



FACULTY OF TECHNOLOGY

**A COMPARATIVE ANALYSIS OF GOOD ENTERPRISE DATA  
MANAGEMENT PRACTICES: INSIGHTS FROM LITERATURE AND  
ARTIFICIAL INTELLIGENCE PERSPECTIVES FOR BUSINESS  
EFFICIENCY AND EFFECTIVENESS**

Md Mehedi Hasan Tito

INDUSTRIAL ENGINEERING AND MANAGEMENT

Master's thesis

July 2023

## ABSTRACT

A comparative analysis of good enterprise data management practices: Insights from literature and artificial intelligence perspectives for business efficiency and effectiveness

Md Mehedi Hasan Tito

University of Oulu, Industrial Engineering and Management, Major in Product Management.

Master's thesis, 74 pp. + 19 appendixes

Supervisor(s) at the university: Janne Härkönen

This thesis aims to conduct a comparative analysis of good data management practices based on both literature and artificial intelligence perspectives. The specific focus is on their impact on data quality, business efficiency, and effectiveness. The study seeks to identify key challenges, potential areas for improvement and propose evidence-based recommendations to enhance data management processes within enterprises. The ultimate goal is to provide valuable insights that can guide companies in optimizing their data management strategies to better support their business objectives and drive overall performance.

A systematic and comprehensive research methodology is employed, consisting of three main stages. The first stage involves an extensive literature review to recognize and define good data management practices. The second stage utilizes artificial intelligence tools such as ChatGPT to examine current data management practices, identify key challenges, and analyze how companies measure data quality. The third stage comprises a comparative review of data management strategies from both the literature and AI viewpoints in order to identify discrepancies, similarities, and potential areas for improvement.

The comparative analysis gives critical insights for improving organizational data management capabilities and overcoming obstacles. The findings emphasize the significance of strong data governance, excellent data quality, efficient data integration, and comprehensive data security measures. AI technology, such as machine learning, improves data management activities such as analytics and decision-making. The research contributes to the scientific debate by demonstrating the transformational

potential of AI in EDM and its implications for information systems, data science, and business management.

The limitations include this study being primarily qualitative in nature and focusing on a comparative analysis of data management practices as presented in the literature and as understood by AI. The general standardization of the findings may be limited by the scope of the sources and the specific AI tools employed.

This research has important implications for firms trying to optimize their data management strategies and employ AI-driven methods to improve data quality, business efficiency, and effectiveness. This study's recommendations can be applied to real-world business circumstances, thereby driving organizations toward improved data management practices.

This thesis uniquely combines literature and AI perspectives in a comparative analysis of good data management practices, highlighting the potential of AI in enhancing traditional data management approaches and offering practical recommendations for enterprises.

*Keywords* – Enterprise data management practices, artificial intelligence, data quality, business efficiency, effectiveness, comparative analysis, literature review, ChatGPT, data management challenges.

## FOREWORD

This thesis embarks on a comprehensive exploration of good data management practices in the context of enterprise data management (EDM). The primary objective of this study is to conduct a comparative analysis of these practices as viewed through the lens of literature and artificial intelligence (AI), with a specific focus on their impact on data quality, business efficiency, and effectiveness.

I am sincerely grateful to my supervisor, Dr. Janne Härkönen, whose deep knowledge in the field and invaluable guidance shaped the trajectory of this work. His insightful inputs and constant encouragement have significantly contributed to the successful completion of this thesis. Furthermore, I extend my appreciation to the Department of Industrial Engineering and Management at the University of Oulu for the administrative support, which has been instrumental in carrying out this research.

My heartfelt thanks extend to my opponent researchers and friends who have provided unyielding support and constructive critique throughout this process. A special thank you to my family for their unwavering belief in me and their encouragement during challenging periods.

Finally, I must mention the extensive use of these master's thesis guidelines by the Faculty of Technology at the University of Oulu, which have served as a roadmap in structuring and formatting this document. I highly recommend these guidelines to all future master's students at the university.

Oulu, Finland

July 31, 2023.

