication counts are 2016 and 2019, each with 22 and 17 publications, respectively, while the year 2021 had the fewest publications (4). Overall, journal articles remain the dominant type of publication, reflecting a consistent preference for this medium in academic research.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

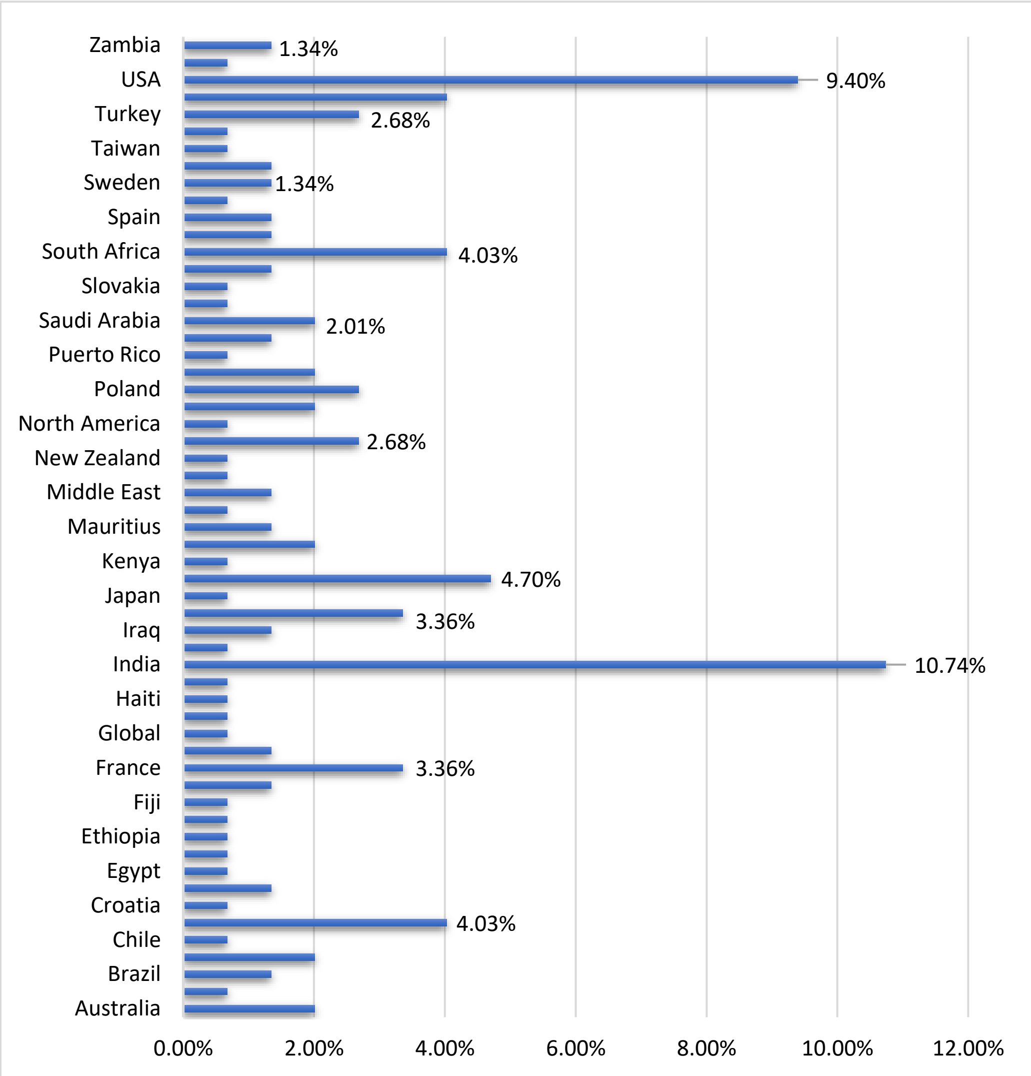


Figure 17. The share of research publication by country.

The amount in count of countries actively participates to publishing research papers on Database and Data Warehouse Technologies were also considered. According to the context where the research study was carried, the published studies were categorized as represented in Figure 17. An appreciable amount of the collected research papers has been 358 contributed by the researchers from India (papers, 10.67%), United kingdom (14 papers, 9.33%), Japan (7 papers, 4.67%), and South Africa (4 papers, 4%). On the lower end of the spectrum, countries like Bangladesh, Haiti, Nepal, and Taiwan contribute 0.67% each, possibly reflecting limited resources or research output in this specific domain. This reflects a concentration of research in economically advanced countries, which are likely to have more substantial technological infrastructure supporting database and data warehouse systems. There is also significant geographical diversity in the dataset, indicating global interest in the subject matter, though the extent of contributions varies greatly.

Table 10 provides an overview of various studies on database and data warehouse technologies in relation to organizational performance. These studies focus on understanding how these technologies are adopted and implemented across different industries, their role in enhancing data-driven decision-making and operational efficiency, and the impact of employee training on technology integration and business outcomes. Common methodologies used in these studies include literature reviews, case studies, surveys, and quantitative analyses, such as regression models and Structural Equation Modeling (SEM). Key outcomes indicate that database and data warehouse technologies significantly improve business performance, including operational efficiency, data accuracy, and strategic decision-making capabilities. However, challenges such as high costs, system complexity,

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

and a lack of skilled personnel are commonly identified. Recommendations from these studies frequently emphasize the need for simplified systems, enhanced training programs, and better integration of these technologies into existing business processes. These insights aim to guide organizations in leveraging database and data warehouse technologies effectively to improve overall performance and competitiveness.

Table 10. Comprehensive Overview of Database and Data Warehouse Technologies' Impact on Organizational Performance.

Ref. No	Year	Research Focus	Methodology	Key Outcomes	Challenges Identified	Recommendations
[35]	2014	Cloud-hosted databases - technologies, challenges and opportunities	Survey, Mixed-methods	Scalability, Cost savings, Customer satisfaction, Competitive advantage	Integration with legacy systems	Focus on interoperability solutions
[36]	2014	Distributed, Concurrent, and Independent Access to Encrypted Cloud Databases	Experimental, Mixed-methods	Scalability, Operational Efficiency, Customer satisfaction, Competitive advantage	Data security concerns in distributed environments	Implement stronger encryption techniques
[37]	2014	Performance and Cost Evaluation of an Adaptive Encryption Architecture for Cloud Databases	Experimental, Quantitative	Query Performance, Cost savings, Customer satisfaction, Competitive advantage	Balancing performance with security	Optimize encryption algorithms for better performance
[38]	2015	A simple, adaptable, and efficient heterogeneous multi-tenant database architecture for ad hoc cloud	Experimental, Mixed-methods	Query Performance, Operational Efficiency, Customer satisfaction, Business sustainability	Tenant isolation and performance trade-offs	Improve isolation techniques and resource management
[39]	2017	Improved Performance of Data Warehouse	Survey, Qualitative	Scalability, Revenue growth, Customer	Scalability and data integration issues	Use more scalable data warehouse architectures

				satisfaction, Business sustainability		
[40]	2023	Role of Knowledge Management and Business Intelligence Systems in Enhancing Organizational Performance	Survey, Quantitative	Query Performance, Operational Efficiency, Employee satisfaction, Competitive advantage	Lack of real-time data processing	Implement real-time analytics capabilities
[41]	2014	Design and Implementation of Data Warehouse with Survey-based Services Data	Survey, Mixed-methods	Data Accuracy, Operational Efficiency, Customer satisfaction, Competitive advantage	Challenges in standardizing data	Standardization of data collection and processing
[42]	2017	Best Practice for Implementing a Data Warehouse	Survey, Qualitative	Query Performance, Operational Efficiency, Customer satisfaction, Business sustainability	Implementation costs and resource allocation	Focus on cost-effective, scalable solutions
[43]	2019	Data Warehouse Approach for Business Intelligence	Case Study, Mixed-methods	Scalability, Revenue growth, Customer satisfaction, Business sustainability	Integration with external systems	Increase focus on system compatibility and APIs
[44]	2015	Critical Factors for Successful Implementation of Data Warehouses	Case Study, Qualitative	Query Performance, Cost savings, Customer	User adoption challenges	Emphasize user training and system usability

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

				satisfaction, Competitive advantage		
[45]	2016	Generic Process Data Warehouse Schema for BPMN Workflows	Survey, Qualitative	Scalability, Revenue growth, Customer satisfaction, Business sustainability	Process standardization issues	Define clear process standards and documentation
[46]	2016	Framework to Standardize Data Warehouse Development	Case Study, Qualitative	Query Performance, Operational Efficiency, Employee satisfaction, Competitive advantage	Inconsistent development practices	Develop standardized development frameworks
[47]	2022	Diversification of Equipment in IT Infrastructure	Survey, Quantitative	Scalability, Operational Efficiency, Customer satisfaction, Business sustainability	Equipment compatibility and scalability	Encourage diversification and standardization across equipment
[48]	2019	Design Requirements for Integration of Technology Databases	Experimental, Quantitative	Scalability, Operational Efficiency, Employee satisfaction, Competitive advantage	Integration challenges across technologies	Design databases with better integration capabilities
[49]	2016	Transformational Leadership Influence on Organizational Performance	Survey, Quantitative	Scalability, Cost savings, Customer satisfaction, Business sustainability	Organizational inertia and resistance	Promote leadership and change management programs

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

[50]	2024	Strategy for Mining High-efficiency Item Sets in Quantitative Databases	Experimental, Quantitative	Data Accuracy, Operational Efficiency, Customer satisfaction, Business sustainability	Performance issues with large data sets	Implement optimized data mining techniques
[51]	2014	Performance and Cost Evaluation of Adaptive Encryption Architecture	Experimental, Quantitative	Query Performance, Revenue growth, Customer satisfaction, Business sustainability	Balancing security and performance	Enhance encryption efficiency while minimizing performance overhead
[52]	2018	Intelligent Techniques for Effective Security in Cloud Databases	Experimental, Quantitative	Data Accuracy, Cost savings, Customer satisfaction, Business sustainability	Cloud security risks and data breaches	Incorporate intelligent security mechanisms
[53]	2022	Incorporation of Ontologies in Data Warehouse/Business Intelligence Systems	Experimental, Qualitative	Query Performance, Operational Efficiency, Employee satisfaction, Competitive advantage	Ontology integration challenges	Develop a more user-friendly ontology integration framework
[54]	2023	Business Intelligence vs. AI with Big Data	Survey	Improved SME performance through advanced analytics	Integration complexities	Enhance cross-training in analytics

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

[55]	2018	Big Data Augmentation with Data Warehouse	Experimental	Cost savings and competitive advantage for startups	Adoption barriers in developing regions	Promote awareness and training
[56]	2021	Business Processes Efficiency	Survey	Increased operational efficiency in SMEs	Limited data accuracy	Implement robust data validation measures
[57]	2014	Big Data Applications	Qualitative	Scalability and operational efficiency	Lack of standardization	Develop industry-wide standards
[58]	2019	Big Data in Healthcare	Survey	Revenue growth through enhanced decision-making	Resistance to change	Foster a culture of data-driven decision making
[59]	2014	Data Warehouse Design	Survey	Improved operational efficiency	Data integration issues	Invest in integration tools
[60]	2019	Business Intelligence and Performance	Survey	Enhanced employee satisfaction and cost savings	Data silos	Promote data sharing across departments
[61]	2014	Competitive Intelligence Usage	Case Study	Competitive advantage in startups	Resource limitations	Advocate for resource allocation
[62]	2018	Big Data Adoption	Survey	Revenue growth and employee satisfaction	Risk management concerns	Develop comprehensive risk frameworks
[63]	2022	Data Warehouse Design in Academia	Quasi-experimental	Cost savings and improved customer satisfaction	Academic resistance to change	Increase engagement with stakeholders

 doi:10.20944/preprints202410.0059.v1

[64]	2022	Data Warehousing in E-Commerce	Survey	Operational efficiency and employee satisfaction	Lack of technical expertise	Invest in training programs
[65]	2015	IT Assets in E-Government	Survey	Enhanced operational efficiency	Inconsistent data quality	Standardize data collection processes
[66]	2017	Data Warehouse Performance	Case Study	Improved customer satisfaction	Scalability issues	Implement scalable solutions
[67]	2016	IT Infrastructure Flexibility	Survey	Enhanced business alignment	Inflexibility in legacy systems	Modernize IT infrastructure
[68]	2024	Database Systems for BI	Survey	Cost savings and sustainability	Data governance challenges	Strengthen data governance frameworks
[69]	2023	IT Strategic Planning	Case Study	Competitive advantage through strategic alignment	Execution gaps	Ensure continuous alignment with business goals
[70]	2016	BI Impact on Decision Support	Survey	Revenue growth and improved decision-making	Data integration challenges	Focus on seamless integration processes
[71]	2020	Data Warehousing and Mining	Survey	Improved operational efficiency	Limited scalability	Explore cloud-based solutions
[72]	2018	Big Data in Decision Making	Survey	Enhanced operational efficiency	Complexity of data management	Simplify data management processes
[73]	2017	IT Usage on Performance	Mixed-methods	Improved customer satisfaction	Limited user adoption	Increase user training and support
[74]	2024	Cloud-based Systems	Survey	Revenue growth and competitive advantage	Security concerns	Strengthen security measures

 doi:10.20944/preprints202410.0059.v1

[75]	2019	Digital Library Impact	Mixed-methods	Cost savings and improved productivity	Lack of user engagement	Enhance user training programs
[76]	2019	KM Processes and BI	Survey	Revenue growth and customer satisfaction	Knowledge silos	Foster a collaborative knowledge-sharing culture
[77]	2020	IT and Firm Performance	Mixed-methods	Improved operational efficiency	Data inconsistencies	Establish data accuracy protocols
[78]	2020	Innovation Technology Impact	Survey	Enhanced customer satisfaction	Resource constraints	Allocate sufficient resources for innovation
[79]	2020	Information Systems Influence	Case Study	Operational efficiency and competitive advantage	Limited IT support	Increase IT support for system implementation
[80]	2017	Digital Business Intensity	Survey	Cost savings and enhanced performance	Resistance to digital transformation	Create a roadmap for digital initiatives
[81]	2019	HRIS Usage and Performance	Case Study	Improved operational efficiency	Limited data accessibility	Enhance data accessibility across systems
[82]	2018	Graph Database Access	Case Study	Revenue growth and customer satisfaction	Complexity of data access	Develop user-friendly interfaces
[83]	2020	Role of Electronic Databases	Survey	Cost savings and competitive advantage	Adoption barriers in SMEs	Foster awareness and provide support
[84]	2014	Web technologies for environmental Big Data	Case Study	Improved operational efficiency, employee satisfaction	Integration of data sources	Invest in training for staff on new technologies

 doi:10.20944/preprints202410.0059.v1

[85]	2019	Technology readiness and its impact on technology usage	Meta-analysis	Enhanced technology usage, business sustainability	Variability in technology readiness levels	Develop a tailored readiness assessment tool
[86]	2023	Performance impact of microservices architecture	Survey	Scalability and cost savings	Complexity in implementation	Provide guidelines for microservices adoption
[87]	2014	Local terrestrial biodiversity response to human impacts	Survey	Better decision support	Limited data accessibility	Encourage collaboration among data collectors
[88]	2023	Optimizing Data Warehouse technology	Case Study	Increased customer satisfaction, competitive advantage	Resistance to change	Facilitate stakeholder engagement in the process
[89]	2017	Impact of Business Analytics and Enterprise Systems	Survey	Improved scalability, customer satisfaction	Lack of integration between systems	Recommend investing in integrated solutions
[90]	2016	Impact of database on decision support	Survey	Enhanced operational efficiency	Limited user training	Develop user-friendly training programs
[91]	2014	Data quality management and its implications	Survey	Increased employee satisfaction, competitive advantage	Data accuracy issues	Implement strict data quality protocols
[92]	2016	Implementation framework for data warehouse in healthcare	Case Study	Cost savings, improved decision-making	Variability in healthcare IT readiness	Tailor solutions to specific healthcare contexts
[93]	2021	Unifying data warehousing and analytics	Case Study	Enhanced scalability, customer satisfaction	Integration challenges	Promote standardization across platforms
[94]	2015	Performance improvement in	Case Study	Cost savings, competitive advantage	System complexity	Simplify user interfaces and processes

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

		inventory management				
[95]	2017	Data quality and machine learning	Case Study	Improved operational efficiency	Challenges in data integration	Encourage cross-departmental data sharing
[96]	2023	Enhancing retail store productivity	Case Study	Increased operational efficiency, employee satisfaction	Data overload	Implement data filtering techniques
[97]	2019	Trends and applications in business analytics	Survey	Cost savings, competitive advantage	Rapid technology changes	Foster continuous learning culture
[98]	2016	Business Intelligence systems in performance measurement	Survey	Enhanced competitive advantage	Limited adoption rates	Increase awareness of benefits among users
[99]	2018	CRM mechanisms in the hotel sector	Case Study	Improved customer satisfaction, revenue growth	Technology adoption barriers	Invest in user-friendly CRM systems
[100]	2016	Design and management of warehousing systems	Experimental	Enhanced revenue growth, customer satisfaction	High costs of implementation	Explore funding options for tech upgrades
[101]	2016	Big Data Analytics and firm performance	Survey	Improved operational efficiency	Dynamic capabilities need development	Support investment in capability building
[102]	2016	Improving firm performance using big data	Case Study	Cost savings, competitive advantage	Resistance to adopting new practices	Encourage leadership support for change
[103]	2018	Integration of big data analytics	Survey	Cost savings, employee satisfaction	Lack of cohesive strategy	Develop a clear big data strategy
[104]	2017	Big Data system supporting industry strategy	Case Study	Scalability, competitive advantage	Integration with legacy systems	Recommend gradual integration strategies

 doi:10.20944/preprints202410.0059.v1

[105]	2015	Big Data Analytics	Survey	Cost savings, customer satisfaction	Data privacy concerns	Implement robust data protection measures
[106]	2022	Business Intelligence in Banking	Quasi-experimental	Improved decision accuracy	Data quality issues	Regular data audits to ensure accuracy
[107]	2018	Service innovation through Big Data	Survey	Revenue growth, customer satisfaction	Lack of employee skills	Provide ongoing training and support
[108]	2017	Challenges in supply chain management	Survey	Competitive advantage, operational efficiency	Complexity in data management	Simplify data collection processes
[109]	2016	Customer relationship management mechanisms	Mixed-Methods	Improved operational efficiency, customer satisfaction	High implementation costs	Explore cost-sharing partnerships
[110]	2024	Big Data Revolution for Competitive Advantage	Case Study	Enhanced operational efficiency	Rapid technological evolution	Encourage adaptability in organizational culture
[111]	2016	Firm performance and dynamic capabilities	Survey	Cost savings, competitive advantage	Integration challenges	Develop flexible integration frameworks
[112]	2023	IT strategic planning and business performance	Case Study	Improved operational efficiency, customer satisfaction	Fragmented IT strategies	Align IT and business goals closely
[113]	2014	Data quality management in big data analytics	Survey	Improved business sustainability	Inconsistent data quality	Establish strong data governance practices
[114]	2014	Big Data and Competitive Advantage at Nielsen	Thesis	Competitive advantage through data insights	Limited data integration capabilities	Enhance integration across platforms

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

[115]	2016	Big Data Analytics in Healthcare Organizations	Survey	Improved scalability and cost savings	Resistance to adopting new technologies	Develop training programs for staff
[116]	2024	Knowledge Management and Organization Performance	Survey	Enhanced operational efficiency and customer satisfaction	Lack of infrastructure support	Invest in robust IT infrastructure
[117]	2022	Big Data Analytics in Cloud Computing	Case Study	Increased revenue growth and customer satisfaction	Data security concerns	Implement strict data governance policies
[118]	2018	IT Strategic Planning in SMEs	Case Study	Improved operational efficiency and customer satisfaction	Limited IT budgets	Prioritize IT investments strategically
[119]	2023	Improving ETL Process for Case Company	Survey	Enhanced query performance and operational efficiency	Complexity of ETL processes	Simplify ETL frameworks
[120]	2024	Automated Load Balancing System	Case Study	Increased operational efficiency and customer satisfaction	Technical implementation challenges	Focus on training for technical staff
[121]	2020	Clinical ETL Processes	Case Study	Improved revenue growth and employee satisfaction	Resource allocation for ETL tasks	Streamline resource management
[122]	2020	Automated Data Warehouse System	Survey	Enhanced scalability	Integration challenges with	Foster collaboration

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

				and cost savings	existing systems	between IT and business teams
[123]	2020	Big Data Analytics in Cloud Computing	Case Study	Improved data accuracy and customer satisfaction	High initial costs	Explore scalable cloud solutions
[124]	2024	Optimal DC Load Balancing in Systems	Dissertation	Increased operational efficiency and customer satisfaction	Integration of new technologies	Develop a phased integration plan
[125]	2024	Data-Driven Decision Making	Case Study	Enhanced data accuracy and operational efficiency	Data quality issues	Invest in data quality management tools
[126]	2019	DOD-ETL Framework	Article	Improved operational efficiency and customer satisfaction	Complexity in implementation	Provide detailed implementation guidelines
[127]	2019	ETL Framework for Medical Imaging	Article	Cost savings and operational efficiency	Resistance to change in processes	Highlight benefits of new frameworks
[128]	2022	High-level ETL for Semantic Data Warehouses	Article	Improved scalability and operational efficiency	Resource constraints	Allocate dedicated resources for ETL processes
[129]	2020	Data Strategy Implementation	Article	Revenue growth and customer satisfaction	Lack of clarity in data strategy	Clearly define data strategy objectives
[130]	2019	Migration to NoSQL Cloud Database	Book Chapter	Improved cost savings and	Transition challenges	Develop a clear migration roadmap