Md Mehedi Hasan Tito

# TABLE OF CONTENTS

ABSTRACT

FOREWORDS

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

1	Introduction .....	8
1.1	Study Background .....	8
1.2	Research problem, objectives, and scope .....	9
2	Literature review .....	10
2.1	Data .....	10
2.2	Product data .....	11
2.3	Master data .....	12
2.4	Product master data .....	12
2.5	Business data .....	13
2.6	Transactional data .....	14
2.7	Data quality .....	15
2.8	Product .....	16
2.9	Product structure .....	17
2.10	Variant logic/ Fit, Form & Function (FFF rule).....	18
2.11	Product Data Management (PDM).....	19
2.12	Business process.....	21
2.13	Enterprise applications and their orchestration .....	22
2.14	EDM & Business performance .....	23
2.15	EDM & Operational performance .....	25
2.16	EDM challenges .....	26
2.17	Theory Synthesis .....	27
3	Data management practices as viewed by AI.....	36
3.1	Research method and collection of material .....	36
3.2	ChatGPT and its application in this study.....	37
3.3	Artificial Intelligence-reviewed EDM Practices .....	37
4	Comparative analysis .....	45
4.1	Comparison of Literature-based and AI-reviewed Practices .....	45
4.2	Addressing key challenges in Effective EDM .....	47
5	recommendations and discussion.....	49

5.1 Development proposals .....	49
5.2 Scientific implications .....	51
5.3 Managerial Implications for Business Efficiency and Effectiveness.....	52
6 Conclusions .....	54
6.1 Key findings .....	54
6.2 Validity and reliability of the study.....	55
6.3 Limitations and future research.....	57
REFERENCES .....	60

#### APPENDICES:

Appendix 1. Questionnaire posed to ChatGPT to ascertain AI-reviewed EDM best practices.

Appendix 2. Collected responses from ChatGPT.

## LIST OF ABBREVIATIONS

AI	Artificial intelligence
BI	Business intelligence
BPM	Business process management
CAD	Computer-aided design
CAE	Computer-aided engineering
CAM	Computer-aided manufacturing
CRM	Customer relationship management
DM	Data management
DW	Data warehouse
EDM	Enterprise data management
ERP	Enterprise resource planning
HW	Hardware
IoT	Internet of Things
KPI	Key performance indicator
MDM	Master data management
ML	Machine learning
NPD	New product development
NLP	Natural language processing
PDM	Product data management
PLM	Product lifecycle management
PM	Product management
PMD	Product master data
MDM	Master data management

# 1 INTRODUCTION

## 1.1 Study Background

The rapid growth of digital technology and the explosion of data generated by various sources have made data management a critical aspect of modern business operations (Lichtenthaler, 2022). Enterprises need to harness the power of data effectively to gain a competitive advantage, improve decision-making processes, and achieve business efficiency and effectiveness (Brynjolfsson, et al., 2011). As a result, data management has emerged as a key area of concern for companies across various industries.

Historically, data management practices have evolved in response to the increasing volume, variety, and complexity of data. Traditional data management approaches have often focused on structured data storage and retrieval (Abu-Elkheir et al., 2013), with limited emphasis on data quality, integration, and analytics. However, the proliferation of digital transformation (Lindstrom, 2020), the advent of advanced analytics techniques (Nadeem et al. 2021; Bidgoli, 2011), and the growing importance of data-driven decision-making (Chen, 2012; Wamba et al., 2015) have necessitated a shift in focus towards more comprehensive and efficient data management practices.

The literature on enterprise data management highlights the importance of adopting a holistic approach to managing data, encompassing aspects such as data governance (Allen & Cervo, 2015; Waddington, 2008), data quality management (Ofner et al., 2013; Shahin, 2014), data integration (Martijn et al., 2015; Kadadi 2014), and business process management (Doebeli et al., 2011). A strong data management strategy is crucial for ensuring data quality, which in turn, has a direct impact on business efficiency and effectiveness (Etwaroo, 2019).

Artificial intelligence (AI) has evolved as a strong tool for analyzing and managing large amounts of data in recent years. AI-driven data management methods (Farshad et al. 2020) offer the potential to improve traditional data management approaches by automating time-consuming operations, using data analytics to deliver more accurate insights, and allowing for the real-time execution of data-driven choices.



This study intends to explore the intersection of these two areas, concentrating on the role of AI in improving data management for business efficiency and effectiveness. This is because data management techniques are expanding and AI is becoming increasingly important. The purpose of this study is to examine the current state of data management practices as observed by both literature and AI in order to provide relevant insights that can assist organizations in improving their data management strategies and utilizing the increasing potential of AI to drive overall performance.

## **1.2 Research problem, objectives, and scope**

The aim of this thesis is to conduct a comparative analysis of good data management practices as viewed by the literature and artificial intelligence with a specific focus to assess their impact on data quality, business efficiency, and effectiveness. Based on the findings, this study aims to integrate the key challenges, potential areas for improvement, and propose evidence-based recommendations to enhance data management processes within enterprises. Ultimately, this study aims to provide valuable insights that can guide companies in optimizing their data management strategies to better support their business objectives and drive overall performance. The aim is divided into three research questions.

Research questions:

RQ1 What is good data management that supports data quality and business efficiency and effectiveness?

RQ2 How are the current data management practices viewed by the artificial intelligence, what are the key challenges in managing the data effectively, and how do companies measure data quality?