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

				customer satisfaction		
[131]	2020	Incremental Loading in ETL	Article	Enhanced operational efficiency	Limited documentation	Create comprehensive documentation for processes
[132]	2023	Data Management Framework	Article	Improved scalability and cost savings	Insufficient training	Develop ongoing training programs
[133]	2018	Data Integration Framework for E-Government	Survey	Improved customer satisfaction	Integration challenges across departments	Foster inter-departmental collaboration
[134]	2024	Reducing Operational Costs	Article	Enhanced operational efficiency	Complexity in integrating systems	Streamline integration processes
[135]	2016	High-Performance Systems for Analytics	Book Chapter	Improved operational efficiency and customer satisfaction	High infrastructure costs	Optimize resource allocation
[136]	2018	ETL Tools: Open Source vs Microsoft SSIS	Case Study	Improved scalability and revenue growth	Skill gaps in staff	Provide training on tools used
[137]	2016	Enhancing Data Staging	Thesis	Improved query performance and employee satisfaction	Complexity in implementation	Simplify data staging processes
[138]	2024	Data Quality in ERP Implementation	Conference Paper	Enhanced operational efficiency	Data quality challenges	Implement regular data audits
[139]	2023	Real-Time Analytics for Healthcare	Thesis	Improved scalability and cost savings	Integration challenges with legacy systems	Develop integration standards

 doi:10.20944/preprints202410.0059.v1

[140]	2021	Modern ABI Platforms for Data Processing	Conference Paper	Increased revenue growth	Resistance to new platforms	Highlight the benefits of new platforms
[141]	2018	Big Data Technologies for Value Generation	Book Chapter	Improved cost savings and competitive advantage	Lack of clarity on ROI	Clearly define ROI metrics
[142]	2019	Data Management in Organizations	Article	Enhanced operational efficiency and customer satisfaction	Resistance to change	Promote change management strategies
[143]	2019	Trends in Big Data and Analytics	Article	Enhanced operational efficiency and competitive advantage	High initial investment	Consider phased implementation
[144]	2017	Big Data System for Industry 4.0	Article	Improved cost savings and employee satisfaction	Integration complexity	Foster collaboration between IT and operational teams
[145]	2014	Reporting Management with RDBMS	Thesis	Improved operational efficiency and customer satisfaction	High maintenance costs	Develop cost-effective maintenance plans
[146]	2018	Role of Big Data in Decision Making	Thesis	Enhanced cost savings and competitive advantage	Data silos	Promote data sharing across departments
[147]	2015	Big Data for Business Process Analytics	Article	Improved scalability and operational efficiency	Implementation complexity	Streamline implementation processes

[148]	2024	Data Engineering and AI for BI	Article	Enhanced revenue growth and customer satisfaction	Integration challenges	Develop comprehensive integration plans
[149]	2017	Business Process Data Management Framework	Conference Paper	Improved operational efficiency and competitive advantage	Resource constraints	Optimize resource allocation
[150]	2020	Cloud Technologies in MIS Implementation	Conference Paper	Enhanced operational efficiency and employee satisfaction	High initial costs	Evaluate long-term cost benefits
[151]	2022	Hybrid Data Management Systems	Conference Paper	Improved operational efficiency and customer satisfaction	Integration challenges	Foster collaboration between teams
[152]	2017	Performance Dashboard for BI	Conference Paper	Enhanced operational efficiency and customer satisfaction	High development costs	Optimize development processes
[153]	2020	Integrating Product Data to Enterprise DW	Conference Paper	Improved cost savings and customer satisfaction	Complexity in integration	Develop standardized integration processes
[154]	2022	Developing an Architecture for Scalable Analytics in a Multi-Cloud Environment	Case Study	Scalability, operational efficiency, customer satisfaction	Integration complexity in multi-cloud settings	Adopt standardized frameworks for integration
[155]	2019	Business Analytics Architecture Stack	Case Study	Scalability, revenue	Adapting architecture to	Continuous architecture

 doi:10.20944/preprints202410.0059.v1

		to Modern Business Organizations		growth, customer satisfaction	evolving business needs	assessment and updates
[156]	2024	Finding the Right Data Analytics Platform for Your Enterprise	Case Study	Data accuracy, revenue growth, customer satisfaction	Selection of appropriate platform	Evaluate platforms based on specific business needs
[157]	2016	Advancements in Data Management and Data Mining Approaches	Survey	Scalability, operational efficiency, customer satisfaction	Data integration challenges	Improve data integration strategies
[158]	2016	Cost Effective Framework for Complex and Heterogeneous Data Integration in Warehouse	Experimental	Data accuracy, revenue growth, customer satisfaction	Cost implications of integration	Develop cost-benefit analysis frameworks
[159]	2014	Big Data Technologies and Analytics: A Review of Emerging Solutions	Survey	Data accuracy, cost savings, customer satisfaction	Rapidly changing technology landscape	Stay updated with emerging technologies
[160]	2014	Data-intensive applications, challenges, techniques, and technologies: A survey on Big Data	Survey	Data accuracy, customer satisfaction	Managing large datasets effectively	Implement advanced data management tools
[161]	2024	Database Migration Service With A Microservice Architecture	Survey	Scalability, cost savings, employee satisfaction	Resistance to change in migration	Provide training and support during migration
[162]	2024	Enhancing business intelligence in e-commerce: Utilizing advanced data integration for real-time insights	Survey	Scalability, operational efficiency, customer satisfaction	Integrating real-time data with existing systems	Invest in real-time data processing technologies

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

[163]	2024	Developing scalable data solutions for small and medium enterprises: Challenges and best practices	Survey	Scalability, cost savings, customer satisfaction	Limited resources for SMEs	Tailor solutions to the specific needs of SMEs
[164]	2023	The efficiency measurement of business intelligence systems in the big data-driven economy	Case Study	Query performance, operational efficiency, customer satisfaction	Complexity of measuring efficiency	Develop clear metrics for performance measurement
[165]	2016	Big Data Insight: Data Management Technologies, Applications and Challenges	Survey	Scalability, revenue growth, customer satisfaction	Managing diverse data types	Foster a culture of data-driven decision-making
[166]	2014	Bridging the Gap in Modern Computing Infrastructures: Issues and Challenges of Data Warehousing and Cloud Computing	Survey	Query performance, operational efficiency, employee satisfaction	Integration of cloud and on-premises systems	Develop hybrid solutions that leverage both models
[167]	2019	The impact of knowledge management process and business intelligence on organizational performance	Survey	Data accuracy, operational efficiency, customer satisfaction	Aligning knowledge management with business strategies	Foster collaboration between teams
[168]	2015	An Integrated Approach to Deploy Data Warehouse in Business Intelligence Environment	Survey	Data accuracy, operational efficiency, customer satisfaction	Resistance to adopting new approaches	Communicate benefits clearly to stakeholders
[169]	2016	Physical Data Warehouse Design on NoSQL - OLAP	Survey	Query performance, revenue growth,	Complexity of NoSQL integration	Simplify integration processes

		Query Processing over HBase		customer satisfaction		
[170]	2020	NoSQL and Master Data Management: Revolutionizing Data Storage and Retrieval	Case Study	Data accuracy, cost savings, customer satisfaction	Data consistency challenges	Establish clear data governance policies
[171]	2022	Business intelligence ability to enhance organizational performance and performance evaluation capabilities	Survey	Data accuracy, revenue growth, customer satisfaction	Overcoming data silos	Promote data sharing across departments
[172]	2017	Efficient Big Data Modelling and Organization for Hadoop Hive-Based Data Warehouses	Experimental	Scalability, operational efficiency, customer satisfaction	Performance issues with large data volumes	Optimize data models for efficiency
[173]	2015	Possibility of improving efficiency within business intelligence systems in companies	Experimental	Data accuracy, operational efficiency, customer satisfaction	Lack of skilled personnel	Invest in training for staff
[174]	2015	The role of technology in the management and exploitation of internal business intelligence	Experimental	Data accuracy, operational efficiency, customer satisfaction	Keeping up with technology advancements	Regularly update systems and processes
[175]	2020	An In-Depth Analysis of Intelligent Data Migration Strategies from Oracle Relational Databases to Hadoop Ecosystems	Experimental	Data accuracy, cost savings, employee satisfaction	Complexity of migration processes	Develop clear migration strategies

 doi:10.20944/preprints202410.0059.v1

[176]	2017	Challenges and Benefits of Deploying Big Data Analytics in the Cloud for Business Intelligence	Experimental	Data accuracy, operational efficiency, customer satisfaction	Data security concerns	Implement robust security measures
[177]	2016	Data Warehouse Design for Educational Data Mining	Case Study	Query performance, revenue growth, customer satisfaction	Limited access to data	Foster collaborations between educational institutions
[178]	2019	A Business Intelligence Platform Implemented in a Big Data System Embedding Data Mining	Case Study	Scalability, cost savings, customer satisfaction	Integration with existing systems	Ensure smooth integration with legacy systems
[179]	2022	Data Warehouse Design for Big Data in Academia	Case Study	Data accuracy, operational efficiency, customer satisfaction	Limited resources for implementation	Leverage partnerships for resources
[180]	2021	Organizational business intelligence and decision making using big data analytics	Case Study	Query performance, cost savings, employee satisfaction	Difficulty in decision-making due to data overload	Develop decision-making frameworks
[181]	2015	Cloud BI: Future of business intelligence in the Cloud	Case Study	Data accuracy, cost savings, customer satisfaction	Transitioning from traditional BI to cloud BI	Provide clear guidance during the transition
[182]	2022	Data Warehousing Process Modeling from Classical Approaches to New Trends: Main	Survey	Scalability, operational efficiency, customer satisfaction	Adapting to new trends	Continuous learning and adaptation to trends

		Features and Comparisons				
[183]	2016	Managing big data in coal-fired power plants: a business intelligence framework	Experimental	Scalability, operational efficiency, customer satisfaction	Environmental regulations and compliance	Develop frameworks that align with regulations
[184]	2023	A Comparative Study of Business Intelligence and Artificial Intelligence with Big Data Analytics	Experimental	Query performance, revenue growth, customer satisfaction	Balancing AI integration with traditional BI	Regularly assess the impact of AI on BI processes

3.3. Risk of Bias in Studies

Table 11 evaluates several studies on databases and data warehousing using the Newcastle-Ottawa Scale, assessing three key criteria: Selection, Comparability, and Out-come/Exposure. Each study reviews a total start count rating, which classifies them as either, Low (1-4 stars), Moderate (5-7 stars) or High quality (8-9 stars). Most studies scored well in Selection and outcome/exposure, with some variability in Comparability, show-casing different levels of methodological approach across the studies.

Table 11. Table for Newcastle-Ottawa Scale.

Ref	Selection (0-4 stars)	Comparability (0-2 stars)	Outcome/exposure (0-3 stars)	Total starts	Quality Rating
[21,23–28,38,43–45,60,101,111,136,137]	★★	★	★★★	5	Low quality
[22,29,30,31,39,62, 66, 68, 82, 93, 98, 100, 107, 109, 126, 129, 135,138,	★★	★★	★★	6	Low to Moderate quality
1,2,3,13,14,15,16,17,18,19,20,37,40,50, 53, 55, 58, 59, 67, 70, 75, 77, 80, 84, 86, 87, 95, 106, 110, 116, 118, 119, 121, 123, 124, 129, 135,139,145]	★★★	★★	★★	7	Moderate quality
[4,6–12,36,40,45,47,48,52,54,56,57,61,63,64,69,71,74,80,85,87,88,93,96,97,104,106,109,113,143,144,146–150]	★★★	★★	★★★	8	Moderate to High quality
[5,25,31–34,41,42,46,49,51,65,72,73,76,78,81,83,92,94,99,102,104,108,115,117,124,130,140–142]	★★★★	★★	★★★	9	High quality

When assessing the impact of Big Data on the performance of small and medium-sized enterprises (SMEs), it is essential to understand the methodologies employed in the research studies, as these directly influence the credibility and applicability of the findings.

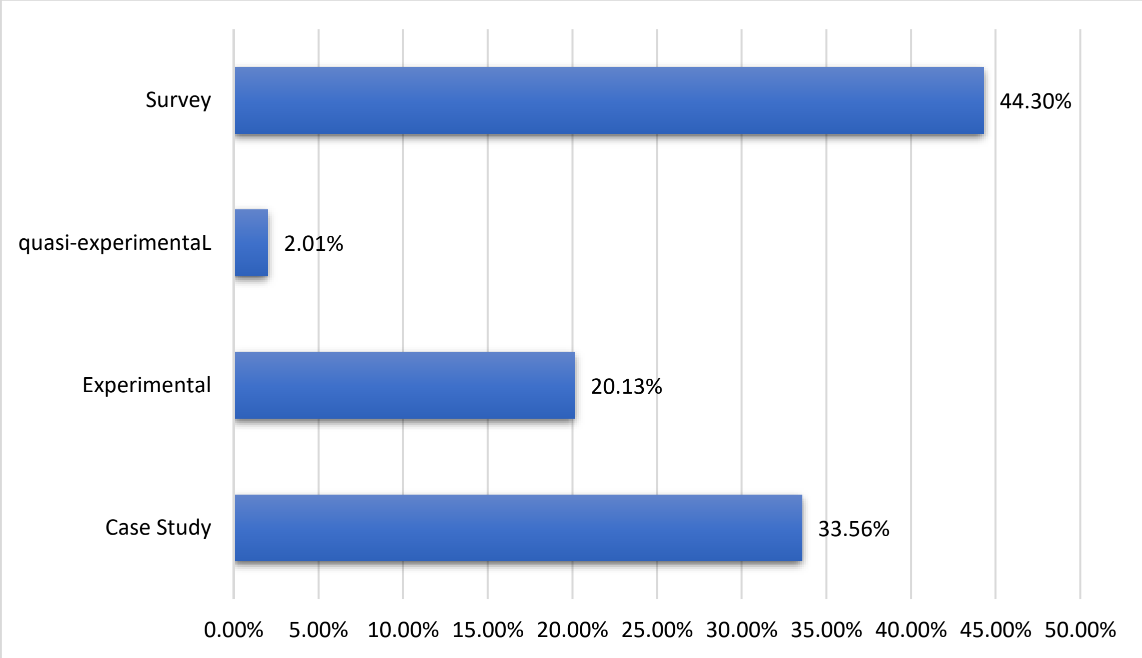


Figure 18. Distribution of Research Methods.

Figure 18 highlights that surveys are the predominant research method, comprising 44.30% of the studies. Surveys are efficient in collecting large volumes of data and can provide valuable insights into trends and perceptions within SMEs. However, they are prone to biases stemming from self-reporting and may struggle to capture complex interactions within organizational contexts. Case studies account for 33.56% of the research. They provide detailed insights and context-specific understanding, which can be beneficial in exploring the nuances of Big Data's influence on SME performance. However, case studies often lack generalizability, making it challenging to apply findings broadly across different contexts. Experimental designs represent 20.13% of the studies. These methods allow for establishing causal relationships and testing hypotheses under controlled conditions. However, the limited external validity can hinder the applicability of findings to real-world scenarios, as the results may not fully reflect the complexities of actual business environments. figure-experimental designs, making up a mere 2.01%, offer another avenue for investigating causal relationships but are similarly constrained by challenges in ensuring rigorous control over variables, which may lead to potential biases.

The data indicates a notable underutilization of alternative methodologies such as document analysis and statistical evaluations. This lack of diversity in research methods can limit the understanding of long-term trends and measurable impacts of Big Data on SME performance, as quantitative methods and historical data analysis provide valuable insights into patterns over time. The predominance of surveys and case studies suggests a reliance on context-specific data, which may introduce biases and limit the generalizability of the findings. To mitigate these risks, future research should consider employing a mixed-methods approach that incorporates a broader array of methodologies, including increased use of experimental designs and quantitative statistical evaluations. This multifaceted approach can enhance the robustness and credibility of conclusions drawn regarding the impact of Big Data on SMEs, ultimately leading to more actionable insights and recommendations for practitioners.

3.4. Results of Individual Studies.

Figure 19 reveals that the focus on various types of database systems is quite distinct, with a clear dominance in relational database technologies. A significant 48.99% of the studies concentrate on relational databases, NoSQL databases, making up 34.90% of the research cloud databases account for 14.77% of the research studies and High-Performance Computing (HPC) constitutes the smallest

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

portion of the research, at just 1.34%, but still represents a critical niche area. The distribution of research topics reflects the ongoing evolution in database technologies, with relational databases maintaining their foundational role, while NoSQL and cloud databases capture increasing interest due to their adaptability in modern data-intensive applications.

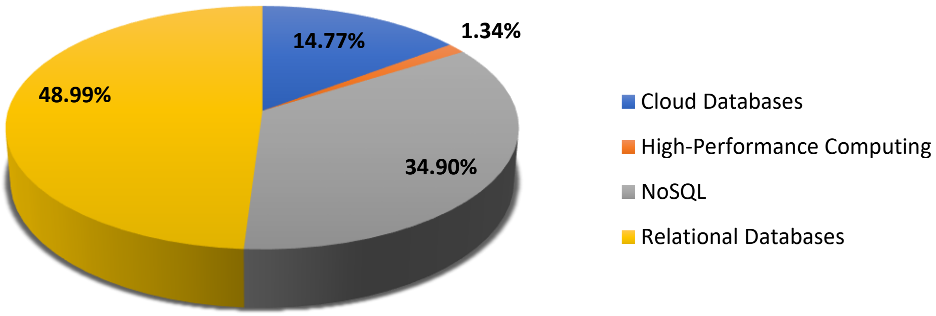


Figure 19. Database Technology types.

Figure 20 depicts that the majority of the studies, 48.32%, focus on data marts And with OLAP systems following closely, accounting for 25.50% of the studies, The ETL (Extract, Transform, Load) tools represent 24.16% of the research, underscoring their role in managing the data integration process. a very small portion of the research addresses CRM systems, data mining, and data warehousing directly, each contributing just 0.67%. This could indicate either a lack of focus on these topics within the reviewed research or that these areas are considered foundational, with most studies assuming their presence rather than focusing on them in detail.

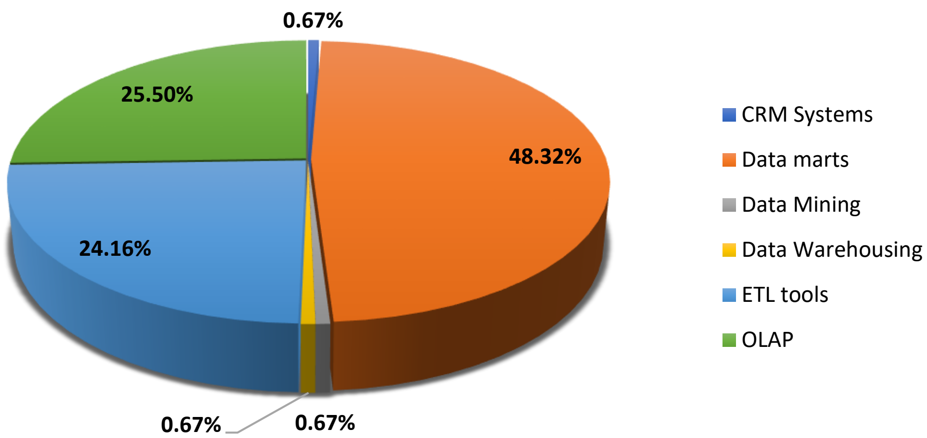


Figure 20. Data Warehouse Technology types.

3.5. Results of Synthesis

3.5.1. Characteristics and Risk of Bias among Contributing Studies

In the systematic review conducted, various data collection methods were employed to gather information from the analyzed papers. The results in Figure 21 indicate a diverse approach to data

 doi:10.20944/preprints202410.0059.v1

collection, with each method contributing uniquely to the overall findings. The most frequently utilized method was document analysis, accounting for 34.46% of the total data collection efforts. This suggests that a significant portion of the research relied on existing literature and documents, which is a common practice in systematic reviews as it allows researchers to synthesize previous findings and establish a comprehensive understanding of the topic at hand. Following document analysis, surveys emerged as another prominent method, representing 31.08% of the total data collection. Surveys are particularly valuable in gathering quantitative data from larger populations, enabling researchers to identify trends and patterns that may not be evident through qualitative methods alone. The integration of surveys into the systematic review highlights an effort to capture a broad range of perspectives and experiences related to the research question. Interviews and observations were also employed but to a lesser extent, with interviews comprising 17.57% and observations 16.89% of the total methods used. Interviews provide in-depth insights into individual experiences and opinions, allowing for a nuanced understanding of complex issues that may not be fully captured through quantitative measures alone. Observations, while less frequently used than other methods, offer direct insight into behaviors and interactions within natural settings, adding another layer of richness to the data collected.

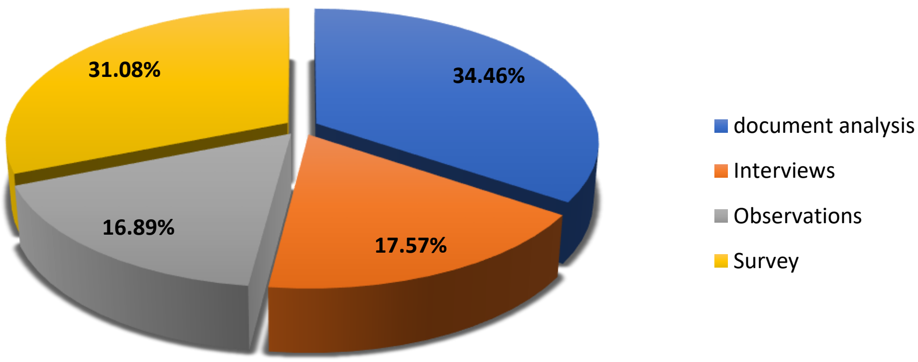


Figure 21. Data Collection methods**.**

3.5.2. Results of Statistical Synthesis.

As illustrated in Figure 22, the synthesis of included studies reveals a predominance of quantitative analysis methods, employed in 62.67% of the studies. This strong preference for numerical and statistical evaluations highlights a robust focus on assessing the impact of database and data warehouse technologies through measurable outcomes. In contrast, thematic analysis, which delves into qualitative patterns and themes, was utilized in 37.33% of the studies. This indicates a significant, though less predominant, role for qualitative insights in understanding the nuances of technology impacts.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

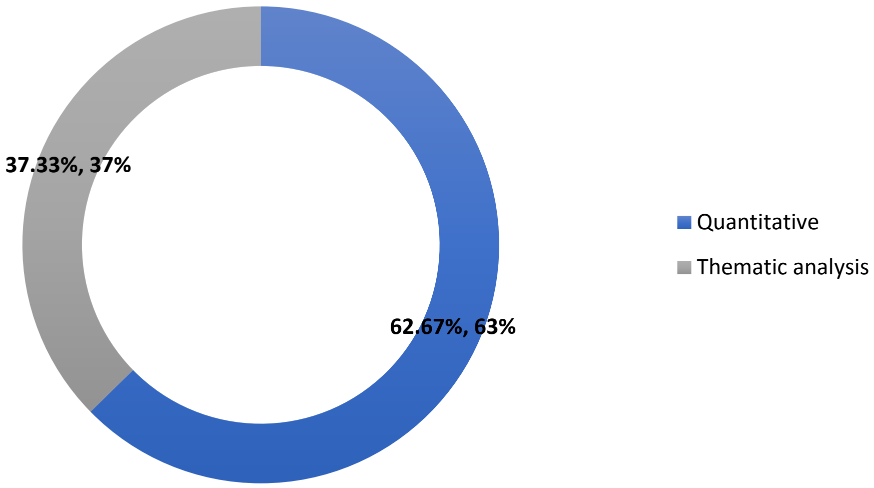


Figure 22. Data analysis techniques.

3.5.3. Investigating Heterogeneity

In the conducted analysis of various papers for the systematic review, the results reveal a significant distribution among four major technology providers: Microsoft SQL Server, IBM, AWS Redshift, and Oracle. Figure 23 indicates that Microsoft SQL Server holds a dominant position with a substantial 54% representation in the reviewed literature.

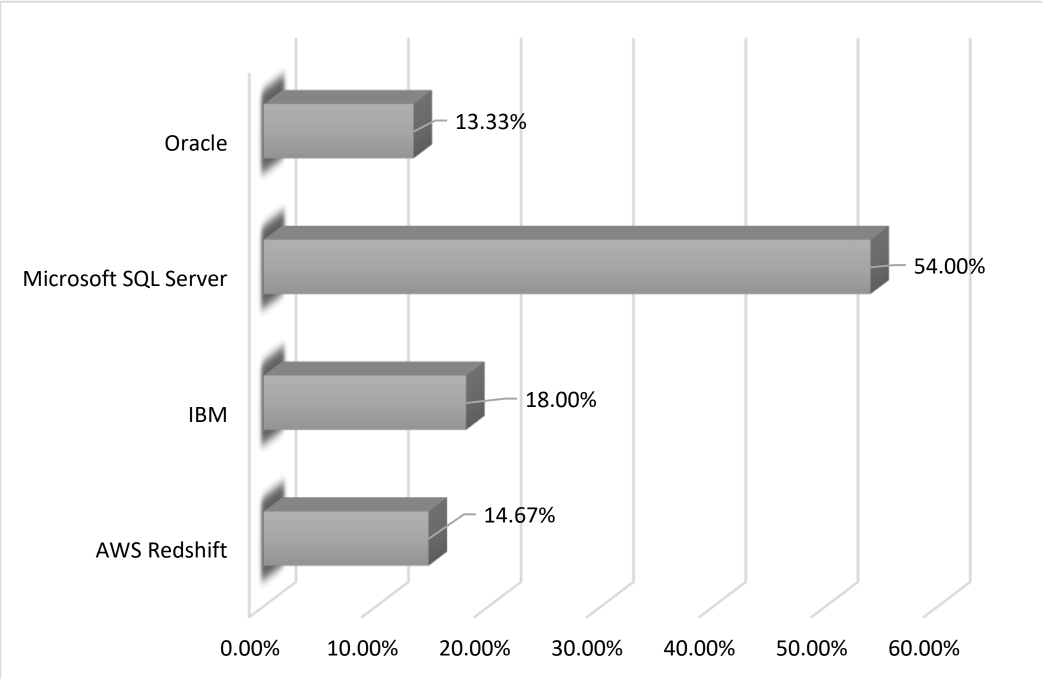


Figure 23. Analysis of Technology Providers.

This suggests that it is widely recognized and utilized within the context of the studies analyzed, potentially due to its robust features, extensive support, and integration capabilities within enterprise environments. Following Microsoft SQL Server, IBM accounts for 18% of the references found in the papers. This indicates a notable presence but suggests that it may not be as prevalent as Microsoft SQL Server in this body of work. IBM's offerings are often associated with advanced analytics and data management solutions, which could explain its relevance in specific contexts within the

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

systematic review. AWS Redshift has a representation of 14.67%, reflecting its growing adoption in cloud-based data warehousing solutions. As organizations increasingly migrate to cloud infrastructures for scalability and flexibility, AWS Redshift's role is likely to expand further in future research. Lastly, Oracle comprises 13.33% of the references noted in the analysis. While it has a smaller share compared to its competitors, Oracle remains a key player known for its comprehensive database solutions and enterprise-grade performance.

3.5.4. Sensitivity Analysis Results

In the context of the analysis conducted on various papers for a systematic review, the results indicate a significant distribution among small businesses, SMEs (Small and Medium Enterprises), and startups. Figure 24 indicates that small businesses account for approximately 36.67% of the total analyzed titles, while SMEs represent a larger portion at 43.33%. Startups, on the other hand, comprise 20.00% of the total count.

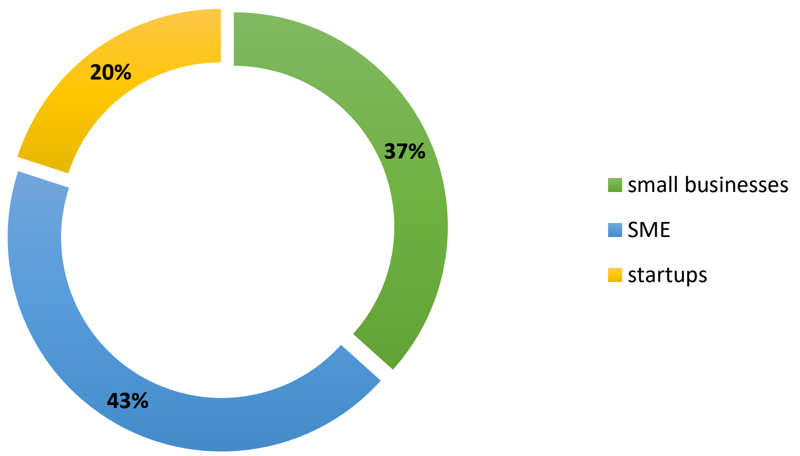


Figure 24. Analysis of Small Businesses, SMEs, and Startups.

This distribution highlights the prominence of SMEs in the industry context, suggesting that they play a crucial role in economic development and job creation. The categorization into small businesses, SMEs, and startups is essential for understanding their respective impacts on the economy. Small businesses often serve as the backbone of local economies by providing employment opportunities and fostering innovation at a community level. They are typically characterized by their limited scale in terms of revenue and number of employees but can be highly influential in niche markets. SMEs encompass a broader range of enterprises that may have more resources than typical small businesses but still operate below certain thresholds defined by national or international standards. Their larger share in this analysis (43.33%) underscores their importance in contributing to GDP and employment across various sectors. SMEs often benefit from government support programs aimed at enhancing competitiveness and sustainability. Startups represent a dynamic segment within this landscape, accounting for 20% of the analyzed titles. These entities are generally characterized by their innovative approaches and potential for rapid growth. However, they also face unique challenges such as securing funding and navigating market entry barriers. The relatively smaller percentage reflects both the high failure rate associated with startups as well as their emerging nature within specific industries. As depicted in figure 25, many of the studies, accounting for 44.67%, are centered around data warehouse systems, indicating a strong research interest in the design, implementation, and optimization of these systems.

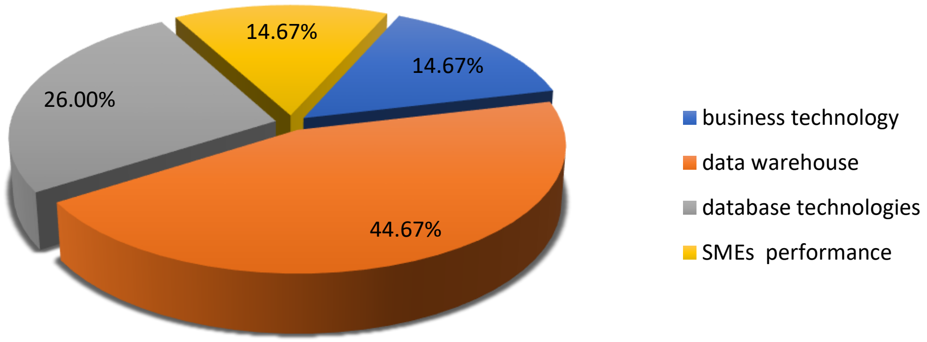


Figure 25. Fields of Study Distribution.

This reflects the growing importance of data warehouses in managing large-scale data and supporting advanced analytics in modern enterprises. Database technologies represent the second-largest category, comprising 26% of the collected research papers. This suggests that, while data warehouses are a key focus, there is still considerable interest in the foundational technologies that underpin data storage, retrieval, and management systems. Research in this area may explore advancements in database architectures, query optimization, and security features. Business technology and SMEs performance, both accounting for 14.67%, highlight a growing focus on the practical applications of these technologies within organizational and business contexts. These studies are likely exploring how innovations in database and data warehouse systems can improve overall business efficiency, decision-making processes, and the operational performance of small and medium-sized enterprises (SMEs).

3.6. Reporting Bias

Figure 26 offers graphical representation of the breakdown of study types and revealed a diverse methodological approach among the included researches. Of the studies reviewed, 39.01% employed quantitative methods, indicating a strong focus on numerical data and statistical analysis. Qualitative studies comprised 33.33% of the sample, highlighting a significant emphasis on in-depth, descriptive insights into organizational performance. Mixed-methods studies accounted for 27.66%, reflecting a balanced integration of both quantitative and qualitative approaches. This distribution underscores the comprehensive nature of the review, incorporating a range of perspectives and methodologies to assess the multifaceted impact of database and data warehouse technologies.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

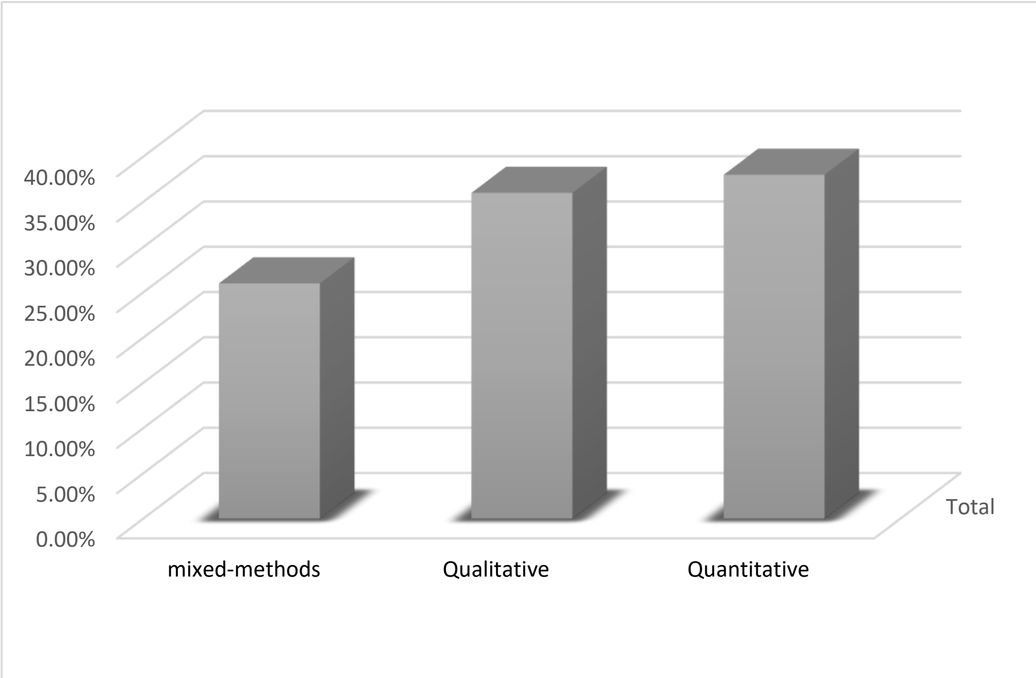


Figure 26. Study type distribution.

The risk of bias was assessed across these studies using established tools, such as the Newcastle-Ottawa Scale (NOS). The studies demonstrated varying levels of risk, with many quantitative studies having moderate to low risk due to their robust statistical methods, while qualitative studies often faced challenges related to subjectivity and bias in interpretation. As depicted in Figure 27, the synthesis of included studies reveals a predominance of quantitative analysis methods, employed in 62.67% of the studies.

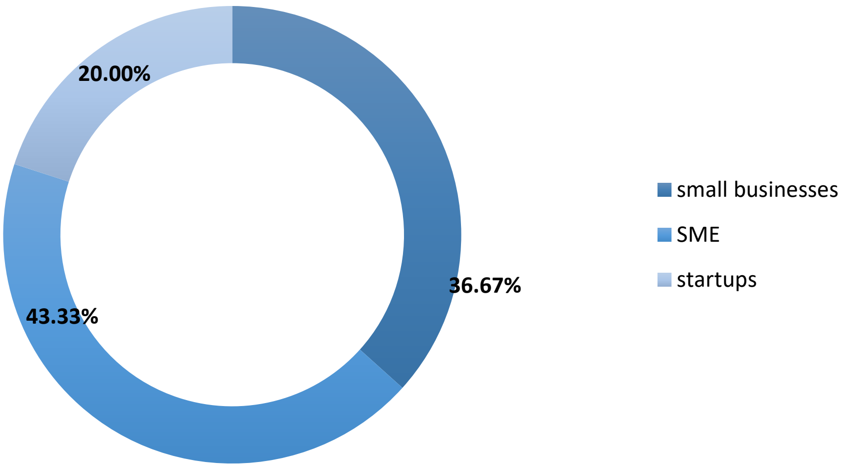


Figure 27. Data analysis techniques.

This strong preference for numerical and statistical evaluations highlights a robust focus on assessing the impact of database and data warehouse technologies through measurable outcomes. In contrast, thematic analysis, which delves into qualitative patterns and themes, was utilized in 37.33% of the studies. This indicates a significant, though less predominant, role for qualitative insights in understanding the nuances of technology impacts.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

3.7. Certainty of Evidence

Figure 28 shows the analysis of business performance metrics that reveals a notable emphasis on operational efficiency, which constitutes 47% of the evaluated metrics. Cost savings follow at 32%, while revenue growth is reported in 19% of the studies. Customer satisfaction, though important, is the least represented metric, accounting for only 1%. This distribution indicates that operational efficiency is the most frequently assessed outcome in relation to these technologies, highlighting its critical role in performance improvements. Conversely, the minimal focus on customer satisfaction suggests a potential area for further investigation to fully understand the breadth of impacts these technologies have on organizational performance.

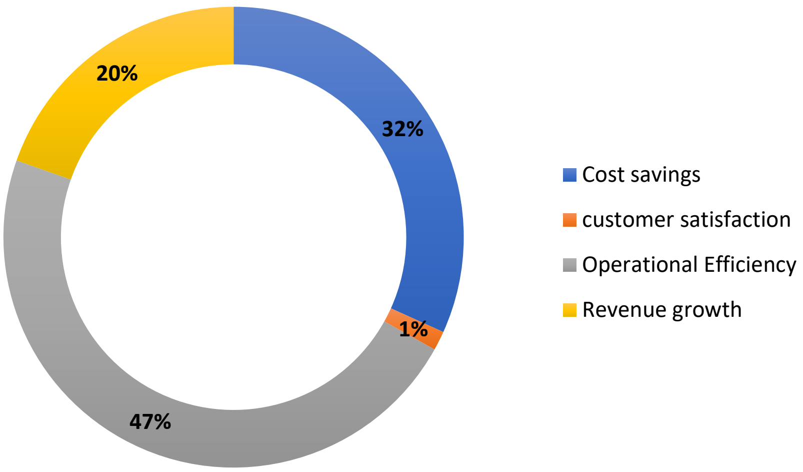


Figure 28. Business Performance metrics.

As shown in Figure 29, when it comes to IT performance, scalability is a big deal. In fact, it accounts for 38% of the metrics analyzed. This means that being able to scale up or down easily is super important.

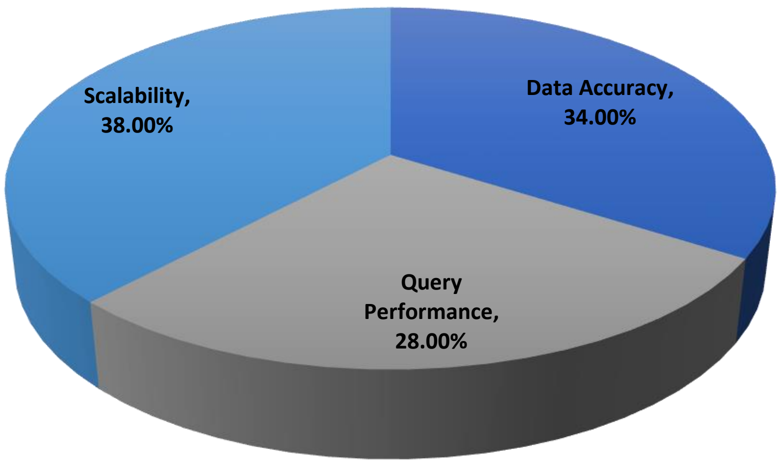


Figure 29. IT Performance Metrics.

Right behind scalability is data accuracy, which makes up 34% of the metrics. This shows just how crucial it is to make sure the information being used is correct and reliable. Query performance, at 28%, may not be as high on the list, but it's still essential. This metric looks at how well the technology can handle and process requests for data. IT performance is a complex issue that can have

 doi:10.20944/preprints202410.0059.v1

a big impact on how well an organization functions. By focusing on scalability, data accuracy, and query performance, businesses can ensure that their IT systems are running smoothly and effectively.

Data analysis techniques utilized across the included studies indicate a predominant use of quantitative analysis methods, employed in 62.67% of the studies. This approach suggests a strong preference for numerical and statistical evaluation of the technologies' impact. In contrast, thematic analysis, which focuses on identifying and interpreting patterns and themes within qualitative data, was used in 37.33% of the studies. These findings highlight a significant emphasis on quantitative methods for assessing performance impacts, while thematic analysis also plays a notable role in understanding the qualitative aspects of database and data warehouse technologies.

3.8. Key Findings and Strategic Implications for Business Leaders.

The systematic review highlights several critical insights regarding the impact of database and data warehouse technologies on business performance. These insights provide valuable strategic guidance for business leaders aiming to enhance their organization's efficiency, reduce costs, and gain a competitive edge. The study highlights the importance of data warehouse systems and relational databases in modern business for advanced analytics and operational efficiency. These systems centralize data sources, enabling better decision-making and access to high-quality data. The latest database technologies, like NoSQL and cloud databases, offer scalability and flexibility, enabling seamless operations and quick adaptation to data management strategies. Top IT performance indicators include scalability, data accuracy, and query performance, emphasizing the need for robust and timely data systems.

The research also established that geographical and organizational variability significantly influences the benefit levels derived from database and data warehouse technologies. Indeed, larger organizations with more sophisticated supporting technological infrastructures have seen overall better performance and increased decision-making capability. This points out the need for fitting approaches depending on the size and maturity of an organization. The small ones or those still at the initial stages of digitization would need more bespoke solutions that answer their specific business needs, while larger entities have the chance to reap more from matured systems to their benefit regarding data-driven decisions. The results show that database technologies are constantly evolving and hence businesses should adopt strategies that fit within their operational and technological landscape.

Strategic investments in data warehouse systems and modern technologies, such as NoSQL and cloud databases, offer scalable cost-effective solutions, reducing operational expenditure by optimizing storage. These technologies avoid redundancy and provide flexible pricing models based on actual usage, scaling infrastructure costs in direct relation to business growth. The same is further enhanced in operational efficiency through optimized relational databases and ETL tools that help expedite data processing and ensure accuracy for speedy decisions. Advanced analytics will help drive insight into customer behavior and market trends, thereby securing a competitive advantage by making better decisions and offering innovation to improve customer service. Moreover, scalability is important, whereby cloud-based and NoSQL databases guarantee flexibility and speed in terms of meeting the ever-growing demands of data storage and applying business for long-term success. Each category in this Table 12 offers strategic recommendations for business leaders looking to improve their operations.

Table 12. Summary of Strategic Opportunities for Business Leaders.

Category	Key Findings	Strategic Recommendations
Cost Efficiency	Cloud-based and NoSQL databases reduce capital expenditure and offer flexible, scalable pricing models.	Invest in scalable, pay-as-you-use technologies to minimize upfront costs and ensure financial flexibility.
Operational Benefits	Optimized databases and ETL tools streamline data processing,	Implement optimized data systems to enhance operational efficiency,

 doi:10.20944/preprints202410.0059.v1

	improving accuracy and reducing delays in information retrieval.	improve data accuracy, and speed up decision-making processes.
Customer Impact	Advanced analytics provide deeper insights into customer behavior, enabling personalized engagement and predictive analytics.	Use real-time data analytics to optimize customer experience, drive satisfaction, improve marketing strategies, and foster long-term loyalty.
Scalability	Scalable databases allow businesses to handle data growth without expensive infrastructure overhauls, adapting to changing demands.	Prioritize scalable solutions that adapt to data growth and market fluctuations without requiring disruptive upgrades.
Data Security	Data encryption and secure access controls safeguard sensitive customer and business data.	Implement robust security measures, including encryption, access control, and compliance with data protection regulations (e.g., GDPR, HIPAA) to mitigate risks.
Regulatory Compliance	Industry-specific regulations (e.g., GDPR, PCI DSS) require strict data handling protocols and reporting.	Ensure database systems are designed with compliance in mind, automating audits and incorporating security features to meet legal standards.
Data Integration	Seamless integration with other systems (e.g., CRM, ERP) enhances data flow and operational efficiency.	Opt for database solutions that easily integrate with existing software ecosystems to avoid data silos and streamline operations across departments.
Data Accessibility	Real-time access to data across multiple platforms improves decision-making and operational agility.	Implement database systems that ensure data is accessible across various devices and platforms to enhance remote work capabilities and on-the-go decision-making.
Disaster Recovery	Cloud-based solutions offer built-in redundancy and disaster recovery, minimizing data loss and downtime.	Invest in solutions that prioritize disaster recovery, ensuring business continuity with minimal data loss and downtime in the event of system failure.
Data Visualization	Advanced reporting and visualization tools enable more informed and quicker decision-making based on real-time insights.	Integrate data visualization and business intelligence tools to transform raw data into actionable insights, improving management's ability to make informed decisions.