RQ3 How do the data management practices as presented by literature compare to those presented by AI?

## 2 LITERATURE REVIEW

### 2.1 Data

Data refers to any collection of information that is gathered and analyzed to gain insights, inform decision-making, or support a particular task. Data has become a critical asset for modern businesses, providing valuable insights and enabling evidence-based decision-making (Alhassan et al., 2018). As the volume of data generated continues to grow exponentially, managing and utilizing these data effectively has become a significant challenge for enterprises (Nadeem et al., 2021). Data management has been identified as a key strategic priority for organizations looking to gain a competitive advantage in today's data-driven economy (Wamba et al., 2015).

Data can originate from various sources, such as customer interactions, business processes, social media, and machine sensors, among others (Nadeem et al., 2021). However, the sheer volume, velocity, and variety of data can make it difficult to manage and analyze effectively (Nadeem et al., 2021). Data quality issues, such as incomplete or inaccurate data, can undermine the usefulness of data and lead to incorrect conclusions (Wang & Strong, 1996).

To enhance the effectiveness and efficiency of enterprise data management, organizations can implement various strategies and best practices. For instance, using data governance frameworks can help establish policies, standards, and procedures for managing data throughout their lifecycle (Alhassan et al., 2018). Adopting data analytics tools and techniques can enable organizations to extract valuable insights from large datasets and make evidence-based decisions (Nadeem et al., 2021). Additionally, utilizing big data and cloud computing technologies can aid firms in effectively processing and storing enormous volumes of data (Wamba et al., 2015). According to Lee et al. (2014), data can be clustered into two categories such as traditional data and big data. Traditional data are the basis for an organization's operations (Lee et al., 2014) where they comprise master data, transactional data, and reference data.

Effective data management has become crucial for organizations to gain a competitive advantage in today's data-driven economy. By implementing appropriate strategies and

best practices, organizations can enhance the efficiency and effectiveness of their data management processes, leading to improved organizational performance (Alhassan et al., 2018).

This thesis aims to explore strategies and best practices for enhancing the effectiveness and efficiency of enterprise data management, from the perspective of a tech product manufacturer, with the goal of improving organizational performance and competitiveness.

## **2.2 Product data**

Product data is crucial to manage and track products throughout the supply chain and to provide accurate and complete information to customers, which is essential for sales, delivery, and service (Silvola et al., 2019). In the Industry 4.0 era, product data can include additional information such as sensor data, maintenance records, and usage data, which can be used to optimize product design and performance as well as enhance operational excellence (Bousdekis et al., 2021). A product, from a business perspective, refers to goods, knowledge, or services that are capable of being sold and provide value to the customer by meeting their requirements. This can take the form of tangible physical objects, intangible services, or a combination of both, as described by Kropsu-Vehkaperä (2012).

However, the quality of product data is crucial to ensure its accuracy, completeness, and consistency (Choi & Luo, 2019). According to Choi (2019), poor data quality can lead to supply chain disruptions, inefficient operations, and negative customer experiences. Managing product data carefully is therefore essential to support efficient sourcing and production operations, optimize product design and performance, and provide a positive customer experience. To ensure high-quality product data, organizations need to implement data quality management practices and data governance frameworks (Abraham et al., 2019). This can be achieved through the use of data quality metrics and tools such as data profiling, data cleansing, and data enrichment (Brous et al., 2016; Abraham et al., 2019).

## 2.3 Master data

Master data is a crucial component of enterprise data management, representing the core data elements related to the transactional and analytical operations (Das & Mishra, 2011). According to Silvola (2018), master data and business data form product data where the master data are created during the NPD (new product development) phase and once it is created, the master data remains unchanged. In another article Silvola et al., (2011) explained that master data includes product development related master data and product item master data which itself contains some attributes like product weight, version code, item group, price, costing. Master data is hence a critical component of an organization's business operations, according to Otto's (2012) definition. It refers to the fundamental entities that are essential to an organization, such as customer, employee, finance, patient, product, location, material, and supplier data (Khatri, 2016). Master data include critical information like product specifications, descriptions, pricing, and related details, which remain the same for the rest of the product lifecycle and they are made available for use by other functional departments and business processes (Silvola et al., 2019; Dahlberg et al., 2011; Vilminko-Heikkinen 2017). Master data are characterized by their independence and behavior, as they stand on their own and are not reliant on other data types. According to Gartner, (2022), master data is highly valued in the business world due to their essential role in describing key business data and other data types their being dependent on them. Customers and sales items are common objects of master data and serve as the foundation for categorizing transactional data, making it an integral part of business reporting (Vilminko-Heikkinen 2017). Master data serve as the foundation for connecting business transactions with business entities, including customers, products, and locations (Dahlberg et al., 2011). The proposed framework by Dahlberg et al. (2011) addresses the challenges of comprehensively managing master data in distributed environments.