Cost efficiency entails the enablement of scalable NoSQL and cloud technologies that reduce fixed costs, allowing companies to pay only for what they use. Operational benefits arise from optimized relational databases and ETL tools, in that these raise data processing speeds, improve accuracy, and further accelerate performance. Advanced analytics from data warehouses would also improve customer impact by showing rich insights into improving customer satisfaction and loyalty. Scalability ensures investment in adaptable solutions to handle the growth of data efficiently, thus enabling businesses to be agile and responsive to market demands. These findings help to align IT investments with key performance metrics, driving operational efficiency to competitive advantage.

3.9. Decision-Making Framework for Implementing Database and Data Warehouse Technologies.

The implementation of database and data warehouse technologies is a critical decision for small and medium enterprises (SMEs) looking to improve data management, optimize operations, and drive business intelligence. However, the decision-making process is highly variable, depending on the unique characteristics and constraints of each SME. Factors such as the size of the enterprise, industry, data complexity, growth plans, budget, and technological expertise all play significant roles in determining the most suitable solution.

For smaller enterprises, the need for cost-efficient, easy-to-manage systems often takes precedence, while medium-sized firms may require more robust, scalable solutions to handle larger data volumes and multi-functional business processes. Additionally, industry-specific requirements, particularly in regulated sectors like healthcare or finance, introduce further complexities such as data security, compliance, and integration challenges. In Table 13, we propose outlines of how these considerations influence decision-making, providing SMEs with a comprehensive framework to select the most appropriate database or data warehouse technology based on their individual needs.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

Table 13. Proposed Decision-making factors for implementing database and data warehouse technologies for different SMEs:.

Decision Factor	Small Enterprises	Medium Enterprises	Data-Intensive Businesses	Transactional Businesses	Regulated Industries	Budget-Conscious SMEs	Growth-Oriented SMEs	Low-Growth SMEs	Tech-Savvy SMEs	Non-Tech-Savvy SMEs	Global Operations	Local Operations
Size and Resources	Limited resources; prioritize cost-effective solutions	Can invest in more sophisticated systems	Needs advanced systems for data analytics and scalability	Simple systems; prioritize real-time transactional data	Must comply with strict regulatory standards	Open-source or cloud-based solutions	Invest in scalable technologies early on	Prioritize affordable, simple solutions	Can customize and manage complex systems	Needs vendor-supported, easy-to-use solutions	Solutions that support global compliance	Less emphasis on international regulations
Nature of Business	Focus on immediate operational needs	Broader, multi-departmental requirements	Data warehouse for analytics, reporting, AI integration	Database for fast transactional processing	High security and compliance requirements	Low-cost, basic systems for current needs	High scalability, with built-in growth support	Long-term cost-efficient system for stability	Opt for advanced, customizable systems	Choose cloud-based, managed services	Must support multi-location data storage and processing	Focus on local compliance
Data Volume and Complexity	Smaller data volumes; structured data	Larger data volumes, mix of structured	High data volume; requires data warehouse	Low-to-moderate data volume; standard	Data volume varies; strong governance	Small data volumes; prioritize affordability	Need scalable data architecture	Moderate data; no need for complex systems	Able to handle unstructured and complex data	Limited data management expertise	Solutions that support international data transfers	Prioritize local data governance and infrastructure

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

		/unstructured data		database sufficient	e is required							
Business Goals and Growth Plans	Immediate operational efficiency	Support multi-functional business operations	Use data for strategic decision-making	Focus on operational efficiencies	Must align with regulatory business goals	Immediate low-cost data solutions	Planning for future scalability and analytics	Long-term cost-efficiency	Plan for advanced business intelligence and analytics	Keep things simple with plug-and-play solutions	Align with multi-region strategic growth	Focus on regional or local performance
IT Infrastructure	Basic IT infrastructure	More developed IT systems; can handle complex setups	Need for data lakes and advanced processing platforms	Require simple, fast-response systems	Advanced IT infrastructure for security compliance	Basic infrastructure; prioritize managed services	Build scalable IT infrastructure	Existing systems meet operational needs	Advanced, in-house IT team	Outsource IT needs for cloud-based systems	Require international IT infrastructure	Focus on local infrastructure needs
Technology Selection	Open-source, affordable databases	Invest in robust, secure systems	Data warehouse for advanced analytics and scalability	Database for real-time operations	Comply with security regulations (e.g., GDPR, HIPAA)	Low-cost cloud solutions (e.g., AWS RDS)	Cloud-based, scalable solutions for growing data needs	Simple, low-maintenance systems	Prefer customizable, on-premise systems	Cloud-based solutions with minimal management	Need cloud solutions for international data compliance	On-premise or local cloud services

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

Scalability	Limited scalability needed	Scalable solutions for future growth	High scalability required	Low scalability; focus on current needs	Need scalable systems to handle regulatory reporting	Limited scalability; focus on immediate needs	High scalability to support future growth	No immediate need for scalability	Customizable and scalable on-premise or hybrid systems	Managed cloud services offering scalable solutions	High scalability for international growth	Moderate scalability for local growth
Cost and Licensing	Low-cost options, open-source, or subscription-based models	Can invest in premium solutions with additional features	Higher investment in analytics and reporting features	Affordable, standard licenses	Must budget for compliance and high security	Open-source or low-cost subscription models	Invest in advanced solutions for long-term ROI	Minimal upfront investment	Willing to invest in premium, customizable solutions	Subscription-based, pay-as-you-go pricing	International licensing and cost considerations	Focus on local pricing and affordability
Cloud vs. On-premise	Cloud-based, low-maintenance solutions	On-premise or hybrid systems	Cloud data warehouses for scalability and accessibility	On-premise or cloud databases	On-premise systems for data control and security	Cloud-based services to reduce infrastructure costs	Cloud-based for scalability; on-premise for control	On-premise for cost efficiency	Preference for on-premise solutions	Cloud-based solutions for ease of use and management	Cloud-based solutions for international compliance	On-premise or local cloud providers
Security	Basic security features	Enhanced security measures for data protection	Advanced security for sensitive,	Standard security protocols	High-level encryption and access controls	Sufficient basic security	Plan for advanced security as the	Moderate security measures to meet	Advanced, customizable	Cloud services with integrated	Must meet global security and compliance	Focus on regional security compliance

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

			large datasets				business grows	current needs	security features	security solutions	e regulations	
Compliance	Minimal compliance requirements	Compliance with general data protection laws	Must meet local and global compliance standards	Standard business data compliance	Strict compliance required (e.g., GDPR, HIPAA)	Limited compliance focus	Must plan for future compliance as the business grows	Focus on cost-efficient compliance measures	Fully compliant with international standards	Rely on vendor-provided compliance solutions	Must meet international data compliance regulations	Focus on local data protection laws
Integration with Systems	Minimal integration with existing systems	Integration with multiple internal systems	Must integrate with advanced BI, CRM, and ERP systems	Integration with basic business applications	Must integrate with secure, compliant systems	Simple integrations with existing platforms	Integration with advanced analytics and business tools	Limited integration with existing systems	Customizable integration capabilities	Plug-and-play integration with minimal setup	Complex integrations for global systems	Simple integration with local tools
IT Support and Expertise	Limited in-house IT expertise	In-house IT team to manage complex systems	Requires skilled IT staff for data warehouse management	Basic IT support for database maintenance	Requires specialized IT expertise for compliance	Minimal IT support; rely on vendor for service	Requires in-house or outsourced IT support	No need for extensive IT support	Skilled in-house team capable of managing systems	Outsource IT support to vendors	Requires international IT support structure	Focus on local IT support needs

The decision-making framework for SMEs highlights the dynamic interplay between technical, financial, and operational factors, underscoring that a one-size-fits-all approach is impractical. Different types of SMEs—whether small, medium-sized, data-intensive, or transactional—require tailored solutions that balance cost, complexity, and functionality. Budget-conscious SMEs often opt for cloud-based or open-source database solutions that minimize upfront costs, providing them with the flexibility to expand as they grow. These solutions are particularly attractive for enterprises with limited IT infrastructure, allowing them to focus on immediate business needs without investing in expensive, on-premises systems. In contrast, growth-oriented SMEs plan for scalability, choosing technologies that can evolve with their expanding data needs. This forward-thinking approach ensures that they are not burdened by costly migrations or system overhauls in the future.

For data-intensive businesses, the decision-making process revolves around the need for advanced analytics and large-scale data processing. Here, the focus shifts to data warehouse technologies that can integrate with existing business intelligence (BI) tools and handle complex queries, making real-time reporting and decision-making more efficient. These businesses prioritize performance, scalability, and integration capabilities, often choosing cloud-based data warehouses that provide flexibility and agility. Transactional businesses, on the other hand, require real-time processing and quick access to data, favoring databases that offer fast transaction speeds and reliable performance. Their focus is less on large-scale analytics and more on day-to-day operational efficiency. This distinction illustrates the importance of understanding the business's core data needs when selecting between a database and a data warehouse. Industry compliance is a pivotal factor for SMEs in regulated sectors, such as healthcare, finance, or legal services. These industries require robust data governance, security features, and adherence to strict regulatory frameworks like GDPR or HIPAA. SMEs in these sectors prioritize systems that can meet compliance demands without compromising operational flexibility, often leaning toward on-premises solutions that give them greater control over data security and privacy. Technological expertise also significantly impacts the decision-making process. SMEs with skilled IT teams may choose more complex, customizable systems that can be tailored to their specific requirements, while those lacking such expertise gravitate toward vendor-managed solutions that reduce the burden of system maintenance and technical oversight.

3.10. Best Practices for Successful Implementation of Database and DW Technologies

Implementing database and data warehouse technologies in SMEs is a strategic decision that requires careful consideration of each business's specific needs, resources, and long-term goals as proposed in Table 14. Whether a small enterprise with limited re-sources or a data-intensive business requiring advanced analytics, the selection and execution of these technologies must align with the operational and regulatory environment. Best practices such as clear goal setting, robust data governance, and security controls ap-ply broadly but are adapted differently based on factors like business size, industry, data complexity, and technical expertise. This customized approach ensures that the chosen systems not only meet current demands but also support future growth and scalability, performance tuning, security updates, and data backups. Finally, identification and selection of technologies that would scale the organization's growth ensure that the system does not face performance-related challenges with increased data volumes and user demands.

Table 14. Proposed Best Practices for Successful Implementation of Database and DW Technologies.

Best Practice	Small Enterprises	Medium Enterprises	Data-Intensive SMEs	Transactional SMEs	Regulated SMEs	Growth-Oriented SMEs	Low-Growth SMEs	Tech-Savvy SMEs	Non-Tech-Savvy SMEs

doi:10.20944/preprints202410.0059.v1

Define Business Requirements	Focus on immediate operational needs, ensure basic requirements are met	Comprehensive planning for broader business functions	Align with data analytics, AI, and reporting needs	Prioritize real-time transactional needs	Must ensure alignment with compliance and regulatory reporting	Long-term planning for growth and scalability	Keep the scope limited to current operational needs	Tailor requirements for advanced analytics and customization	Rely on simple, easy-to-understand business objectives
Data Governance Strategy	Focus on basic data quality and accessibility	Implement broader data governance, ensuring consistency across functions	High focus on data quality, especially for analytics	Prioritize transactional data accuracy and consistency	Strict governance to meet compliance requirements	Implement advanced governance for scalability and future growth	Basic governance focused on immediate needs	Advanced governance, integrating security and customization	Depend on vendor-provided governance frameworks
Technology Selection	Choose open-source or low-cost solutions	Invest in scalable, robust solutions	High-performance data warehouse for scalability and reporting	Prioritize simple database for fast transactions	Select technologies that ensure compliance (e.g., GDPR, HIPAA)	Choose flexible and scalable solutions for future growth	Low-cost, simple solutions for current needs	Prefer customizable, advanced technologies	Choose cloud-based, vendor-supported solutions
Data Architecture	Design for immediate	Plan for flexibility across multiple	Complex architecture	Simple, efficient architecture	Secure architecture ensuri	Scalable architecture to	Basic architecture design	Advanced, flexible	Vendor-managed

	needs, with minimal complexity	departments	supporting large-scale data processing	cture for operational data	ng compliance and privacy	accommodate growth	ed for stability	architecture	architecture for simplicity
Pilot Before Deployment	Conduct minimal pilot with small dataset	Pilot across departments to ensure integration	Full pilot with extensive datasets and use case simulations	Pilot for transactional consistency	Test pilot with strict compliance protocols	Pilot with growth potential in mind	Minimal pilot required to test basic functionality	Extensive pilot to validate customization	Run vendor-led pilot to ensure smooth transition
Security and Access Control	Basic security features for essential data	Implement role-based access across departments	Advanced security for sensitive and large datasets	Standard access control for operational data	Ensure strict compliance with data security laws	Plan for future security expansions	Basic security focusing on operational needs	Customize security features for specific user groups	Rely on vendor-provided security
Training and Support	Provide minimal training for basic system use	Broad training across functions	Comprehensive training for analytics and reporting	Focus training on operational efficiency	Extensive training focused on compliance	Provide ongoing training to adapt to growth	Basic training focusing on daily operations	Advanced training for technical staff and users	Focus on vendor-provided training
Performance Monitoring and	Monitor basic performance metric	Implement broader performance monitor	Continuously optimize for data processi	Focus on optimizing transac	Strict performance monitoring for	Implement scalable monitoring	Minimal monitoring for curren	Customizable performance metrics	Rely on vendor-managed

Optimization	s (uptime, response times)	ing and optimization	ng and analytics	tion speed	compliance	and optimization solutions	t needs	and optimization	performance monitoring
Plan for Scalability	Scalability is not a priority	Plan for future growth, ensuring flexibility	High scalability for data growth and increased data demand	Minimal need for scalability	Ensure scalability in line with regulatory reporting demands	Design scalable solutions for rapid business expansion	Scalability only to meet immediate operational needs	Advanced scalability features for future needs	Choose vendor-supported scalable solutions
Data Migration and Change Management	Minimal data migration requirements	Detailed migration plan for integrating various systems	Comprehensive migration strategy to avoid data loss	Basic migration strategy to minimize downtime	Strict migration controls to ensure compliance	Plan for future system upgrades during migration	Focus on minimizing disruption during migration	Detailed migration with extensive testing	Rely on vendor for migration manager

The variation in best practices across different types of SMEs underscores the im-portance of a tailored approach to database and data warehouse implementation. Small enterprises typically prioritize cost-efficiency and simplicity, while medium-sized and growth-oriented businesses focus on scalability and long-term flexibility. For da-ta-intensive companies, the ability to handle large-scale analytics and performance optimization is crucial, while transactional SMEs emphasize real-time processing capabilities. Regulated businesses must place security and compliance at the forefront of their decision-making, integrating strict data governance strategies. The technological expertise within the SME also plays a significant role; tech-savvy businesses are more likely to customize their systems, while less technical companies rely on vendor-managed solutions. This diversity highlights the need for SMEs to adopt a decision-making process that aligns with their specific operational context, ensuring that the chosen technology delivers both immediate value and future readiness.

3.11. Metrics and KIPs for Measuring Performance of Database and DW Technologies.

The performance of database and data warehouse technologies in SMEs can be optimized by using relevant metrics and KPIs that align with the specific characteristics of each business type as proposed in Table 15. These metrics help measure system reliability, scalability, data quality, security, and overall business value. As SMEs vary in size, growth potential, industry requirements, and technological capacity, the importance of these KPIs also shifts. Customizing the selection of performance indicators ensures that each SME can monitor and enhance its database and DW systems in a way that supports its operational and strategic goals.

Table 15. Proposed Metrics and KPIs for Measuring Performance of Database and DW Technologies based on different types of SMEs.

Category	Metric/KPI	Description	Small Enterprises	Medium Enterprises	Data-Intensive SMEs	Transactional SMEs	Regulated SMEs	Growth-Oriented SMEs	Low-Growth SMEs	Tech-Savvy SMEs	Non-Tech-Savvy SMEs
System Performance	Query Response Time	Measures the time taken to retrieve data from queries. Shorter times indicate better performance.	Moderate relevance	High relevance	Critical	Highly relevant	Relevant	Necessary	Low relevance	High relevance	Vendor-managed
	Data Load Time	Time taken to load data into the system.	Moderate relevance	Important	Critical for ETL	Relevant	Critical	Important as data grows	Low relevance	Important	Vendor-managed
	Uptime/Downtime	Percentage of time the system is	Important	Critical	Essential	Highly relevant	Highly critical	Key for growth	Necessary	Essential	Vendor-managed

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

		operational.									
	Throughput	Rate at which data is processed, measured in transactions per second (TPS).	Moderate relevance	Necessary	Critical	Highly relevant	Relevant	Important	Low relevance	Critical	Vendor-managed
	Concurrency	Number of users/processes that can access the system simultaneously without performance degradation.	Low relevance	Relevant	Critical	Relevant	Critical	Important	Low relevance	Critical	Vendor-managed
Data Quality	Data Accuracy	Measures the correctness of data compared to real-	Low relevance	Important	Highly relevant	Critical	Critical	Highly important	Low relevance	Critical	Vendor-managed

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

	Data Completeness	world values. Percentage of data fields filled without missing values.	Moderate relevance	Important	Critical	Relevant	Essential	Important	Low relevance	Critical	Vendor-managed
	Data Consistency	Ensures data remains consistent across different systems or database replicas.	Low relevance	Necessary	Highly relevant	Low relevance	Critical	Important	Low relevance	Critical	Vendor-managed
	Error Rate	Percentage of errors encountered during data processing or loading.	Low relevance	Important	Critical	Moderate relevance	Critical	Important	Low relevance	Critical	Vendor-managed
Scalability and	System Scalability	Ability to handle	Low relevance	Relevant	Essential	Low relevance	Critical	Critical	Low rele	Critical	Vendor-man

Flexibility		increased data volume or users without performance loss.							vance		aged
	Capacity Utilization	Measures how much of the system's capacity (storage or processing power) is being used.	Low relevance	Important	Critical	Low relevance	Relevant	Critical	Low relevance	Critical	Vendor-managed
	Elasticity	Ability to dynamically adjust resources (compute, storage) based on workload	Low relevance	Important	Essential	Low relevance	Critical	Critical	Low relevance	Critical	Vendor-managed

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

Security and Compliance	Data Security Incidents	(especially in cloud). Tracks number of security incidents, such as unauthorized access or breaches.	Moderate relevance	Relevant	Important	Moderate relevance	Highly critical	Important	Low relevance	Critical	Vendor-managed
	Access Control Violations	Measures unauthorized access attempts.	Low relevance	Important	Relevant	Relevant	Critical	Important	Low relevance	Critical	Vendor-managed
	Encryption Compliance	Percentage of data encrypted at rest and in transit.	Low relevance	Important	Important	Moderate relevance	Critical	Important	Low relevance	Critical	Vendor-managed
Operational Efficiency	ETL Process Efficiency	Measures the efficiency of ETL processes (Extract,	Moderate relevance	Important	Critical	Moderate relevance	Relevant	Important	Low relevance	Critical	Vendor-managed

		Transform, Load).									
	Cost per Query	Measures the cost of executing queries in cloud environments.	Low relevance	Important	Critical	Moderate relevance	Relevant	Critical	Low relevance	Critical	Vendor-managed
	Resource Utilization	Percentage of resources (CPU, memory, storage) being used.	Low relevance	Important	Critical	Moderate relevance	Important	Critical	Low relevance	Critical	Vendor-managed
Business Value	Time to Insight	Measures how quickly the business can retrieve and analyze data to generate insights.	Moderate relevance	Important	Highly critical	Moderate relevance	Relevant	Critical	Low relevance	Critical	Vendor-managed

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

Return on Investment (ROI)	Financial return vs. investment in database or DW technology. Positive ROI shows value.	Moderate relevance	Important	Critical	Moderate relevance	Important	Critical	Low relevance	Critical	Vendor-managed
User Satisfaction	Feedback from users on system performance, ease of use, and value.	Relevant	Important	Critical	Relevant	Critical	Important	Low relevance	Critical	Vendor-managed

The diversity of SMEs requires a tailored approach to measuring database and data warehouse performance. Small enterprises benefit most from focusing on simplicity and basic operational efficiency metrics, such as uptime and resource utilization, while me-dium-sized businesses must prioritize scalability, data accuracy, and concurrency to manage increasing data volumes. Data-intensive SMEs rely heavily on ETL efficiency and system performance metrics, as their operations depend on handling large datasets effec-tively. Transactional SMEs need real-time metrics like query response time and through-put to ensure seamless operations, while regulated SMEs prioritize security and compli-ance KPIs to meet industry standards. Growth-oriented businesses emphasize scalability, ensuring their systems can expand as their needs evolve, while low-growth SMEs focus on maintaining stability with basic performance indicators. Tech-savvy SMEs take a more sophisticated approach, tracking advanced KPIs such as elasticity and resource utiliza-tion, while non-tech-savvy SMEs rely on vendor-managed solutions, focusing on vendor uptime and security compliance. These distinctions highlight the need for a customized, KPI-driven strategy to optimize database and data warehouse performance in SMEs.

3.12. Customizing The Database and Data Warehouse Technologies for Different SME Industries.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

The selection and customization of database and data warehouse technologies for SMEs vary significantly depending on the industry. Different industries have unique op-erational requirements, data complexities, compliance needs, and customer engagement models. As a result, businesses in sectors like healthcare, finance, retail, and manufactur-ing need to adopt tailored technology solutions that address their specific data manage-ment, security, and scalability needs. By choosing the appropriate database and data warehouse technologies, SMEs can enhance their efficiency, gain actionable insights, and remain competitive in their respective markets using the proposed Table 16.

Table 16. Proposed Customization of Database and Data Warehouse Technologies for Different SME Industries.