## 2.4 Product master data

Product master data is a comprehensive and accurate source of information about a company's products and is used to support various business functions such as sales, marketing, supply chain, and logistics. Product master data plays a critical role in ensuring that accurate and consistent product information is available across different business

units and systems and is essential for making informed business decisions (Silvola, 2018). Product master data is the basic key information for product related business processes (Silvola et al. 2011).

In today's business environment, effectively taking care of product master data is critical for organizations to stay competitive in the market. For example, the increasing importance of digital channels for sales and marketing and the rise of e-commerce have made it essential for companies to manage their product master data effectively. According to Wang and Strong (1996), effective product master data management involves maintaining accurate and consistent information about products and ensuring that this information is readily accessible to customers and other stakeholders.

However, people who create product master data must understand that their actions affect many downstream life-cycle phases in the organization (Kropsu-Vehkapera, 2012). If the product data content is not correctly understood, the quality or purpose of using the product data across IT systems for other business processes is adversely affected (Kempainen, 2010). Key internal stakeholders for creating product master data include sales and marketing, supply chain, and after-sales teams (Kropsu-Vehkapera and Haapasalo, 2011). Ultimately, effective product master data management can help organizations improve their performance and competitiveness in today's data-driven business environment (Wang & Strong, 1996).

## **2.5 Business data**

Business data refers to the vast amounts of information generated by various business processes, including sales, customer interactions, and operational activities. According to Silvola (2018), business data pertains to the processes of a business, such as marketing and sales data related to the products, supply chain data, and data related to services and customer care. The author further noted that these types of data play a significant role in supporting business transactions by being utilized within various business processes. According to Harkonen et al., (2019), business data plays a crucial role in enabling effective data management and analysis for the productization or product structure as it provides relevant information about ownership, stakeholders, requirements, and parameters of the related product offerings.

Effective analysis of business data can lead to numerous benefits, including better customer engagement and retention. Business data and the product master data are different in nature, as business data can be changed and are not as much as critical in terms of minor issues related to the data quality (Harkonen et al., 2019). Because business data also includes information that is critical to the day-to-day operation of the organization, such as customer and vendor data, supply chain data, and service-related data. According to the Harkonen et al., (2019), product master data is kind of DNA (human identical genom) for the product whereas business data can be termed as the product related additional business information. This data is used by various departments within the organization, including sales, marketing, finance, and customer service.

## **2.6 Transactional data**

Transactional data is a structured and detailed type of data generated through commercial transactions and business applications like ERP and CRM, where transactions are mandatorily performed based on product master data (Hannila, 2019). This data can be used to analyze purchasing habits, consumer behavior, spending patterns, and other trends (Ghinita et al., 2011). According to Prayitno and Astuty (2017), transactional data analysis is useful in customer relationship management since it provides insights into customer behavior, preferences, and purchase history, which can be used to personalize marketing campaigns, improve customer experience, and increase customer loyalty.

The emergence of IoT has transformed business operations, where transactional data plays a vital role in the digital landscape. This data are crucial for real-time monitoring and control of IoT devices and identifying their usage pattern (Petnik et al., 2019). Furthermore, online shopping contributes significantly to the massive volumes of transactional data in e-commerce. Thus, the emergence of the Internet and other technologies have ensured that these large transactional data are stored, processed, and analyzed with speed (Kwarteng, et al., 2017). The exponential increase in online shopping is clear evidence of the growth in electronic buying and selling (Kwarteng et al., 2017). According to multiple sources (Schaffer, 2017; Silvola, 2011; Silvola, 2019; Kumar and Mishra, 2011; Thornthwaite, 2015), master data is a necessity for all transactions in business applications. With the emergence of smart products equipped with sensors,

microprocessors, and software, as well as connectivity to other products, equipment, people, and ecosystems, vast amounts of usage and connectivity data are generated, which provide new analytics and business intelligence opportunities for companies (Borga, 2014; Porter, 2014).

Transactional data records specific actions or events that occur within a business, such as sales, deliveries, purchases, emails, and revocations. Purchase orders, business invoices, payments, deliveries, and storage records are examples of transactional data, which organizations often need to analyze customer behavior, preferences, and buying patterns. These data can provide valuable insights into consumer behavior, policy analysis, and product innovations as well (Allcott et al., 2019; Harding et al., 2012; Jaravel, 2019; Macedoni et al., 2020). According to Chalupa (2017), transactional data are particularly useful in market segmentation since they are a great source of customer behavioral data, which allows marketing managers to identify the most effective marketing strategies for different customer segments.