SME Industry	**Database Technologies**	**Data Warehouse Technologies**	**Key Features and Customization**	**Key Benefits**
Retail	Relational Databases (e.g., MySQL, PostgreSQL)	Cloud Data Warehouses (e.g., Google BigQuery, Amazon Redshift)	Focus on high-volume transaction processing, customer data, and sales trends. Integration with e-commerce platforms and POS systems.	Real-time transaction tracking, inventory management, customer behavior insights.
	NoSQL Databases (e.g., MongoDB)	Hybrid Data Warehouses	For handling unstructured data such as customer reviews, social media data, and dynamic content.	Improved customer experience through personalized marketing.
Healthcare	Relational Databases with Compliance (e.g., Oracle, SQL Server)	On-Premise Data Warehouses (e.g., Teradata, SAP BW)	Compliance with healthcare regulations (e.g., HIPAA, GDPR). Focus on patient data security, medical records, and interoperability.	Secure and compliant storage of sensitive patient data, quick retrieval for medical decisions.
	HL7-Compliant Databases	Cloud Hybrid Solutions	For integrating clinical systems with non-clinical (e.g., billing, patient feedback). Ensures scalability without compromising security.	Seamless integration of clinical and business systems with secure scalability.
Finance & Banking	Relational Databases with Encryption (e.g., SQL Server, Oracle)	Enterprise-Grade Data Warehouses (e.g., Snowflake, Azure Synapse)	High-security, real-time data processing for transactions and fraud detection. Compliance with regulations (e.g., PCI DSS, GDPR).	Enhanced financial reporting, fraud detection, and regulatory compliance.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

	Blockchain-Based Databases	Distributed Data Warehouses	For secure, tamper-proof transaction records and ensuring auditability.	Improved transparency and fraud prevention through immutable records.
Manufacturing	Industrial IoT Databases (e.g., InfluxDB, TimescaleDB)	Cloud Data Warehouses (e.g., AWS Redshift, Google BigQuery)	Focus on production data, machine telemetry, and real-time equipment monitoring. Data integration from sensors and automation systems.	Predictive maintenance, optimized production lines, and cost reductions.
	Relational Databases (e.g., MySQL, PostgreSQL)	On-Premise Data Warehouses	For managing supply chain, inventory, and operational data across factories.	Efficient supply chain management, real-time inventory tracking.
E-commerce	Relational Databases (e.g., MySQL, PostgreSQL)	Cloud Data Warehouses (e.g., Snowflake, Redshift)	Handling large-scale product catalogs, customer transactions, and web analytics. Real-time reporting on sales and customer engagement.	Personalized marketing, enhanced sales forecasts, and real-time customer insights.
	Document-Based Databases (e.g., MongoDB)	Hybrid Cloud Data Warehouses	Storing unstructured data such as user-generated content, reviews, and multi-channel data integration.	Multi-channel sales data integration and enhanced customer experience.
Logistics & Supply Chain	Relational Databases with Real-Time Analytics (e.g., PostgreSQL)	Cloud Data Warehouses with AI Integration (e.g., Google BigQuery)	Focus on real-time tracking, shipment data, and demand forecasting. Integration with GPS systems and third-party logistics providers.	Real-time shipment tracking, enhanced demand forecasting, and optimized route planning.
	Graph Databases (e.g., Neo4j)	Distributed Data Warehouses	For mapping complex relationships between suppliers, routes, and inventory nodes.	Better route optimization and supplier management through relationship mapping.

 doi:10.20944/preprints202410.0059.v1

Education (Higher Ed)	Relational Databases (e.g., SQL Server, PostgreSQL)	Cloud Data Warehouses with BI Tools (e.g., Azure Synapse, AWS Redshift)	Manage student records, faculty data, and research outputs. Integration with LMS (Learning Management Systems) and ERP systems.	Improved student data tracking, better reporting on performance, and scalable infrastructure for research data.
	NoSQL Databases for Unstructured Data	Hybrid Solutions	Manage multimedia files, e-learning resources, and unstructured student feedback.	Enhanced digital learning experiences and efficient resource management.
Hospitality & Tourism	Relational Databases (e.g., MySQL, MariaDB)	Cloud Data Warehouses (e.g., Snowflake, Google BigQuery)	Focus on managing customer bookings, loyalty programs, and occupancy rates. Integration with third-party booking platforms and CRM systems.	Real-time booking data, personalized marketing, and improved customer satisfaction.
	Graph Databases for Customer Relationship Management	Hybrid Cloud Solutions	Manage customer relationships, preferences, and booking histories.	Enhanced customer loyalty programs and personalized experiences.
Energy & Utilities	Time-Series Databases (e.g., InfluxDB, TimescaleDB)	Cloud Data Warehouses for IoT Data (e.g., AWS Redshift, Azure Synapse)	Real-time monitoring of energy consumption, grid performance, and predictive maintenance for infrastructure.	Improved energy efficiency, predictive maintenance, and reduced downtime.
	Relational Databases for Billing & Regulatory Compliance	On-Premise Data Warehouses for Compliance	Managing customer billing, compliance reporting, and integrating IoT data from smart meters.	Accurate billing, energy usage monitoring, and compliance with energy regulations.
Telecommunications	Relational Databases for Subscriber Management (e.g., Oracle)	Cloud Data Warehouses with AI (e.g., Google BigQuery, Snowflake)	Handling subscriber data, call records, network data, and customer support. Real-time data for fraud	Real-time fraud detection, improved network performance, and customer churn analysis.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

			detection and customer behavior.	
	Graph Databases for Network Analysis	Hybrid Data Warehouses	Mapping network connections and relationships between devices and services.	Optimized network performance and reduced downtime through predictive analytics.
Professional Services	Relational Databases (e.g., SQL Server, MySQL)	Cloud Data Warehouses with BI Tools (e.g., Azure Synapse)	Focus on managing customer projects, billing data, and service contracts. Integration with project management tools and CRMs.	Efficient project tracking, real-time billing insights, and better client management.
	Document-Based Databases (e.g., MongoDB)	Hybrid Data Warehouses	Storing client documents, contracts, and unstructured communication data.	Centralized client document management and faster service delivery.
Media & Entertainment	NoSQL Databases for Content Delivery (e.g., Cassandra)	Cloud Data Warehouses for Streaming Data (e.g., Snowflake)	Managing large volumes of streaming data, content metadata, and user preferences. Integration with recommendation engines.	Enhanced user experience through personalized content recommendations and faster content delivery.
	Graph Databases for Social Media Integration	Hybrid Solutions for Analytics	Handling relationships between users, content, and social interactions.	Improved social media engagement, targeting, and content performance analysis.

The customization of database and data warehouse technologies for different indus-tries highlights the diverse data management challenges SMEs face. Retail and e-commerce SMEs focus on real-time data analytics to track sales and customer behaviors, benefiting from cloud-based solutions for scalability and cost-efficiency. Healthcare and finance SMEs, on the other hand, prioritize security and regulatory compliance, opting for on-premises or hybrid solutions that safeguard sensitive data. Manufacturing and logistics SMEs require integration with IoT systems to monitor real-time data from equipment and supply chains, enabling predictive maintenance and operational optimization. Similarly, media and telecommunications SMEs rely heavily on NoSQL and graph databases to handle vast amounts of unstructured and streaming data, driving customer engagement and content personalization. Each industry's approach to selecting and customizing these technologies ensures that SMEs can leverage their data for improved decision-making, enhanced customer experiences, and streamlined operations.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

3.13. Proposed Industry-Specific Frameworks for Database and Data Warehouse Technologies

Table 17 customizes the framework for various industries, covering the essential considerations for selecting and deploying database and data warehouse technologies, such as data complexity, compliance needs, performance requirements, and scalability. Each industry has unique data needs, compliance standards, and operational goals, which are reflected in this customized framework.

Table 17. Proposed Industry-Specific Frameworks for Database and Data Warehouse Technologies.

Industry	Data Complexity	Compliance Requirements	Database Technologies	Data Warehouse Technologies	Key Performance Requirements	Scalability Needs	Data Integration Needs	Use Cases
Retail	Moderate to High	Consumer Protection (GDPR, CCPA)	Relational Databases (MySQL, PostgreSQL)	Cloud DW (Google BigQuery, Snowflake)	High transaction throughput, real-time analytics, and low-latency processing for customer interactions	Requires horizontal scalability to handle seasonal spikes and growing customer data	Integration with e-commerce platforms, CRM, ERP, and POS systems	Inventory management, personalized marketing, real-time customer insights, sales reporting
			NoSQL Databases (MongoDB)	Hybrid DW for Structured & Unstructured Data	Handles unstructured data (e.g., product reviews, social media). Supports multi-channel sales.	Scalability required for large product catalogs and customer personalization	Multi-channel data (web, mobile, social media) and IoT integration	Customer experience optimization, product recommendations, demand forecasting
Healthcare	High	HIPAA, GDPR, POPIA	Relational Databases with Compliance	On-Premise DW (SAP BW,	High data security, rapid access to medical records,	Vertical scalability to handle high data	Integration with EHR (Electronic Health Records),	Patient data management, medical record integration

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

			(Oracle, SQL Server)	Teradata)	low-latency for clinical decision-making	growth, while ensuring compliance and data security	PACS, billing systems, and external labs	, compliance reporting, predictive healthcare analytics
			HL7-Compliant Databases	Cloud Hybrid Solutions	Seamless integration across clinical and non-clinical systems. Ensures scalability and compliance.	Needs secure scalability, especially when handling telemedicine and digital health records	Medical device data, lab systems, clinical research systems	Secure data sharing, cross-hospital collaboration, real-time diagnostic support
Finance & Banking	High	PCI DSS, GDPR, AML Regulations	Relational Databases (SQL Server, Oracle)	Cloud DW (Snowflake, AWS Redshift)	Real-time transaction processing, fraud detection, high availability, and financial reporting	Elastic scalability for rapid data growth and real-time fraud detection	Integration with payment systems, CRM, AML systems, and regulatory bodies	Fraud detection, real-time transaction analysis, compliance reporting, customer insights
			Blockchain Databases	Distributed DW (Azure Synapse)	Tamper-proof, secure transactions for auditing and compliance	Requires horizontal scaling for financial transactions across multiple global regions	Multi-system integration for trading platforms, core banking systems, and mobile payments	Audit trails, cross-border payments, secure financial transactions

 doi:10.20944/preprints202410.0059.v1

Manufacturing	High	Industry Standards Compliance	Industrial IoT Databases (InfluxDB, TimescaleDB)	Cloud DW (AWS Redshift, Snowflake)	Real-time data processing from IoT sensors, predictive maintenance, low-latency analytics	Scalability for monitoring multiple production lines, large-scale telemetry, and sensor data	Integration with ERP, MES (Manufacturing Execution System), IoT sensors, and SCADA systems	Predictive maintenance, equipment monitoring, real-time supply chain tracking, production optimization
			Relational Databases (PostgreSQL, MySQL)	On-Premise DW for Secure Operations	Stable performance for managing production schedules, inventory, and supply chain	Requires scalability to handle growing production data and supply chain complexity	Seamless integration between factories, suppliers, logistics providers, and automated systems	Supply chain management, just-in-time inventory, production forecasting
E-commerce	High	Consumer Protection (GDPR, CCPA)	Relational Databases (PostgreSQL, MySQL)	Cloud DW (Snowflake, Google BigQuery)	High-volume transactions, real-time product recommendations, low-latency customer behavior analysis	Horizontal scalability to handle high traffic surges, especially during sales or peak seasons	Integration with CRM, web platforms, mobile apps, social media, and payment gateways	Personalized marketing, sales forecasting, customer segmentation, product catalog management

doi:10.20944/preprints202410.0059.v1

			Document-Based Databases (MongoDB)	Hybrid DW	Unstructured data handling, product review analysis, and multi-channel data processing	Scalability required for expanding product catalogs and managing user-generated content	Multi-channel sales data, integration with marketing and loyalty platforms	Cross-channel customer behavior tracking, dynamic pricing, demand forecasting
Logistics & Supply Chain	Moderate to High	Trade and Export Regulations	Relational Databases (PostgreSQL, Oracle)	Cloud DW (Google BigQuery, AWS Redshift)	Real-time shipment tracking, inventory management, demand forecasting, and low-latency processing	Scalability required for global operations and high-volume shipment tracking	Integration with GPS, RFID, WMS (Warehouse Management Systems), and third-party logistics providers	Real-time shipment tracking, route optimization, inventory management, demand forecasting
			Graph Databases (Neo4j)	Distributed DW	Mapping complex relationships between suppliers, products, and customers. Optimized route planning.	Needs flexible scalability to handle global supply chain operations and multiple distribution centers	Integration across multiple suppliers, logistics partners, and warehousing systems	Optimized route planning, supply chain transparency, inventory optimization

 doi:10.20944/preprints202410.0059.v1

Education (Higher Ed)	Moderate to High	FERPA, GDPR	Relational Databases (PostgreSQL, SQL Server)	Cloud DW (Azure Synapse, AWS Redshift)	Managing student records, research data, and faculty information. Compliance with education regulations	Horizontal scalability to manage fluctuating student enrollments and research data	Integration with LMS (Learning Management Systems), ERP, research management systems	Student data tracking, institutional performance analytics, research data management
			NoSQL Databases for Multimedia Data	Hybrid Solutions for Unstructured Data	Handling multimedia files, digital resources, and student feedback for e-learning platforms	Needs scalability to accommodate growing multimedia resources and digital learning platforms	Integration with digital libraries, video conferencing tools, and online learning platforms	Multimedia content management, personalized learning experiences, digital resource management
Hospitality & Tourism	Moderate	GDPR, PCI DSS	Relational Databases (MySQL, MariaDB)	Cloud DW (Google BigQuery, Snowflake)	Managing bookings, customer preferences, and loyalty programs with real-time data insights	Scalability needed to handle peak booking seasons, loyalty programs, and multiple properties	Integration with booking engines, CRM systems, third-party travel sites, and payment gateways	Real-time booking management, personalized guest experiences, loyalty program management

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

			Graph Databases for Customer Relationships	Hybrid Cloud Solutions	Managing guest preferences, booking histories, and customer relationships. Optimized marketing.	Scalability needed for global chains and handling large customer datasets	Seamless integration with CRM, property management systems (PMS), and booking platforms	Customer relationship management, loyalty tracking, personalized experiences
Energy & Utilities	High	Regulatory Compliance (ISO, NERC, GDPR)	Time-Series Databases (InfluxDB, TimescaleDB)	Cloud DW for IoT Data (AWS Redshift)	Real-time monitoring of energy usage, grid performance, predictive maintenance, and customer billing	Horizontal scalability to handle massive IoT data from smart meters and energy grids	Integration with smart meters, grid monitoring systems, billing platforms, and regulatory reporting	Smart grid monitoring, predictive maintenance, energy usage optimization, compliance reporting
			Relational Databases for Billing Systems	On-Premise DW for Regulatory Compliance	Secure and compliant management of customer billing data and energy usage records	Scalability needed for expanding smart grid deployments and customer bases	Integration with smart meters, energy storage systems, renewable energy sources	Accurate billing, grid efficiency management, regulatory reporting
Telecommunications	High	FCC, GDPR	Relational Databases for Subscriber	Cloud DW with AI Integration (Google	High availability, real-time analytics for	Elastic scalability to handle increasing data	Integration with CRM, billing systems, network	Real-time fraud detection, network performance

			Data (Oracle, SQL Server)	BigQuery)	network performance, fraud detection, and customer behavior	volumes, growing user bases, and IoT device management	infrastructure, and customer support systems	monitoring, customer churn analysis
			Graph Databases for Network Relationships	Hybrid DW for Real-Time Analysis	Optimizing network connections and relationships between devices, users, and services	Scalability required to handle growing IoT device ecosystems and customer bases	Integration with network monitoring systems, IoT devices, and 5G infrastructure	Predictive network management, device performance optimization, customer engagement analytics
Professional Services	Moderate to High	GDPR, Data Protection Laws	Relational Databases (SQL Server, MySQL)	Cloud DW with BI Tools (Azure Synapse)	Managing client projects, billing, and service contracts with real-time visibility and reporting	Horizontal scalability to handle multiple client projects and large data volumes	Integration with CRM, project management tools, and client portals	Project tracking, real-time billing insights, client performance management
			Document-Based Databases (MongoDB)	Hybrid DW	Storing client documents, contracts, and communication data for unstructur	Scalability needed for growing client documentation and long-	Integration with document management systems, CRMs, and client	Document management, service delivery optimization, client collaboration

 doi:10.20944/preprints202410.0059.v1

					ed client management	term project archiving	collaboration tools	
Media & Entertainment	High	GDPR, Copyright Laws	NoSQL Databases for Content Delivery (Cassandra)	Cloud DW for Streaming Data (Snowflake)	Managing high-volume streaming data, content metadata, and user preferences with real-time insights	Elastic scalability to handle high traffic, especially during content releases and live events	Integration with content distribution networks (CDNs), recommendation engines, and social media	Content personalization, user engagement analysis, streaming performance monitoring
			Graph Databases for Social Media Integration	Hybrid Solutions for Real-Time Data	Managing relationships between users, content, and social interactions. Optimizing engagement.	Scalability required for expanding user bases and content distribution channels	Integration with social media platforms, user-generated content, and recommendation engines	Social media engagement, content targeting, real-time feedback analysis

Each industry faces unique challenges when it comes to data management, compliance, and performance requirements. Retail and e-commerce industries rely heavily on scalable cloud data warehouses integrated with customer interaction data and real-time analytics to optimize customer experience and inventory management. Healthcare and finance, with their stringent compliance regulations, require highly secure, encrypted database technologies paired with scalable on-premise or cloud-based warehouses to ensure data privacy and regulatory compliance. Manufacturing and logistics sectors prioritize real-time data processing from IoT sensors and predictive analytics for operational optimization. In contrast, media and entertainment businesses focus on high-performance databases capable of handling unstructured streaming data and user preferences, ensuring seamless content delivery and personalization.

a. Real Case Studies and How They Relate to the Database and Data Warehouse Performance

The application of database and data warehouse technologies in small and medium-sized enterprises (SMEs) has been explored through several real-world case studies, which provide insights into the practical implementation and performance of these systems. These cases reinforce the

 doi:10.20944/preprints202410.0059.v1

findings of the systematic review by highlighting how the adoption of these technologies contributes to operational efficiency, scalability, and compliance. Through the examination of both successes and challenges, the case studies offer a comprehensive understanding of the technological landscape for SMEs, aligning with broader academic insights. One case study involves a small retail company that successfully integrated a cloud-based data warehouse to manage real-time sales and inventory data. Before the implementation, the company faced challenges with data fragmentation and slow retrieval speeds, which hampered decision-making processes. After adopting a cloud solution, they were able to centralize data from multiple sources, enabling faster data access and more accurate sales forecasting. The scalability of the cloud system allowed the company to handle seasonal peaks in demand without significant infrastructure upgrades. This example directly links to the findings of the systematic review, which emphasize the importance of performance and scalability in retail, particularly for real-time analytics.

Another case study focuses on a healthcare provider that adopted a HIPAA-compliant data warehouse to manage electronic medical records (EMR). Prior to implementation, the provider struggled with regulatory compliance and data security issues, particularly regarding patient confidentiality. The introduction of encrypted data storage and secure access protocols resolved these concerns, while the system's predictive analytics capabilities improved patient care by identifying at-risk individuals early. This case mirrors the systematic review's findings that highlight data sensitivity and regulatory compliance as key priorities in the healthcare industry, underscoring the importance of security in database systems. In the manufacturing sector, a small electronics company implemented a database system integrated with IoT sensors to monitor equipment performance in real time. Before the system was introduced, the company experienced frequent equipment downtime, leading to costly delays in production. The new system enabled predictive maintenance by analyzing sensor data, significantly reducing unplanned downtime and improving operational efficiency. The case aligns with the systematic review’s insights into the role of databases in supporting real-time data processing and predictive analytics in manufacturing environments.

A financial institution offers another relevant case study. The company transitioned from a traditional on-premises database system to a cloud-based solution to handle real-time financial transactions and manage risk. The previous system struggled with high transaction volumes and increasing compliance demands, but the cloud solution provided the scalability and security required to meet these challenges. The new system also incorporated real-time fraud detection mechanisms, enhancing both operational efficiency and customer trust. This case directly relates to the systematic review’s findings, which highlight the critical need for high performance, scalability, and regulatory adherence in the finance industry. Table 18 summarizes these real-world case studies and illustrates the key metrics before and after database and data warehouse implementation.

Table 18. Summary of real-case studies related to database and data warehouse technologies.

Ref.	Case Study	Challenges	Solutions Implemented	Outcomes	Data Sensitivity	Performance Improvements	Scalability	Regulatory Compliance	Challenges Faced
[185]	Retail	Data fragmentation, slow data retrieval speeds	Cloud-based data warehouse	Centralized data, improved sales forecasting	Medium (customer purchase history)	Faster data access, improved decision-making	High (handled seasonal peaks)	GDPR (customer data privacy)	Integrating with existing e-commerce platforms

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

[186]	Healthcare	Compliance issues, data security challenges	HIPAA-compliant data warehouse	Enhanced patient care, secure data storage	High (patient records, medical data)	Quicker access to patient data, predictive analytics	High (accommodated growing patient data)	HIPAA, GDPR	Ensuring data security during migration
[187]	Manufacturing	Equipment downtime, inefficient process management	IoT-integrated database for predictive maintenance	Reduced equipment downtime, higher operational efficiency	Medium (production techniques, supply chain)	Real-time data processing, optimized maintenance	Moderate (integrated IoT data from multiple locations)	ISO compliance (industry-specific)	Data standardization from IoT devices
[188]	Finance	High transaction volume, fraud detection inefficiencies	Cloud-based database with real-time fraud detection	Scalable transactions, enhanced fraud prevention	High (financial transactions, personal data)	Real-time transaction processing, risk management	High (handled increased transaction volumes)	PCI DSS, SOX, GDPR	Ensuring real-time fraud detection at scale

These case studies provide clear examples of how SMEs have leveraged database and data warehouse technologies to overcome common challenges in their industries. The systematic review's emphasis on performance, scalability, and regulatory compliance is reflected in each case, demonstrating the practical relevance of the research. These cases not only validate the academic findings but also offer a roadmap for SMEs looking to implement similar technologies to achieve competitive advantage and operational efficiency. The comparison table helps visualize key performance improvements across different industries, while the case study timelines can illustrate the various phases of implementation, from initial challenges to final outcomes. Each case offers important lessons for SMEs, showing how the right technological solutions can drive both short-term efficiency gains and long-term sustainability.

b. Roadmap for SME Businesses and Policy Recommendations.