According to Madera C., (2020), there is a significant relationship between master data and transactional data. The author explained that the relationship between master data and transactional data can be viewed as a noun/verb relationship where transactional data captures the verbs and master data the nouns (Madera, 2020). While master data is typically less volatile than transactional data, entities with attributes that do not change at all may not require a master data solution (Madera, 2020, p.127-129).

## **2.7 Data quality**

Data quality is a crucial factor in ensuring that transactions fulfill their intended purpose (Haug & Arlbjørn, 2011). The fitness for use of data can be assessed by evaluating its quality dimensions, which include accuracy, reliability, timeliness, relevance, and completeness (Etwaroo, 2019). However, organizations face challenges in managing data quality due to the lack of data management actions, including centralized data management, data owners, data stewards, data governance, and adequate data procedures and tools (Etwaroo, 2019).

Quality of data can be poor due to the lack of data management actions, which can lead to significant costs for organizations, such as lost opportunities, mistakes, and failed deliveries or invoicing problems (Silvola et al., 2011; Etwaroo, 2019). Therefore, organizations need to implement effective data management actions to ensure high-quality master data and avoid costly data quality issues (Haug & Arlbjørn, 2011; Etwaroo, 2019).

As the world moves towards the Industry 4.0 standard, the volume of data generated by machines, processes, and services will significantly increase in the future, as noted by Awwad et al. (2018) in their literature review on big data analytics in the supply chain. According to Knolmayer and Rothlin (2006), the increasing amount of data in companies is causing data quality problems, particularly in the management of material master data, which can affect the usefulness of distributed ERP systems. This is a common issue faced by companies in today's digital age.

However, the availability of high-quality data is crucial for effective demand forecasting and related operations such as inventory planning. Without it, these processes can be negatively impacted, as highlighted by Choi and Luo (2019) in their study on data quality challenges for sustainable fashion supply chain operations in emerging markets. They also suggest potential solutions such as the use of blockchain technology, government sponsors, and environmental taxes.

## **2.8 Product**

A product can take many different forms, and may be a combination of physical objects, services, or knowledge. According to Kropsu-Vehkaperä and Haapasalo (2011), a product can be difficult to define as it can be a combination of hardware, software, or both. From a business perspective, a product is a saleable item that provides value to the customer and meets their requirements (Kropsu-Vehkaperä, 2012). This can include goods, knowledge, or services.

Products can be tangible physical objects or intangible services, or a combination of both (Kropsu-Vehkaperä, 2012). As existing products are upgraded, new versions are created that replace the old ones, often with improvements in quality and performance, and



reduced costs (Kropsu-Vehkaperä et al., 2011b). These new versions can include both big and small enhancements, which may be referred to as product revisions (Kropsu-Vehkaperä, 2012) or variants.

As products are upgraded by the time being, new versions are created with enhancements and revisions that can improve quality and performance and reduce costs. These new versions may also include customer requirements, which can lead to a sequence of revisions and upgrades. Additional customer requirements can be incorporated into new product variants, but this does not necessarily increase the number of sellable products. Instead, it increases the sequence of revisions or upgrades to the existing product (Kropsu-Vehkaperä, 2012).

## **2.9 Product structure**

The concept of product structure is used to refer productization of a physical product (Hardware) and services commercially and technically (Tolonen et al., 2014; Harkonen et al., 2017). Product structure is a hierarchical, item-based representation of a product (Janardanan et al., 2008). The product structure concept is useful to respond, manage, and fulfil stakeholder requirements (Salvador 2007). According to Haines (2014), a product is not just a single item, rather it consists of a hierarchy of products and services. It might be a part of another product or bundle of parts assembled purposefully to create a feature. There might be valid reasons for the change in the product feature i.e., cost revision, quality enhancement (Harkonen et al. 2017). Therefore, the physical or service products require productization to make it commercially and technically feasible (Harkonen et al., 2018a, 2018b, 2018c). Productization is thereby widely interpreted as the process of transforming features or customer requirements into a sellable product offering (Harkonen et al., 2015; Suominen et al., 2009; Flamholtz 1995). Like the product structure concept, productization can also be applied to any type of products or any service (Harkonen et al. 2017; Harkonen et al., 2018a; 2018b; 2018c; Mustonen et al., 2019a). Service productization helps in developing more concrete and object-like service products (Wirtz et al., (2021). In their article, Wirtz et al. (2021) suggest to productize a service product to become less complex, less-failure prone, and a more enjoyable service encounter to maximize customer satisfaction. So, it is important to understand the logic of the productization that how to create sales items and associated version items, and how

to describe a company's product portfolio. Hierarchically, solution level is the highest level of the product structure/productization. The solution level can be divided into different product families or sub-families (Tolonen et al. 2014). Going further to divide the product family and sub-family, we get the smaller entity of the product which is the product configuration level. In this level we get configurable products. The product is configured according to the customer's needs (Lahtinen et al., 2021).

According to Janardanan et al. (2008), the product structure is a vital component of the product database, which facilitates the exchange of product information among various users and systems. The integration of product structure into contemporary enterprise information systems, such as product data management systems, has prompted researchers to study it extensively. A clear product structure that defines the relationships between product components and their attributes is essential for effective productization. Productization involves organizing product data in a structured and standardized format that can be easily managed and shared across different stakeholders. In organizational data management, an effective product structure is crucial for efficiently managing product data and facilitating technical product upgrades, ultimately resulting in faster time-to-market commercially.

## **2.10 Variant logic/ Fit, Form & Function (FFF rule)**

In response to the recent growing demand for personalized products that meet the unique needs of different customer segments, variant logic has emerged as a popular product development strategy (ElMaraghy et al., 2013). Variants of a product are created to improve functional parts, reduce its price, performance and to improve quality (Lahtinen et al., 2021). So, this approach involves creating multiple versions of a product with different features and improvements, which allows businesses to better serve their customers and increase customer loyalty while reducing the risk of losing customers to competitors (Lestari & Putra, 2022; Retnowati et al., 2021). The use of common components across different product variations also enables companies to reduce production and development costs, providing greater value to customers and improving overall customer satisfaction and loyalty (Hakim, 2022). Because there are several factors that influence purchasing decisions, including product variants and brand image strength that eventually affect customer loyalty (Retnowati et al., 2021). This also increases

opportunities to gain sales growth and win competitor's market share by offering a wider range of products to their customers (Hakim, 2022; Solomon, 2015; Khayru & Issalilla, 2021).

The FFF rule, which stands for Fit, Form, and Function, is an important aspect of the variant logic approach (Lahtinen et al., 2021). This rule ensures that new product versions or variants meet the intended purpose, have an appropriate appearance, and function effectively and efficiently. By applying this approach, companies can create new product variants that meet the needs of specific customer segments or niches, while also improving the overall product family and complying with regulatory requirements (ElMaraghy et al., 2013). Adopting a variant management approach that considers the product, process, and market views can help companies control the range of product variants offered and manage their effects throughout their life cycle. Therefore, by considering the fit, form, and function of a product, engineers can create new versions or variants that improve quality, performance, or cost-effectiveness, while also meeting customer expectations (Lahtinen et al., 2021).

## **2.11 Product Data Management (PDM)**

Product Data Management (PDM) is part of a product lifecycle which ensures the availability of up-to-date, real-time data to different stakeholders throughout the product development process (Arnold et al., 2011; Lachmayer, 2015; Vajna et al., 2009; Feldhusen, 2013; VDI 2219, 2016). It is the process of exchanging, creating, maintaining, and controlling the information related to a product throughout its lifecycle (Eynard et al., 2004; Gu & London, 2010). PDM is also an effective way of controlling knowledge and learning processes as it maintains a central data repository (Silvola, 2018). Effective PDM supports collaborative engineering and management of product development projects, product structures, documents, and quality (Yang et al., 2007; Rouibah and Ould-Ali, 2007). PDM systems are used during the product development phase to save, manage, and provide product models and information (Eigner & Stelzer, 2009; Scheidel, 2017) which conventionally focuses on managing product models, product structure, documents, and development-relevant information (Arnold et al., 2011; Lindner, 2015). These pieces of information are stored in various data management systems which can be

further shared and used by various departments of the organization for the product related decision making throughout its lifecycle (Zhang et al., 2017).

In another perspective, Silvola et al., (2011) explained PDM as a system that is used to manage all product master data and product-related business data where this system focuses on data creation, storing, disseminating, and archiving, and is considered as a sub-function of the Product Lifecycle Management (PLM) system. According to Scheidel et al., (2017), PDM systems collect lifecycle data, verify product related information, and transfer them back to the product development phase. Therefore, PDM system provides a compatible infrastructure with appropriate interfaces for controlling and distributing data for users, connecting product data related product and process management. Under such infrastructure, PDM stores typical product information such as brand name, part number, part description, supplier details, unit of measure, cost and price, detailed CAD drawings, material datasheets, and vendor descriptions with version control and document security in database (Silvola et al., 2011; Scheidel et al., 2017).