The proposed roadmap in Table 19 ties the technological evolution of SMEs to industry-specific needs and integrates policy recommendations that can help businesses scale, compete globally, and operate sustainably. Retail SMEs benefit from integrating advanced CRM and ERP systems to enhance customer engagement and optimize inventory management, with government policies focused on supporting digital marketing and e-commerce expansion. Healthcare SMEs require secure and compliant data management systems, with governments incentivizing the adoption of AI for diagnostics. Finance SMEs need support for adopting AI and blockchain to enhance fraud detection and secure transactions, while manufacturing SMEs can increase productivity through IoT and predictive maintenance.

In the logistics and supply chain sector, government subsidies for adopting advanced tracking and AI-based route optimization can enhance global supply chain management. Similarly, energy SMEs should receive incentives to adopt renewable energy technologies and IoT systems for real-time grid monitoring, ensuring sustainability and regulatory compliance.

Table 19. Proposed Roadmap for SME Businesses and Policy Recommendations.

Industry	Business Objective	Technology Adoption Focus	Workforce Development	Operational Strategy	Scalability & Digital Transformation	Policy Recommendations
Retail	Grow market presence, increase customer engagement, and optimize operations	Implement advanced CRM, e-commerce platforms, and data analytics. Focus on mobile integration for customers.	Digital marketing, customer engagement, CRM management training.	Implement real-time customer engagement and predictive sales analytics through CRM.	Adopt data-driven decision-making through analytics and BI tools to optimize marketing and inventory.	Support for Digital Commerce - Governments should provide subsidies for e-commerce platform adoption and offer tax credits for digital marketing investments.
	Scale operations locally and regionally, manage inventory efficiently	Deploy ERP for inventory, sales, and financials integration; integrate AI for personalized customer experience.	Train staff in ERP management, supply chain optimization, and personalized marketing strategies using AI tools.	Automate inventory tracking and streamline supply chain management with real-time visibility.	Implement cloud-based ERP and AI-driven sales forecasting to prepare for regional expansion.	Subsidized Technology Training - Government grants should be made available for training programs on advanced ERP and AI systems for retailers.
	Develop a global e-commerce platform,	Implement cross-border e-commerce, integrate global payment	Train staff on global trade regulations, international	Build a global presence by optimizing logistics, improving	Scale operations by integrating global ERP systems to	International Market Entry Support - Policymakers should reduce

doi:10.20944/preprints202410.0059.v1

	compete globally	gateways, and global supply chain management tools.	payment systems, and cross-border logistics management.	cross-border customer support, and localizing product offerings for diverse markets.	manage cross-border supply chains, payments, and customer engagement.	tariffs, streamline regulations for cross-border e-commerce, and provide export subsidies.
Healthcare	Ensure compliance and secure patient data	Deploy relational databases with HL7 compliance, integrate EHR systems, and ensure secure data storage.	Train staff on data privacy regulations (HIPAA, GDPR) and EHR management.	Implement secure EHR systems for real-time access to patient data, ensuring compliance with healthcare regulations.	Expand to cloud-based EHR systems and implement AI tools for predictive diagnostics and patient care analytics.	Grants for Healthcare Data Security - Governments should provide grants to help SMEs in healthcare adopt secure EHR systems and ensure compliance with regulations.
	Improve healthcare outcomes through data-driven decision-making	Integrate AI for predictive analytics in patient care, enhance data integration across systems (billing, patient management).	Train staff in AI tools for healthcare, advanced data analytics, and machine learning for diagnostics.	Optimize patient care through predictive analytics, and automate billing and patient record management.	Implement scalable AI tools to analyze large datasets and predict healthcare trends, allowing for proactive care.	Incentives for AI Adoption in Healthcare - Governments should incentivize AI adoption in healthcare by offering subsidies for AI-based diagnostic tools.
Finance & Banking	Secure financial transactions and ensure compliance	Deploy encrypted relational databases for real-time transactions and compliance (PCI DSS, AML).	Train staff on financial compliance (AML, PCI DSS), secure transaction processing, and fraud	Implement real-time fraud detection tools and ensure secure customer data management, compliant with	Adopt scalable cloud-based systems for global financial transactions and integrate blockchain for	Regulatory Support for Compliance - Governments should offer financial assistance for SMEs to adopt compliance

			detection systems.	financial regulations.	secure audit trails.	systems (AML, PCI DSS) and secure databases.
	Enhance customer experience and strengthen fraud detection capabilities	Implement AI for customer behavior analysis and fraud detection. Deploy real-time analytics for customer insights.	Train data scientists, fraud analysts, and machine learning specialists for AI-powered customer analysis.	Optimize customer interactions through AI-based personalization while enhancing fraud detection systems to reduce financial crime risks.	Scale using global financial ERP systems and integrate blockchain for tamper-proof financial records and transaction histories.	AI and Blockchain Incentives - Provide grants for SMEs adopting AI in fraud detection and blockchain technology for secure financial transactions.
Manufacturing	Automate production lines and improve supply chain management	Implement IoT-based databases and ERP systems to track production and inventory in real-time.	Train staff in IoT systems, automation tools, and ERP management for real-time operations monitoring.	Automate manufacturing processes through IoT and ERP integration, ensuring continuous production monitoring and reducing downtime.	Scale using AI-driven predictive maintenance and integrate machine learning to optimize production and reduce waste.	Subsidies for IoT Adoption - Governments should subsidize IoT and automation technologies for manufacturers to boost productivity and reduce downtime.
	Predictive maintenance and reduce operational downtime	Deploy AI for predictive maintenance, integrate IoT sensors for real-time equipment monitoring.	Provide training in AI tools for predictive maintenance, IoT data analysis, and real-time monitoring systems.	Automate predictive maintenance to reduce operational downtime and optimize equipment life cycles.	Expand IoT networks to cover all production lines and implement AI for real-time production adjustments.	Incentives for Predictive Maintenance - Offer tax breaks and financial support for SMEs adopting predictive maintenance technologies and AI.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

E-commerce	Enhance customer experience and drive sales growth through data analytics	Implement advanced CRM and ERP systems, integrate AI for product recommendations and dynamic pricing.	Train staff on personalized marketing, data-driven sales strategies, and CRM management.	Personalize the shopping experience through AI-powered product recommendations, and optimize pricing using dynamic algorithms.	Scale operations by implementing real-time analytics and personalized marketing strategies through AI-driven CRM tools.	Grants for E-commerce Expansion - Governments should provide grants for SMEs adopting advanced CRM, ERP, and AI tools to enhance customer experience and sales.
	Optimize supply chain and logistics management for fast delivery	Deploy cloud-based ERP systems for supply chain optimization and real-time inventory tracking.	Train staff in logistics management, inventory tracking using ERP, and real-time customer support systems.	Optimize supply chain through automated tracking, ensuring fast delivery and inventory optimization.	Scale supply chain operations globally, integrate with cross-border logistics, and optimize for international markets using AI-driven forecasting.	Export and Logistics Support - Policymakers should provide subsidies for SMEs adopting supply chain optimization technologies and reduce logistical barriers for cross-border e-commerce.
Logistics & Supply Chain	Optimize shipping routes and improve real-time shipment tracking	Implement graph databases for complex supply chain management and IoT sensors for real-time tracking.	Train staff in real-time logistics tracking, route optimization, and IoT integration for shipment management.	Optimize shipping routes and warehouse management through real-time tracking and IoT sensor integration.	Scale using AI for predictive demand and route optimization, ensuring efficient resource utilization.	Support for Logistics Technology - Governments should provide financial support for adopting advanced logistics technologies, including AI for route optimization

 doi:10.20944/preprints202410.0059.v1

						and IoT for real-time tracking.
	Enhance global supply chain visibility and optimize operations	Deploy distributed ERP systems and AI for global supply chain visibility, real-time demand forecasting.	Provide training in global supply chain management, AI-driven demand forecasting, and risk mitigation strategies.	Enhance supply chain transparency and optimize operations by integrating data from global logistics partners.	Scale using global ERP systems to integrate suppliers, distributors, and third-party logistics providers into a single platform.	Trade Facilitation & Global Supply Chains - Offer financial incentives for SMEs to participate in global supply chains and streamline regulations for international logistics operations.
Education (Higher Ed)	Digital transformation for remote learning and research management	Implement cloud-based LMS systems, integrate data analytics for tracking student performance and learning outcomes.	Train educators and administrative staff in LMS management, e-learning tools, and data analytics for performance tracking.	Transition to digital learning platforms for enhanced student engagement and optimized learning outcomes.	Scale e-learning platforms using cloud-based infrastructure, and integrate AI for personalized learning experiences.	Grants for E-learning Platforms - Governments should offer financial support for adopting digital learning platforms and provide tax incentives for developing e-learning technologies.
	Enhance research management and collaboration through data integration	Deploy cloud-based research data management systems, integrate collaborative	Train researchers in data management systems, cloud-based collaboratio	Optimize research collaboration by enabling real-time data sharing and collaborative	Scale research capabilities by integrating global research data platforms and ensuring	Support for Research Collaboration - Policymakers should provide grants for adopting cloud-

 doi:10.20944/preprints202410.0059.v1

		tools for global research networks.	n tools, and digital library systems.	research across institutions.	secure data sharing across institutions.	based research collaboration tools and digital data-sharing platforms.
Hospitality & Tourism	Enhance customer experience through personalized service	Deploy CRM systems for loyalty program management, integrate AI for personalized guest recommendations and dynamic pricing.	Train staff in CRM tools, guest data management, and personalized marketing strategies using AI.	Optimize guest experiences through real-time data insights, improving loyalty programs and personalized recommendations.	Scale guest services using global CRM systems and AI for real-time pricing optimization, ensuring personalized experiences.	Subsidies for Tourism Tech - Provide subsidies for SMEs adopting CRM and AI-based personalized marketing technologies to enhance customer satisfaction.
	Optimize booking systems and streamline property management	Implement cloud-based booking engines, integrate real-time property management systems (PMS).	Train staff in digital booking management, property management systems, and cross-channel marketing strategies.	Automate booking processes, ensure seamless property management, and optimize customer engagement through real-time insights.	Scale property management operations by integrating cloud-based PMS systems and optimizing cross-channel booking platforms globally.	Incentives for Cloud-based Booking Systems - Policymakers should incentivize SMEs to adopt cloud-based property management and booking systems through tax incentives and financial support.
Energy & Utilities	Monitor energy usage and optimize operations for sustainability	Deploy time-series databases for smart meter data, integrate IoT for real-time grid performance monitoring.	Train staff in real-time energy usage monitoring, IoT data analysis, and energy	Optimize energy usage through real-time grid monitoring, ensuring predictive maintenance for	Scale using AI-driven predictive analytics for energy management and expand IoT systems to cover all	Green Energy Incentives - Governments should offer subsidies for SMEs adopting renewable energy technologies, IoT

		optimization strategies.	energy infrastructure.	energy distribution points.	monitoring systems, and energy optimization platforms.
Ensure compliance with environmental and regulatory standards	Implement carbon footprint tracking systems, integrate renewable energy systems and smart grid technologies.	Train staff in carbon tracking systems, environmental management tools, and compliance reporting.	Automate compliance with environmental standards, optimize energy usage, and reduce carbon footprint through renewable energy adoption.	Scale energy distribution systems using IoT and AI for real-time grid management, ensuring compliance with evolving regulatory frameworks.	Grants for Sustainability and Compliance - Governments should offer grants for adopting carbon footprint tracking and renewable energy systems, promoting sustainable practices in SMEs.

To successfully implement database and data warehouse technologies, SMEs need a structured roadmap that outlines clear steps for short-, medium-, and long-term goals. This roadmap will guide businesses through key phases such as assessment, implementation, scaling, and optimization while also considering the challenges and resources needed to adopt these technologies as shown in Figure 30. This leads to the development of policies by governments and industry associations to assist SMEs in adopting data technologies. Financial support may be available through subsidies, tax incentives, or grants that cover part of the initial costs associated with acquiring technology, training, and data migration. Training and education through online courses, certification, and workshops will help SMEs understand the benefits accrued and the best practices concerning database technologies. Public-private partnerships could further enable collaboration between SMEs and technology vendors to develop appropriate solutions. Clear guidance on the issue of data privacy and security compliance, especially in high regulated verticals such as healthcare and finance, would better support SMEs in handling the respective compliance obligations. This will also put SMEs in a position where they can make informed choices regarding their technological investments

Figure 30. Roadmap for Database and Data Warehouse Technologies Adoption in SMEs.

The proposed roadmap to the implementation of SMEs covers four steps: adopting database and data warehouse technologies through initial assessment, implementation, scaling, and optimization. Initial assessments allow for identifying needs and challenges, while at this stage, transitions are made from old legacy systems to modern databases or data warehouses. This phase scales database capabilities for increased demand without major overhauls. Therefore, the optimization phase chiefly focuses on long-term improvement by enabling the SMEs to leverage advanced analytics, driving innovation, and thereby maintaining competitiveness. Policies recommend reducing financial barriers while making available necessary resources.

4. Discussion

a. General Interpretation of the Results in the Context of Other Evidence

The results of this systematic review reveal that database and data warehouse technologies have a significant impact on organizational performance, with a strong focus on enhancing customer satisfaction, operational efficiency, and competitive advantage. The emphasis on relational databases and data warehousing systems aligns with broader trends in the literature, which highlight the foundational role of these technologies in modern data management. Studies focused on NoSQL databases and cloud computing also underscore the growing interest in scalable and flexible systems, a key concern in the current technological landscape. The prominence of business intelligence systems and their integration with knowledge management aligns with other studies that emphasize how data-driven decision-making boosts organizational performance. This review's findings also support the growing body of evidence showing that business intelligence and data analytics systems are instrumental in achieving long-term business sustainability and gaining a competitive edge. While the literature largely reflects the use of these technologies in developed regions, this review highlights a critical need for more research in developing economies to fully understand the global impact of these technologies.

b. Limitations of the Evidence Included

A notable limitation in the included studies is the geographical concentration of research, with most studies originating from developed regions. This creates potential challenges in generalizing the findings to less economically developed regions where the infrastructure for database and data

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 October 2024 doi:10.20944/preprints202410.0059.v1

warehouse technologies might be less advanced. Additionally, there is an over-representation of studies focusing on customer satisfaction, while employee satisfaction and other internal organizational metrics are less frequently analyzed. This suggests a gap in the literature regarding how these technologies influence internal operations and employee experiences. Another concern is underrepresentation of specific technologies, such as High-Performance Computing (HPC) and CRM systems, despite their importance in certain industries. The limited attention to these areas indicates either a lack of research focus or challenges in implementation that are not well-documented.

c. Limitations of the Review Process Used

The review process utilized stringent inclusion and exclusion criteria, which helped maintain focus but may have excluded some relevant studies. For instance, the reliance on specific keywords in the search strategy might have unintentionally missed research that employs different terminology or conceptual frameworks. Furthermore, the exclusion of non-English studies could have led to an incomplete representation of global research trends. The lack of automation tools and manual efforts to contact study investigators for missing information introduced some limitations. Although this approach enhanced the accuracy of the review, it may have led to potential biases in study selection or interpretation. Also, the use of the Newcastle-Ottawa Scale (NOS) and GRADE 3.0 frameworks provided a robust quality assessment but may not capture all nuances, especially in newer or interdisciplinary studies.

d. Implications of the Results for Practice, Policy, and Future Research

The results indicate that organizations should prioritize investments in data warehouse systems and relational databases, as these technologies have been shown to significantly enhance operational efficiency and customer satisfaction. Organizations may also benefit from incorporating NoSQL and cloud database systems to enhance scalability and flexibility in managing large datasets. The results suggest that businesses aiming to improve customer satisfaction, a key driver of success, should focus on data-driven decision-making through business intelligence systems.

From a policy perspective, there is a clear need for governments and regulators, especially in developing countries, to invest in technological infrastructure that supports the adoption of database and data warehouse systems. Encouraging the use of these technologies in small and medium-sized enterprises (SMEs) through incentives and subsidies could lead to greater economic growth and business performance. Furthermore, standardization of data management practices could help reduce variability and enhance the implementation of these systems across industries.

Future research should focus on addressing the geographical bias observed in the current evidence by conducting more studies in developing regions to understand the challenges and benefits of adopting database technologies in these contexts. Moreover, there is a need to explore employee satisfaction and internal organizational outcomes, as they are often overlooked in favor of customer-centric metrics. Expanding the scope of research to include High-Performance Computing and CRM systems could also uncover new insights into niche technological areas. The intersection between business sustainability and competitive advantage deserves further exploration to understand how organizations can leverage database technologies for both short-term and long-term benefits. This would be particularly valuable for businesses in industries undergoing rapid technological transformation.

5. Conclusion

This systematic review provides a detailed assessment of the impact of database and data warehouse technologies on organizational performance, based on a comprehensive analysis of 150 studies conducted between 2014 and 2024. The review highlights a predominant focus on customer satisfaction and operational efficiency, with substantial interest in data warehouse systems and relational databases. However, there is a notable imbalance in research emphasis, with less attention given to employee satisfaction and foundational topics like CRM systems and data mining. The findings reveal a concentration of research in economically advanced countries, which may limit the generalizability of results to less developed regions. The dual focus on competitive advantage and business sustainability underscores the strategic importance of these technologies in achieving both immediate and long-term organizational goals. The predominance of quantitative methods and the emphasis on scalability and data accuracy in IT performance metrics reflect a strong inclination towards numerical evaluations and technical performance aspects. This methodological diversity, encompassing quantitative, qualitative, and mixed methods approaches, ensures a robust understanding of the multifaceted impacts of these technologies. Future research should aim to address the gaps identified, particularly in exploring employee-related outcomes and foundational technologies. Additionally, there is a need to broaden the research scope to include a more diverse range of geographic and economic contexts to provide a more comprehensive understanding of the global impact of database and data warehouse technologies. This review serves as a foundational study for advancing knowledge in this field and guiding future research directions.

Author Contributions: N.G.M., N.M.F., and J.K.K., carried out the data collection, and investigations, wrote, and prepared the article under supervision of B.AT. B.A.T. & L.M were responsible for conceptualization, reviewing, and editing the article. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank all the researchers for their contribution in the database.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. T. Bradley, “Small Business Growth in The Digital Age,” Forbes. https://www.forbes.com/sites/tonybradley/2024/05/08/small-business-growth-in-the-digital-age/
2. L. Olmstead, “9 Critical Digital Transformation Challenges to Overcome (2022) - Whatfix,” The Whatfix Blog | Drive Digital Adoption, Dec. 02, 2022. https://whatfix.com/blog/digital-transformation-challenges/
3. P. C. Verhoef et al., “Digital transformation: a Multidisciplinary Reflection and Research Agenda,” Journal of Business Research, vol. 122, no. 122, pp. 889–901, Jan. 2021, doi: https://doi.org/10.1016/j.jbusres.2019.09.022.
4. X. Teng, Z. Wu, and F. Yang, “Impact of the Digital Transformation of Small- and Medium-Sized Listed Companies on Performance: Based on a Cost-Benefit Analysis Framework,” Journal of Mathematics, vol. 2022, pp. 1–14, Sep. 2022, doi: https://doi.org/10.1155/2022/1504499.
5. Vollcom Digital GmbH, “10 Key Steps to Creating an Effective Digital Transformation Roadmap,” Medium, Apr. 29, 2024. https://vollcomdigital.medium.com/10-key-steps-to-creating-an-effective-digital-transformation-roadmap-f67563332ddf (accessed Sep. 04, 2024).
6. C. Stedman, “What Is Data Management and Why Is It Important?,” SearchDataManagement, 2019. https://www.techtarget.com/searchdatamanagement/definition/data-management
7. Altexsoft, “Database Management Systems (DBMS) Comparison: MySQL, Postgr,” AltexSoft, May 16, 2023. https://www.altexsoft.com/blog/comparing-database-management-systems-mysql-postgresql-mssql-server-mongodb-elasticsearch-and-others/
8. “Data Warehouse vs. Data Database | Key Differences,” www.xenonstack.com. https://www.xenonstack.com/blog/data-warehouse-vs-data-database
9. P.Carnell, “Benefits and Challenges of Data Warehouse Automation,” Lonti.com, Nov. 08, 2023. https://www.lonti.com/blog/benefits-and-challenges-of-data-warehouse-automation (accessed Sep. 04, 2024).
10. J. O. Martins, R. Sahandi, and F. Tian, “Critical analysis of vendor lock-in and its impact on cloud computing migration: a business perspective,” Journal of Cloud Computing, vol. 5, no. 1, Apr. 2016, doi: https://doi.org/10.1186/s13677-016-0054-z.
11. S. Baroth, “Types of Databases,” GeeksforGeeks, May 08, 2020. https://www.geeksforgeeks.org/types-of-databases/
12. “What Are The Various Types of Databases? | Simplilearn,” Simplilearn.com, Jan. 12, 2023. https://www.simplilearn.com/tutorials/dbms-tutorial/what-are-various-types-of-databases
13. C. Cuello, “What Are the Different Types of Databases in 2023?,” Rivery, Jun. 19, 2023. https://rivery.io/data-learning-center/database-types-guide/