Another important aspect of product data management is data governance, which involves establishing policies and procedures for managing data throughout the organization (Abraham et al., 2019). According to Abraham et al., 2019, effective data governance can help to ensure that product data is consistent and accurate across different departments and functions within the organization. Because data governance is related to the decisions and the decision makers, whereas product data management is related to making those decisions as a part of the day-to-day execution of governance policies for the particular product (Dyché & Levy, 2006, pp. 150; Haggmann, 2013, pp. 234; Khatri & Brown, 2010, p. 148; Otto, 2013, p. 96). This is particularly important in the case of product manufacturers, where product data may be used by different teams for a variety of purposes, including data exchange, product design, manufacturing, and marketing (González et al., 2019; Lucas-Estañ et al., 2018).

In the context of Industry 4.0, the role of product data management is highlighted by Bousdekis et al., (2021) where the author demonstrated that the PDM has become even more critical for companies with the advent of advanced technologies like Internet of Things (IoT) and Artificial Intelligence (AI) to remain competitive in the market.

According to Bousdekis et al., (2021), PDM significantly help companies streamline their processes, reduce costs, and improve overall efficiency.

## **2.12 Business process**

Business processes provide a way of working for organizations by referring to a set of coordinated activities or tasks to achieve specific goals, produce products, or deliver services (Mansar & Reijers, 2005; Silvola et al., 2011; Silvola et al., 2016). Mansar & Reijers (2005) also explain that these processes are the backbone of operational efficiency and effectiveness, as they streamline workflows and facilitate better resource utilization. Consequently, organizations are increasingly focusing on business process management (Dallas & Wynn, 2014) to optimize their operations, identify bottlenecks, and enhance overall performance (Hammer, 2010).

Business process management (BPM) is an approach that involves continuous improvement, automation, and monitoring of business processes to achieve strategic objectives and satisfy customer needs (van der Aalst, 2013). According to Doebeli et al. (2011), BPM is a part of company's strategy. It is needed by organizations which often entail the use of technology, such as BPM software, to support the design, execution, and analysis of processes (Harmon, 2014). According to Doebeli et al. (2011), organizational targets are part of a company's strategy, and thus it is necessary that senior management understand the role of BPM as a holistic management system, including the IT systems. BPM is not a clear-cut topic, however, as company needs and the use of BPM varies due to practicalities and common sense (Doebeli et al., 2011). The adoption of BPM can result in significant benefits, including increased agility, improved compliance, and cost reduction (Rosemann & vom Brocke, 2015).

In the context of data management, business processes play a crucial role in ensuring that data is collected, stored, and processed efficiently and securely. Effective data management processes help organizations to make data-driven decisions, maintain data quality, and protect sensitive information (Chen, 2012).

### 2.13 Enterprise applications and their orchestration

The growing significance of data management in the modern digital economy has led organizations to invest in advanced enterprise IT systems for improved management, analysis, and utilization of their data (Wixom et al., 2013). Data gathered from business operations is stored and maintained within these IT systems. Additionally, business processes are interconnected through IT systems, where data pertaining to process control, performance management, and reporting for internal and external stakeholders is acquired (Genaroro & Lourero, 2015; Tian & Quan, 2008). IT systems manage enterprise data, ensuring their uniqueness, consistency, traceability, and dependability through various solutions like Enterprise Resource Planning (ERP), Customer Relationship Management (CRM), and Data Warehouse (DW) systems (Moss, 2007). CRM helps enterprises to enhance customer loyalty and satisfaction and sustain them for long-term (Herman et al., 2020). The research findings by Elfarmawi (2019) explains that CRM systems have a perfect positive linear relationship with customer satisfaction and has a significant association with product innovation. Payne and Frow (2005) and Azad and Ahmadi (2015) provide comprehensive reviews of CRM systems, which can be helpful for companies to avoid potential problems due to the lack of properly defined CRM. Coltman (2007) discusses the challenges and opportunities associated with CRM systems implementation where human, technological and business capabilities require orchestration.

ERP systems serve as comprehensive software solutions that unify various organizational functions, streamlining processes and enhancing data management (Silvola, 2018). They are often seen as the central IT system for supply chain operations (Hannila, 2019). Significant research on ERP systems and their effects on data management have been conducted by Oliveira et al. (2014) and Ram et al. (2014). PDM systems function as a single, centralized database designed to support all business processes (Silvola, 2018). The management of product data attributes is not limited to PDM systems but can also be handled by systems such as ERP (Gomez et al., 2014; Sonzini et al., 2015; Sriti et al., 2015; Tidstam and Malmqvist, 2015).

CRM, ERP and service/care systems are the main IT systems needed during the product life-cycle (Silvola et al., 2019). CRM and ERP systems form a precondition for one

product data (Silvola et al., 2019) where a CRM system keeps track of interactions with existing and potential new customers (Gomez et al., 2014; Meyliana et al., 2016), storing data throughout the entire interaction lifecycle. An ERP system encompasses critical business processes such as manufacturing, planning, marketing, sales, shipping, and invoicing, acting as a central system for business transactions (Jain, 2016).