14. "What Are the Different Types of Databases?," Indeed Career Guide. https://www.indeed.com/career-advice/career-development/types-of-databases
15. A. P. Singh, "15 Types of Databases and When to Use Them," blog.algomaster.io. https://blog.algomaster.io/p/15-types-of-databases
16. C. Team, "Data Warehouse," Corporate Finance Institute, Nov. 21, 2023. https://corporatefinanceinstitute.com/resources/data-science/data-warehouse/#:~:text=Once%20in%20the%20data%20warehouse (accessed Sep. 04, 2024).
17. BluEnt, "What is Data Warehousing? Types, Benefits, & Use Cases," www.bluent.com, May 29, 2023. https://www.bluent.com/blog/what-is-data-warehousing/
18. "What is a Data Warehouse?," Oracle.com, 2020. https://www.oracle.com/za/database/what-is-a-data-warehouse/
19. "What is Data Warehousing? How it Works, Types, and General Stages," www.datachannel.co. https://www.datachannel.co/blogs/introduction-to-data-warehousing
20. Tubis, A.A. and Rohman, J. (2023). Intelligent Warehouse in Industry 4.0—Systematic Literature Review. Sensors, [online] 23(8), p.4105. doi:https://doi.org/10.3390/s23084105.
21. Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G. and Gnanzou, D. (2019). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. International Journal of Production Economics, [online] 165(165), pp.234–246. doi:https://doi.org/10.1016/j.ijpe.2014.12.031.
22. Yusuf Ayokunle, Shukla, N., Rahman, T., Chakraborty, S. and Kumari, S. (2024). The Performance Evaluation of Digital Technologies in Warehouse Management: A Systematic Literature Review. [online] doi:https://doi.org/10.2139/ssrn.4862260.
23. Llave, M.R. (2017). Business Intelligence and Analytics in Small and Medium-sized Enterprises: A Systematic Literature Review. Procedia Computer Science, [online] 121, pp.194–205. doi:https://doi.org/10.1016/j.procs.2017.11.027.
24. Lagorio, A., Zenezini, G., Mangano, G. and Pinto, R. (2020). A systematic literature review of innovative technologies adopted in logistics management. International Journal of Logistics Research and Applications, 25(7), pp.1–24. doi:https://doi.org/10.1080/13675567.2020.1850661.
25. Ahmad, S., Miskon, S., Alkanhal, T.A. and Tlili, I. (2020). Modeling of Business Intelligence Systems Using the Potential Determinants and Theories with the Lens of Individual, Technological, Organizational, and Environmental Contexts-A Systematic Literature Review. Applied Sciences, [online] 10(9), p.3208. doi:https://doi.org/10.3390/app10093208.
26. Kamble, S.S. and Gunasekaran, A. (2019). Big data-driven supply chain performance measurement system: a review and framework for implementation. International Journal of Production Research, 58(1), pp.1–22. doi:https://doi.org/10.1080/00207543.2019.1630770.
27. Plotnikova, V., Dumas, M. and Milani, F. (2020). Adaptations of data mining methodologies: a systematic literature review. PeerJ Computer Science, 6, p.e267. doi:https://doi.org/10.7717/peerj-cs.267.
28. Tallon, P.P., Queiroz, M., Coltman, T. and Sharma, R. (2019). Information technology and the search for organizational agility: A systematic review with future research possibilities. The Journal of Strategic Information Systems, [online] 28(2), pp.218–237. doi:https://doi.org/10.1016/j.jsis.2018.12.002.
29. Govindan, K., Kannan, D., Jørgensen, T.B. and Nielsen, T.S. (2022). Supply Chain 4.0 performance measurement: A systematic literature review, framework development, and empirical evidence. Transportation Research Part E: Logistics and Transportation Review, 164, p.102725. doi:https://doi.org/10.1016/j.tre.2022.102725.
30. Robershaw, K. and Wolf, B. (2023). Research Analytics: A Systematic Literature Review. SSRN Electronic Journal. doi:https://doi.org/10.2139/ssrn.4363262.
31. Gunther Misahuaman, Alfredo Sánchez Daza and Zavaleta, E. (2021). Web-based systems for inventory control in organizations: A Systematic Review. [online] doi:https://doi.org/10.1109/snpd51163.2021.9704993.
32. Antunes, A.L., Cardoso, E. and Barateiro, J. (2022). Incorporation of Ontologies in Data Warehouse/Business Intelligence Systems - A Systematic Literature Review. International Journal of Information Management Data Insights, 2(2), p.100131. doi:https://doi.org/10.1016/j.jjimei.2022.100131.
33. McManus, D.J. and Snyder, C.A. (2003). Synergy between data warehousing and knowledge management: three industries reviewed. International Journal of Information Technology and Management, 2(1/2), p.85. doi:https://doi.org/10.1504/ijitm.2003.002450.
34. Bogojevic, P. (2020). Project Management in Data Warehouse Implementations: A Literature Review. IEEE Access, pp.1–1. doi:https://doi.org/10.1109/access.2020.3045072.
35. Tsiu, S.; Ngobeni, M.; Mathabela, L.; Thango, B. Applications and Competitive Advantages of Data Mining and Business Intelligence in SMEs Performance: A Systematic Review. *Preprints* 2024, 2024090940. https://doi.org/10.20944/preprints202409.0940.v1

36. Ferretti, L., Colajanni, M. & Marchetti, M., 2014. Distributed, Concurrent, and Independent Access to Encrypted Cloud Databases. *IEEE Transactions on Parallel and Distributed Systems*, 25(2), pp.437-446. http://doi.ieeecomputersociety.org/10.1109/TPDS.2013.154 [Accessed 8 Aug. 2024].
37. Ferretti, L., Pierazzi, F., Colajanni, M. & Marchetti, M., 2014. Performance and Cost Evaluation of an Adaptive Encryption Architecture for Cloud Databases. IEEE Transactions on Cloud Computing, 2(2), pp.171-184.
38. Pippal, S.K. & Kushwaha, D.S., 2013. A simple, adaptable and efficient heterogeneous multi-tenant database architecture for ad hoc cloud. Journal of Cloud Computing: Advances, Systems and Applications, 2(1), p.5. https://doi.org/10.1186/2192-113X-2-5 [Accessed 8 Aug. 2024].
39. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., 2017. 'Improved Performance of Data Warehouse', International Conference on Inventive Communication and Computational Technologies, 16(7), India: SME.
40. Gold A., Malhotra A. & Segars A. 2023. 'The Role of Knowledge Management Processes and Business Intelligence Systems in Enhancing Organizational Performance in the Banking Sector', Journal of Banking and Financial Studies, 12(9), Jordan: SME.
41. Seah, B.K. and Selan, N.E., 2014. Design and implementation of data warehouse with data model using survey-based services data. In: *Proceedings of the Fourth International Conference on Innovative Computing Technology (INTECH 2014)*. Luton, UK, 13-15 August 2014. IEEE. DOI: 10.1109/INTECH.2014.6927748. https://ieeexplore.ieee.org/document/6927748.
42. Mkhize, A.; Mokhothu, K.; Tshikhotho, M.; Thango, B. Evaluating the Impact of Cloud Computing on SMEs Performance: A Systematic Review. Preprints 2024, 2024090882. https://doi.org/10.20944/preprints202409.0882.v1.
43. Garani, G., Chernov, A., Savvas, I. and Butakova, M. 2019. 'A Data Warehouse Approach for Business Intelligence', IEEE Conference, 16(17), Russia: SME.
44. Ojeda, A. 2015. Critical Factors for Successful Implementation of Data Warehouses, Issues in Information Systems, 12(11), Puerto Rico: SME.
45. Benker, T. 2016. A Generic Process Data Warehouse Schema for BPMN Workflows', BIS 2016, 9(19), Germany: Startups.
46. Asrani D, Jain R. 2016. Designing a Framework to Standardize Data Warehouse Development Process for Effective Data Warehousing Practices', International Journal of Database Management Systems, 5(15), India: SME.
47. Janiszewska D, Korczak J. 2022. Diversification of Equipment in the IT Infrastructure of Enterprises in Central Pomerania in Poland', Energies, 0(15), Poland: SME.
48. Masior, J. 2019. 'Design Requirements on the Integration of Technology Databases', Measuring Business Excellence, 19(23), England: Small Businesses.
49. Kgakatsi, M.; Galeboe, O.; Molelekwa, K.; Thango, B. The Impact of Big Data on SME Performance: A Systematic Review. Preprints 2024, 2024090985. https://doi.org/10.20944/preprints202409.0985.v1.
50. Huynh, B., Tung, N.T., Nguyen, T.D.D., Bui, Q.-T., Nguyen, L.T.T., Yun, U., and Vo, B., 2024. An efficient strategy for mining high-efficiency itemsets in quantitative databases. *Knowledge-Based Systems*, 299, p.112035. [Accessed 8 Aug. 2024].
51. Ferretti, L., Pierazzi, F., Colajanni, M. & Marchetti, M. 2014. Performance and Cost Evaluation of an Adaptive Encryption Architecture for Cloud Databases', IEEE Transactions on Cloud Computing, 7(23), Italy: Small Businesses.
52. Molete, O. B.; Mokhele, S. E.; Ntombela, S. D.; Thango, B. A. The Impact of IT Strategic Planning Process on SME Performance: A Systematic Review. Preprints 2024, 2024091024. https://doi.org/10.20944/preprints202409.1024.v1.
53. Antunes, A.L., Cardoso, E. & Barateiro, J., 2022. Incorporation of ontologies in data warehouse/business intelligence systems - a systematic literature review. *International Journal of Information Management Data Insights*, 2(2), p.100131.
54. Bharadiya, J. 2023. 'A comparative study of business intelligence and artificial intelligence with big data analytics', *American Journal of Artificial Intelligence*, vol. 9, no. 164.
55. Siddiqui, G.F, Aftab, U. 2018. 'Big data augmentation with data warehouse: A survey', *IEEE International Conference on Big Data*, vol. 19, no. 46.
56. Abu Salma, A.J., Prasolov, V., Glazkova, I., Rogulin, R. (2021) 'The impact of business processes on the efficiency of small and medium-sized enterprises', *Montenegrin Journal of Economics*, vol. 17, no. 21.
57. Zhang, C. Philip Chen, C.L., 2014. 'Data-intensive applications, challenges, techniques and technologies: A survey on big data', *Information Sciences*, vol. 15, no. 3692.
58. Chen, P.-T., Lin, C.-L., & Wu, W.-N., 2019. 'Big data management in healthcare: Adoption challenges and implications', *International Journal of Information Management*, vol. 16, no. 192.

59. Seah, B.K. and Selan, N.E., 2014. 'Design and implementation of data warehouse with data model using survey-based services data', *IEEE Conference*, vol. 7, no. 2.
60. Stemberger, M.I., Jaklic, j., Bach, P.M., Vuksic, V.B., Vugec, S.D, 2019. 'Business intelligence and organizational performance: The role of alignment with business process management', *The Role of Alignment with Business Process Management*, vol. 26, no. 21.
61. Tustin, D., Venter, P, 2014. 'The availability and use of competitive and business intelligence in South African business organizations', *Southern African Business Review*, vol. 13, no. 13.
62. Raguseo, E., 2018. 'Big data technologies: An empirical investigation on their adoption, benefits and risks for companies', *International Journal of Information Management*, vol. 9, no. 508.
63. Rudniy, A., 2022. 'Data warehouse design for big data in academia', *Computers, Materials & Continua*, vol. 10, no. 20. .
64. Krishnamurthi, M., 2022. 'The role of data warehousing in the infrastructure of e-commerce', *International Journal of Advances in Engineering and Management*, vol. 4, no. 6.
65. Mathew, K.S., Dahiya D., 2015. 'IT assets, IT infrastructure performance and IT capability: A framework for e-government', *Transforming Government-people Process and Policy*, vol. 10, no. 19.
66. Gayakwad, M., 2017. 'Improved performance of data warehouse', *International Conference on Inventive Communication and Computational Technologies*, vol. 8, no. 16.
67. Putra, E., Hidayanto, A., Pikarti, G., Isal, Y., 2016. 'Analysis of IT infrastructure flexibility impacts in IT-business strategic alignment', *Journal of Industrial Engineering and Management*, vol. 9, no. 12.
68. Kumar, A., 2024. 'Data-driven decision making: Advanced database systems for business intelligence', *Journal of Business Intelligence and Analytics*, vol. 1, no. 150.
69. Mothapo, M.; Thango (Y2-rated Researcher), B.; Matshaka, L. Tracking and Measuring Social Media Activity: Key Metrics for SME Strategic Success – A Systematic Review. Preprints 2024, 2024091757. https://doi.org/10.20944/preprints202409.1757.v1.
70. Afshari, S., Ravasan, A.Z., Ashrafi, A., Rouhani, S., 2016. 'The impact model of business intelligence on decision support and organizational benefits', *Journal of Enterprise Information Management*, vol. 12, no. 12.
71. Gupta, P., Vaish, P., 2020. 'Business intelligence: Escalation of data warehousing and data mining for effective decision making', *International Journal of Advanced Science and Technology*, vol. 8, no. 4.
72. Patil, Y., Kumari, S., Jeble, S., 2018. 'Role of big data in decision making', *Operations and Supply Chain Management*, vol. 9, no. 5.
73. Al-Habsi, N.A.S., Jalagat, R., 2017. 'Evaluating the impacts of IT usage on organizational performance', *European Academic Research*, vol. 5, no. 8.
74. Yass, A., 2024. 'The impact of cloud-based information systems on organizational performance', *IOSR Journals*, vol. 7, no. 23.
75. Ahmad, K., JianMing, Z., Rafi, M., 2019. 'Evaluating the impact of digital library database resources on the productivity of academic research', *Information Discovery and Delivery*, vol. 14, no. 3.
76. Abualous, S., Abusweilem, M.A., 2019. 'The impact of knowledge management process and business intelligence on organizational performance', *Management Science Letters*, vol. 15, no. 10.
77. Rafi, M., 2020. 'A resource-based perspective on information technology and firm performance: A meta-analysis', *Industrial Management & Data Systems*, vol. 13, no. 19.
78. Saleh, S.S., Nakshabandi, A.O., Ismael, G.Y., Zeebaree, M. 2020. 'Impact of innovation technology in enhancing organizational management', *IEEE Conference*, vol. 20, no. 23.
79. Ramirez-Correa, P., Alfaro-Perez, J.L., Grandon, E.E., Foster, P., Guzman, S.A., 2020. 'Influence of information systems in organizational performance', *IEEE Conference*, vol. 21, no. 55.
80. Datta, P., Nwankpa, J.K., 2017. 'Balancing exploration and exploitation of IT resources: The influence of digital business intensity on perceived organizational performance', *European Journal of Information Systems*, vol. 14, no. 333.
81. Bach, P.M., Poor, J., Barisic, A.F., 2019. 'The intensity of human resources information systems usage and organizational performance', *INDECS*, vol. 11, no. 246.
82. Fabregat, A., Korninger, F., Viteri, G., Sidiropoulos, K., Marin-Garcia, P., Ping, P., Wu, G., Stein, L. D'Eustachio, P. and Hermjakob, H., 2018. 'Reactome graph database: Efficient access to complex pathway data', *PLOS Computational Biology*, vol. 7, no. 46.
83. Ngcobo, K.; Bhengu, S.; Mudau, A.; Thango, B.; Matshaka, L. Enterprise Data Management: Types, Sources, and Real-Time Applications to Enhance Business Performance - A Systematic Review. Preprints 2024, 2024091913. https://doi.org/10.20944/preprints202409.1913.v1.
84. Vitolo, C., Elkhatib, Y., Reusser, D., Macleod, C.J.A. and Buytaert, W., 2014. 'Web technologies for environmental big data', *Environmental Modelling & Software*, vol. 14, no. 77.
85. Wang, C., Blut, M., 2019. 'Technology readiness: A meta-analysis of quantitativeizations of the construct and its impact on technology usage', *Journal of the Academy of Marketing Science*, vol. 10, no. 41.
86. Ramu, B.V., 2023. 'Performance impact of microservices architecture', *The Review of Contemporary Scientific and Academic Studies*, vol. 21, no. 3330.

87. Hudson, L.N., Newbold, T., Contu, S., and Hill, S.L.L., 2014. 'The PREDICTS database: A global database of how local terrestrial biodiversity responds to human impacts', *Ecological Solutions and Evidence*, vol. 23, no. 168.
88. Jiang, C., Bao, Y., Chen, X., Liu, C., Zhu, G., Jiang, F., & Chen, M., 2023. 'Optimizing data warehouse technology', *IEEE Transactions on Power Electronics*, vol. 5, no. 120.
89. Appelbaum, D., Kogan, A., Vasarhelyi, M. and Yan, Z., 2017. 'Impact of business analytics and enterprise systems on data warehouse technology', *Accounting, Organizations and Society*, vol. 6, no. 43.
90. Qaderi, A., Ghoul, L., Jaradat, L., 2016. 'The impact model of database on decision support and organizational benefits', *Journal of Enterprise Information Management*, vol. 5, no. 8.
91. Kwon, O., Lee, N., and Shin, B., 2014. 'Data quality management, data usage experience and acquisition intention of big data analytics', *International Journal of Information Management*, vol. 10, no. 22.
92. Kuziemsky, C., Foshay, N., 2016. 'Towards an implementation framework for data warehouse in healthcare', *International Journal of Information Management*, vol. 12, no. 46.
93. Armbrust, M., Ghodsi, A., Xin, R. and Zaharia, M., 2021. 'Lakehouse: A new generation of open platforms that unify data warehousing and advanced analytics', *Conference Paper*, vol. 15, no. 50.
94. Atieh, A.M., Kaylani, H., Al-abdallat, Y., Qaderi, A., Ghoul, L., Jaradat, L. and Hdairis, I., 2015. Performance Improvement of Inventory Management System Processes by an Automated Warehouse Management System (2015) *Conference Paper*. Vol. 10. SME performance. Jordan.
95. Ding, J., Apon, A., Gudivada, V.N., 2017. Data Quality Considerations for Big Data and Machine Learning (2017) *International Journal on Advances in Software*. Vol. 2. Database technologies. SME. USA.
96. Leto, L., Gobbo, R.D., Castellano, N., 2023. Using Big Data to Enhance Data Envelopment Analysis of Retail Store Productivity (2023) *International Journal of Productivity and Performance Management*. Vol. 14. Data warehouse. Startups. USA.
97. Nweke, H.F., Ajah, I.A., 2019. Big Data and Business Analytics: Trends, Platforms, Success Factors and Applications (2019) *Big Data Cogn. Comput.* Vol. 32. Business technology. Small businesses. Nigeria.
98. Peters, M.D., Wieder, B., Sutton, S.G. and Wakefield, J., 2016. Business Intelligence Systems Use in Performance Measurement Capabilities: Implications for Enhanced Competitive Advantage (2016) *Journal of Accounting & Organizational Change*. Vol. 23. SME performance. SME. Australia.
99. Talon-Ballestero, P., González-Serrano, L., Soguero-Ruiz, C., Muñoz-Romero, S. and Rojo-Álvarez, J.L., 2018. Using Big Data from Customer Relationship Management Information Systems to Determine the Client Profile in the Hotel Sector (2018) *Tourism Management*. Vol. 11. Database technologies. SME. Sweden.
100. Accorsi, R., Manzini, R. & Maranesi, F., 2016. A Decision-Support System for the Design and Management of Warehousing Systems (2016) *Computers & Industrial Engineering*. Vol. 10. Data warehouse. Small businesses. Japan.
101. Fosso Wamba, S., Gunasekaran, A., Akter, S., Ren, S.J.-F., Dubey, R. and Childe, S.J., 2016. Big Data Analytics and Firm Performance: Effects of Dynamic Capabilities (2016) *Journal of Business Research*. Vol. 5. Business technology. SME. China.
102. Akter, S., Fosso Wamba, S., Gunasekaran, A., Dubey, R. & Childe, S.J., 2016. How to Improve Firm Performance Using Big Data Analytics Capability and Business Strategy Alignment (2016) *International Journal of Production Economics*. Vol. 1. SME performance. Small businesses. USA.
103. Jain, S., Saggi, M.K., 2018. A Survey Towards an Integration of Big Data Analytics to Big Insights for Value-Creation (2018) *Information Processing & Management*. Vol. 5. Database technologies. SME. Switzerland.
104. Santos, M.Y., Oliveira e Sá, J., Andrade, C., Vale Lima, F., Costa, E., Costa, C., Martinho, B. & Galvão, J., 2017. A Big Data System Supporting Bosch Braga Industry 4.0 Strategy (2017) *Information & Management*. Vol. 10. Data warehouse. SME. Portugal.
105. Xiong, Y., Duan, L., 2015. Big Data Analytics (2015) *Volume 16, Issue II*. Vol. 12. Business technology. SME. Ethiopia.
106. Mohammad, A.B., Al-Okaily, M., Al-Majali, M. and Masa'deh, R., 2022. Business Intelligence and Analytics (BIA) Usage in the Banking Industry Sector: An Application of the TOE Framework (2022) *J. Open Innov. Technol. Mark. Complex.* Vol. 15. SME performance. SME. Jordan.
107. Mohlala, T. T.; Mehlwana, L. L.; Nekhavhambe, U. P.; Thango, B.; Matshaka, L. Strategic Innovation in HRIS and AI for Enhancing Workforce Productivity in SMEs: A Systematic Review. Preprints 2024, 2024091996. https://doi.org/10.20944/preprints202409.1996.v1.
108. Arunachalam, D., Kumar, N. and Kawalek, J.P., 2017. Understanding Big Data Analytics Capabilities in Supply Chain Management: Unravelling the Issues, Challenges and Implications for Practice (2017) *Technological Forecasting and Social Change*. Vol. 19. Data warehouse. Small businesses. Chile.
109. Navimipour, N.J., Soltani, Z., 2016. Customer Relationship Management Mechanisms: A Systematic Review of the State of the Art Literature and Recommendations for Future Research (2016) *Computers in Human Behavior*. Vol. 4. Business technology. SME. Brazil.
110. Sheikh, N., Shahid, U., 2024. The Big Data Revolution: Leveraging Vast Information for Competitive Advantage (2024) *Revista Española de Documentación Científica*. Vol. 3. Data warehouse. SME. Haiti.

111. Akter, S., 2016. Big Data Analytics and Firm Performance: Effects of Dynamic Capabilities (2016) *Journal of Business Research*. Vol. 9. Data warehouse. SME. China.
112. Min, S.K., Suh, E.H. and Kim, S.Y., 2023. An Integrated Approach for IT Strategic Planning and Business Performance (2023) *International Journal of Information Management*. Vol. 15. Data warehouse. Startups. Europe.
113. AI-Mamun, M.A., Naeem, Z., 2014. Data Quality Management, Data Usage Experience and Acquisition Intention of Big Data Analytics (2014) *International Journal of Information Management*. Vol. 20. Data warehouse. SME. South Korea.
114. Prescott, M.E., 2014. Big Data and Competitive Advantage at Nielsen (2014) *Management Decision*. Vol. 14. Data warehouse. Small businesses. USA. Emerald Insight.
115. Wang, Y., Kung, L. & Byrd, T.A., 2016. Big Data Analytics: Understanding Its Capabilities and Potential Benefits for Healthcare Organizations (2016) *Technological Forecasting and Social Change*. Vol. 28. Data warehouse. SME. USA.
116. Abualoush, S., Masa'deh, R., Bataineh, K. and Alrowwad, A., 2024. The Role of Knowledge Management Process and Intellectual Capital as Intermediary Variables between Knowledge Management Infrastructure and Organization Performance (2024) *Informing Science Institute*. Vol. 12. Data warehouse. Small businesses. Jordan.
117. Shabani, I., Meziu, E., Berisha, B., 2022. Big Data Analytics in Cloud Computing: An Overview (2022) *Journal of Cloud Computing*. Vol. 16. Data warehouse. Various sectors including businesses, healthcare, and education. Mauritius.
118. Williams Jr, R.I., Smith, A., Aaron, J.R., Manley, S.C. and McDowell, W.C., 2018. Managing IT Strategic Planning in SMEs: A Strategic Alignment Perspective (2018) *International Journal of Business and Management*. Vol. 20. Data warehouse. SME. South Africa.
119. Meziu, E., 2023. Improving the ETL Process for a Case Company (2023) *Metropolia University of Applied Sciences*. Vol. 21. Database technologies. SME. Finland.
120. Pereira, C.M., 2024. Developing an Automated database System (2024) *IEEE* Vol. 10. Data warehouse. Small businesses. Global.
121. Abranches, M., 2020. Monitoring Framework for Clinical ETL Processes and Associated Performance Resources (2020) *IEEE Conference*. Vol. 12. Business technology. Small businesses. Portugal.
122. Sharma, S., Goyal, S.K. and Kumar, K., 2020. An Approach for Implementation of Cost Effective Automated Data Warehouse System (2020) *International Journal of Computer Information Systems and Industrial Management Applications*. Vol. 9. SME performance. Small businesses. India.
123. Yilmaz, N., Demir, T., Kaplan, S. and Demirci, S., 2020. Demystifying Big Data Analytics in Cloud Computing (2020) *Fusion of Multidisciplinary Research, An International Journal (FMR)*. Vol. 2. Database technologies. Small businesses. Turkey.
124. Salem, R., 2024. Optimal performance of database and data warehouse: A Review (2024) *IEEE* Vol. 1. Data warehouse. Small businesses. USA.
125. Orunbon, N.O., AI-Mamun, M.A., Naeem, Z., 2024. Data-Driven Decision Making: Advanced Database Systems for Business Intelligence (2024) *Nanotechnology Perceptions*. Vol. 18. Business technology. SME. Bangladesh.
126. Oliveira, L.O., Pereira, C.M., 2019. DOD-ETL: Distributed On-Demand ETL for Near Real-Time Business Intelligence (2019) *Journal of Internet Services and Applications*. Vol. 15. SME performance. Small businesses. Brazil.
127. Costa, C., Lebre, R., 2019., ETL Framework for Real-Time Business Intelligence over Medical Imaging Repositories (2019) *Journal of Digital Imaging*. Vol. 21. Database technologies. Startups. Venezuela.
128. Deb Nath, R.P., Romero, O., Pedersen, T.B. & Hose, K., 2022. High-level ETL for Semantic Data Warehouses (2022) *Semantic Web*. Vol. 20. Data warehouse. Startups. Denmark.
129. Das, S., Balakrishnam, R., 2020. Implementing Data Strategy: Design Considerations and Reference Architecture for Data-Enabled Value Creation (2020) *Australasian Journal of Information Systems*. Vol. 21. Business technology. Startups. India.
130. Tomar, D., Bhati, J.P., Tomar, P. and Kaur, G., 2019. Migration of Healthcare Relational Database to NoSQL Cloud Database for Healthcare Analytics and Management (2019) *Healthcare Data Analytics and Management*. Vol. 4. SME performance. Small businesses. USA.
131. Mondal, K.C., 2020. Efficient Incremental Loading in ETL Processing for Real-Time Data Integration (2020) *Innovations in Systems and Software Engineering*. Vol. 5. Database technologies. Startups. UK.
132. Dongmo, C., Van der Poll, J.A., Mositsa, J., 2023. Towards a Qualitative Framework for Data Management in Business Intelligence (2023) *IEEE Conference*. Vol. 15. Database technologies. Startups. South Africa.
133. Salem, R., 2018. A Cloud-Based Data Integration Framework for E-Government Service (2018) *IEEE Conference*. Vol. 16. Database technologies. Small businesses. Egypt.

134. Nwosu, T.N., 2024. Reducing Operational Costs in Healthcare Through Advanced BI Tools and Data Integration (2024) OSCM Publications, 2024. *Reducing Operational Costs in Healthcare Through Advanced BI Tools and Data Integration*. *Journal of Operations and Supply Chain Management*, vol. 2, pp. 787.
135. Chelliah, P.R., Raman, A., Nagaraj, D. and Duggirala, S., 2015. High-Performance Integrated Systems, Databases, and Warehouses for Big and Fast Data Analytics (2016) Semantic Web, 2016. *High-Performance Integrated Systems, Databases, and Warehouses for Big and Fast Data Analytics*. In: *Database Technologies* vol. 1, pp. 420. Book Chapter.
136. Ranjit, K., Sreemath Tirumala, S. & Nandigam, D., 2018. ETL tools for Data Warehousing: An empirical study of Open Source Talend Studio versus Microsoft SSIS (2018) Semantic Web, 2018. *ETL tools for Data Warehousing: An empirical study of Open Source Talend Studio versus Microsoft SSIS*. In: *Conference Paper* vol. 18, pp. 461.
137. Kimwele, M., Cheruiyot, W., Gatimu, M.R., 2016 Enhancing Data Staging as a Mechanism for Fast Data Access (2016) Science Data and Learning Machine Intelligence, 2016. *Enhancing Data Staging as a Mechanism for Fast Data Access*. *Thesis*, vol. 25, pp. 25.
138. Elbadri, M., Altaher, A. and Alkawan, S., 2024. Data Quality Considerations for ERP Implementation: Techniques for Effective Data Management (2024) Communications in Computer and Information Science, 2024. *Data Quality Considerations for ERP Implementation: Techniques for Effective Data Management*. In: *Conference Paper* vol. 32, pp. 96.
139. Mohamed, A., 2023. Framework of Big Data Analytics in Real Time for Healthcare Enterprise Performance Measurements (2023) Electronic Theses and Dissertations, 2023. *Framework of Big Data Analytics in Real Time for Healthcare Enterprise Performance Measurements*. *Thesis*, vol. 18, pp. 78.
140. Ahmed, N.A.S., Durmic, N., Shaari, H., 2021. Modern ABI Platforms for Healthcare Data Processing (2021) Lecture Notes in Networks and Systems, 2021. *Modern ABI Platforms for Healthcare Data Processing*. *Conference Paper*, vol. 5, pp. 89.
141. Arora, B., 2018. Big Data Analytics: The Underlying Technologies Used by Organizations for Value Generation (2018) Understanding the Role of Business Analytics, 2018. *Big Data Analytics: The Underlying Technologies Used by Organizations for Value Generation*. In: *Book Chapter* vol. 17, pp. 5.
142. Crowe, M., Offia, C., 2019. A Qualitative Exploration of Data Management and Integration in Organisation Sectors (2019) Electronic Theses and Dissertations, 2019. *A Qualitative Exploration of Data Management and Integration in Organisation Sectors*. *Journal of SME Performance*, vol. 15, pp. 90.
143. Nweke, F.H., Ajah, I.A., 2019. Big Data and Business Analytics: Trends, Platforms, Success Factors and Applications (2019) Big Data Cogn. Comput., 2019. *Big Data and Business Analytics: Trends, Platforms, Success Factors and Applications*. *Journal*, vol. 5, pp. 5.
144. Elbadri, M., Altaher, A., & Alkawan, S., 2017. A Big Data system supporting Bosch Braga Industry 4.0 strategy (2017) International Journal of Information Management, 2017. *A Big Data system supporting Bosch Braga Industry 4.0 strategy*. *Journal*, vol. 1, pp. 153.
145. Aromaa, A., Metso, L., Karri, T., 2014. Improving Reporting Management with Relational Database Management System (2014) Ecological Solutions and Evidence, 2014. *Improving Reporting Management with Relational Database Management System*. *Thesis*, vol. 16, pp. 27.
146. Jeble, S., Kumari, S. and Patil, Y., 2018. Role of Big Data in Decision Making (2018) OSCM Publications, 2018. *Role of Big Data in Decision Making*. *Thesis*, vol. 21, pp. 67.
147. Vera-Baquero, A., Colomo-Palacios, R., Stantchev, V., & Molloy, O., 2015. Leveraging Big-Data for Business Process Analytics (2015) The Learning Organization, 2015. *Leveraging Big-Data for Business Process Analytics*. *Journal*, vol. 16, pp. 45.
148. Fargo, W., Verma, R., 2024. Data Engineering and AI for BI Efficiency: Roles, Permissions, and Cloud Storage (2024) International Numeric Journal of Machine Learning and Robots, 2024. *Data Engineering and AI for BI Efficiency: Roles, Permissions, and Cloud Storage*. *Journal*, vol. 13, pp. 89.
149. Gahnouchi, S.A., Hassani, A., 2017. A framework for Business Process Data Management based on Big Data Approach (2017) Procedia Computer Science, 2017. *A framework for Business Process Data Management based on Big Data Approach*. *Conference Paper*, vol. 9, pp. 14.
150. Morawiec, P. & Sołtysik-Piorunkiewicz, A., 2020. The new role of Cloud Technologies in Management Information Systems Implementation Methodology (2020) Proceedings of the Future Technologies Conference (FTC) 2020, Volume 3, 2020. *The new role of Cloud Technologies in Management Information Systems Implementation Methodology*. *Conference Paper*, vol. 5, pp. 3.
151. Somasundaram, P., 2022. Hybrid Data Management Systems: Integrating Data Warehouses and Data Lakes (2022) Science Data and Learning Machine Intelligence, 2022. *Hybrid Data Management Systems: Integrating Data Warehouses and Data Lakes*. *Conference Paper*, vol. 20, pp. 75.
152. Belwal, M., Kumar, S.M, 2017. Performance Dashboard: Cutting-Edge Business Intelligence and Data Visualization (2017) 2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon), 2017. *Performance Dashboard: Cutting-Edge Business Intelligence and Data Visualization*. *Conference Paper*, vol. 14, pp. 13.

153. Shubham, K., 2020. Integrating Product Data to Enterprise Data Warehouse (2020) Master's Thesis, 2020. *Integrating Product Data to Enterprise Data Warehouse*. *Conference Paper*, vol. 12, pp. 33.
154. Abdel-Rahman, M. and Younis, F.A., 2022. Developing an Architecture for Scalable Analytics in a Multi-Cloud Environment for Big Data-Driven Applications (2022) International Journal of Business Intelligence and Big Data Analytics, 2022. *Developing an Architecture for Scalable Analytics in a Multi-Cloud Environment for Big Data-Driven Applications*. *Journal*, vol. 7, pp. 55. .
155. Palanivel, K., Manikandan, J., 2019. Business Analytics Architecture Stack to Modern Business Organizations (2019) International Journal of Computer Sciences and Engineering, 2019. *Business Analytics Architecture Stack to Modern Business Organizations*. *Conference Paper*, vol. 32, pp. 63.
156. Somepalli, S., Sikha, V.K., Korada, L., 2024. Finding the Right Data Analytics Platform for Your Enterprise (2024) International Journal of Science and Research (IJSR), 2024. *Finding the Right Data Analytics Platform for Your Enterprise*. *Journal*, vol. 7, pp. 45.
157. Younesi, E., Patel, K., Kalaitzopoulos, D., 2016. Advancements in Data Management and Data Mining Approaches (2016) Translational Medicine, 2016. *Advancements in Data Management and Data Mining Approaches*. In: *Book Chapter* vol. 7, pp. 28.
158. Mohanapriya, M., Amuthabala, P., 2016. Cost Effective Framework for Complex and Heterogeneous Data Integration in Warehouse (2016) Advances in Intelligent Systems and Computing (Volume 465), 2016. *Cost Effective Framework for Complex and Heterogeneous Data Integration in Warehouse*. *Conference Paper*, vol. 18, pp. 1.
159. Abdelhafez, H.A., 2014. Big Data Technologies and Analytics: A Review of Emerging Solutions (2014) International Journal of Business Analytics (IJBAN), 2014. *Big Data Technologies and Analytics: A Review of Emerging Solutions*. *Journal*, vol. 12, pp. 16.
160. Chen, C.L.P. & Zhang, C.-Y., 2014. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data (2014) Information Sciences, 2014. *Data-intensive applications, challenges, techniques and technologies: A survey on Big Data*. *Journal*, vol. 23, pp. 2474.
161. Nyirenda, M., Kaluba, Z., 2024. Database Migration Service With A Microservice Architecture (2024) IOSR Journal of Computer Engineering, 2024. *Database Migration Service With A Microservice Architecture*. *Conference Paper*, vol. 25, pp. 56.
162. Bibire Seyi-Lande, O., Johnson, E., Adeleke, G.S., Amajuoyi, C.P. and Simpson, B.D., 2024. Enhancing business intelligence in e-commerce: Utilizing advanced data integration for real-time insights (2024) International Journal of Management & Entrepreneurship Research, 2024. *Enhancing business intelligence in e-commerce: Utilizing advanced data integration for real-time insights*. *Journal*, vol. 17, pp. 798.
163. Haider, S. & Sen, S., 2024. Developing scalable data solutions for small and medium enterprises: Challenges and best practices (2024) International Journal of Management & Entrepreneurship Research, 2024. *Developing scalable data solutions for small and medium enterprises: Challenges and best practices*. *Journal*, vol. 6, pp. 2246. .
164. Al-Okaily, A., Teoh, A.P., Al-Okaily, M., Iranmanesh, M., & Al-Betar, M.A., 2023. The efficiency measurement of business intelligence systems in the big data-driven economy: a multidimensional model (2023) Information Discovery and Delivery, 2023. *The efficiency measurement of business intelligence systems in the big data-driven economy: a multidimensional model*. *Journal*, vol. 2, pp. 364.
165. Krishnasamy, S., 2016. Big Data Insight: Data Management Technologies, Applications and Challenges (2016) International Journal of Control Theory and Applications, 2016. *Big Data Insight: Data Management Technologies, Applications and Challenges*. *Journal*, vol. 25, pp. 580.
166. Awoyelu, I., Omodunbi, T., & Udo, J., 2014. Bridging the Gap in Modern Computing Infrastructures: Issues and Challenges of Data Warehousing and Cloud Computing (2014) Computer and Information Science, 2014. *Bridging the Gap in Modern Computing Infrastructures: Issues and Challenges of Data Warehousing and Cloud Computing*. *Journal*, vol. 18, pp. 23.
167. Yang, J., Bao, M., 2019. The impact of knowledge management process and business intelligence on organizational performance (2019) Management Science Letters, 2019. *The impact of knowledge management process and business intelligence on organizational performance*. *Journal*, vol. 12, pp. 53.
168. Ghosh, R., Haider, S. & Sen, S., 2015. An Integrated Approach to Deploy Data Warehouse in Business Intelligence Environment (2015) Proceedings of the 2015 Third International Conference on Computer, Communication, Control and Information Technology (C3IT), 2015. *An Integrated Approach to Deploy Data Warehouse in Business Intelligence Environment*. *Conference Paper*, vol. 5, pp. 5.
169. Scabora, L. de C., Brito, J. J., Ciferri, R. R. and Ciferri, C. D. de A., 2016. Physical Data Warehouse Design on NoSQL - OLAP Query Processing over HBase (2016) Proceedings of the 18th International Conference on Enterprise Information Systems - Volume 1: ICEIS, 2016. *Physical Data Warehouse Design on NoSQL - OLAP Query Processing over HBase*. *Conference Paper*, vol. 18, pp. 1230.
170. Lourenço, J.R., Cabral, B., Carreiro, P., Vieira, M. & Bernardino, J., 2020. NoSQL and Master Data Management: Revolutionizing Data Storage and Retrieval (2020) Open Journal Systems, 2020. *NoSQL and Master Data Management: Revolutionizing Data Storage and Retrieval*. *Journal*, vol. 8, pp. 34.

 doi:10.20944/preprints202410.0059.v1

171. Wang, J., Omar, A.H., Alotaibi, F.M., Daradkeh, Y.I. & Althubiti, S.A., 2022. Business intelligence ability to enhance organizational performance and performance evaluation capabilities by improving data mining systems for competitive advantage (2022) Information Processing & Management, 2022. *Business intelligence ability to enhance organizational performance and performance evaluation capabilities by improving data mining systems for competitive advantage*. *Journal*, vol. 9, pp. 19.
172. Costa, E., Costa, C. A., & Santos, M. Y., 2017. Efficient Big Data Modelling and Organization for Hadoop Hive-Based Data Warehouses (2017) European, Mediterranean, and Middle Eastern Conference on Information Systems (EMCIS), 2017. *Efficient Big Data Modelling and Organization for Hadoop Hive-Based Data Warehouses*. *Conference Paper*, vol. 27, pp. 14.
173. Kubina, M., Koman, G. & Kubinova, I., 2015. Possibility of improving efficiency within business intelligence systems in companies (2015) Procedia Economics and Finance, 2015. *Possibility of improving efficiency within business intelligence systems in companies*. *Conference Paper*, vol. 12, pp. 36.
174. Harrison, R., Parker, A., Brosas, G., Chiong, R., & Tian, X., 2015. The role of technology in the management and exploitation of internal business intelligence (2015) Journal of Systems and Information Technology, 2015. *The role of technology in the management and exploitation of internal business intelligence*. *Conference Paper*, vol. 8, pp. 78.
175. Shekhar, S. 2020. An In-Depth Analysis of Intelligent Data Migration Strategies from Oracle Relational Databases to Hadoop Ecosystems: Opportunities and Challenges (2020) International Journal of Applied Machine Learning and Computational Intelligence, 2020. *An In-Depth Analysis of Intelligent Data Migration Strategies from Oracle Relational Databases to Hadoop Ecosystems: Opportunities and Challenges*. *Journal*, vol. 3, pp. 49.
176. Prasad, S., Balachandran, B.M., 2017. Challenges and Benefits of Deploying Big Data Analytics in the Cloud for Business Intelligence (2017) Procedia Computer Science, 2017. *Challenges and Benefits of Deploying Big Data Analytics in the Cloud for Business Intelligence*. *Journal*, vol. 7, pp. 563.
177. Moscoso-Zea, O., Andres-Sampedro, & Luján-Mora, S., 2016. Data Warehouse Design for Educational Data Mining (2016) 15th International Conference on Information Technology Based Higher Education and Training (ITHET), 2016. *Data Warehouse Design for Educational Data Mining*. *Conference Paper*, vol. 18, pp. 20.
178. Massaro, A., Vitti, V., Savino, N., 2019. A Business Intelligence Platform Implemented in a Big Data System Embedding Data Mining: A Case of Study (2019) International Journal of Data Mining & Knowledge Management Process (IJDKP), 2019. *A Business Intelligence Platform Implemented in a Big Data System Embedding Data Mining: A Case of Study*. *Journal*, vol. 16, pp. 1204.
179. Rudniy, A., 2022. Data Warehouse Design for Big Data in Academia (2022) Computers, Materials & Continua, 2022. *Data Warehouse Design for Big Data in Academia*. *Journal*, vol. 11, pp. 45.
180. Niu, Y., Ying, L., Yang, J., Bao, M. and Sivaparthipan, C.B., 2021. Organizational business intelligence and decision making using big data analytics (2021) Information Processing & Management, 2021. *Organizational business intelligence and decision making using big data analytics*. *Journal*, vol. 8, pp. 124.
181. Al-Aqrabi, H., Liu, L., Hill, R. and Antonopoulos, N., 2015. Cloud BI: Future of business intelligence in the Cloud (2015) Journal of Computer and System Sciences, 2015. *Cloud BI: Future of business intelligence in the Cloud*. *Journal*, vol. 4, pp. 43.
182. Data Warehousing Process Modeling from Classical Approaches to New Trends: Main Features and Comparisons (2022) Data, 2022. *Data Warehousing Process Modeling from Classical Approaches to New Trends: Main Features and Comparisons*. *Journal*, vol. 7, pp. 56.
183. Chongwatpol, J., 2016. Managing big data in coal-fired power plants: a business intelligence framework (2016) Industrial Management & Data Systems, 2016. *Managing big data in coal-fired power plants: a business intelligence framework*. *Journal*, vol. 12, pp. 2473.
184. Bharadiya, J., 2023. A Comparative Study of Business Intelligence and Artificial Intelligence with Big Data Analytics (2023) American Journal of Artificial Intelligence, 2023. *A Comparative Study of Business Intelligence and Artificial Intelligence with Big Data Analytics*. *Journal*, vol. 23, pp. 77.
185. Nicoleta (2023). Big Data in Retail: *A Revolution [Use Cases and Examples]*. [online] www.tokinomo.com. Available at: https://www.tokinomo.com/blog/big-data-in-retail.
186. Zarour, M., Alenezi, M., Ansari, M.T.J., Pandey, A.K., Ahmad, M., Agrawal, A., Kumar, R. and Khan, R.A. (2021). *Ensuring data integrity of healthcare information in the era of digital health. Healthcare Technology Letters*, [online] 8(3), pp.66–77. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8136763/.
187. Kondabolu, S.S. and Nasina, J. (2010). A Virtual Data Warehouse for Manufacturing Industry. SSRN Electronic Journal. doi:https://doi.org/10.2139/ssrn.1544806.
188. Dutta, S. (n.d.). *Big Data in Finance: Benefits, Use Cases, and Examples*. [online] www.turing.com. Available at: https://www.turing.com/resources/big-data-in-finance.

 doi:10.20944/preprints202410.0059.v1