Within the literature on enterprise data management, data warehousing (DW) is identified as a suite of decision-support systems designed to empower information professionals (e.g., leaders, administrators, and analysts) in improving and accelerating decision-making (Simões 2010; Al-Debei 2011). This highlights the significance of data warehouse systems as a critical element of business intelligence, which is widely adopted by medium and large organizations for effective data management.

The integration and orchestration of different IT systems are crucial to achieving efficient data management in organizations (Irani et al., 2003; Wixom et al., 2013). Key studies in this area include those by Al-Mashari and Zairi (2000) on ERP system integration, as well as Ranganathan and Brown (2006) on the integration of CRM systems. Moreover, the role of Service-Oriented Architecture (SOA) in orchestrating and integrating these systems has been explored by Papazoglou and van den Heuvel (2007). Recently there is an increasing trend to adopt cloud-based solutions and AI techniques in replacing of traditional IT systems integrations (Bidgoli, 2011). This trend is allowing organizations to meet challenges of implementing and orchestrating IT applications in terms of data privacy and other security concern, which is eventually improving organizational data management processes (Ghods et al., 2017).

## **2.14 EDM & Business performance**

The impact of Enterprise Data Management (EDM) on business efficiency has been a recurring theme in recent literature. Centralized data management and the elimination of data silos streamline business processes and workflows, leading to increased operational efficiency (Steinfeld et al., 2011). Recently organizations have striven to make sense of the vast amounts of data they generate and collect. The effective management of this data can significantly influence business performance, particularly in terms of data quality, business efficiency, and effectiveness.

Data quality is a fundamental aspect of EDM, as it directly affects the accuracy and reliability of business insights derived from data (Ahola, 2022). Several studies have emphasized the importance of data quality in decision-making and the subsequent impact on business performance. Improved data quality has been associated with enhanced decision-making capabilities, leading to better strategic and operational decisions (Brous et al., 2016; Chakravorty, 2020; Ashdown et al., 2021). Chakravorty (2020) added that maintenance process, data ownership and their responsibilities are to be significantly clear to improve data quality. Because quality data ensures transparency and cooperation between different business units that eventually improves the business operations.

Research highlights the crucial role of data governance, standardization, and validation processes in ensuring high-quality data and enhancing business performance (Brous et al., 2016). Data governance establishes processes and structures that transform an organization's behavior to improve and safeguard data as a strategic asset, ultimately boosting performance (Cleven & Wortmann, 2010; Dreibelbis et al., 2008; McGilvray, 2006).

Data governance, as a key data management practice, directly impacts business performance by addressing both organizational and technical aspects that require leadership, authority, control, and resource allocation (Lucas, 2010; Mosley & Brackett, 2010; Ashdown et al., 2021). With a focus on shared decision-making over the management of an organization's data assets, data governance exercises authority and control, aiming to ensure the quality, integrity, security, and usability of the data (Brous et al., 2016; van Gils, 2020; Ashdown et al., 2021). This ultimately leads to enhanced trust in the data, which is vital for driving business performance (Ashdown et al., 2021).

Data integration and analytics tools also enhance business efficiency. Organizations that successfully implement data integration solutions achieve significant improvements in efficiency through reduced manual labor and quicker access to critical information (Martijn et al., 2015). Advanced analytics and big data technologies optimize resource allocation and reduce operational costs, resulting in efficiency gains (Bughin et al., 2010). EDM contributes to business effectiveness by aligning strategic goals with operational activities, leading to improved performance (Chaki, 2015). As data has evolved from a



byproduct of business processes to a valuable asset (Lake and Crowther 2013), measuring business performance and compliance reporting have driven demand for improved data quality in enterprise systems (Martijn et al., 2015).

Effective data governance and management are crucial for companies to survive in complex business environments (Brous et al., 2016). Well-integrated processes and consistent data management result in real-time business intelligence (BI), a competitive advantage (Thusoo & Sharma, 2017). Holistic data governance increases business performance compared to competitors, positively affecting time-to-market, profitability, and labor costs when faultless and reliable data is shared quickly and automatically (Allen & Cervo, 2015).

Business process management (BPM), closely tied to EDM, captures organizational targets, and presents them in a way that stakeholders can understand the changes needed to meet targets (Doebeli et al., 2011). BPM is based on process leadership, governance, performance, strategic alignment, people capability, project execution, and technology (Jeston & Nelis, 2008).

## **2.15 EDM & Operational performance**

EDM enhances operational performance by guaranteeing that data are precisely maintained, readily available, and efficiently managed across the organization (Kropsu-Vehkaperä, 2012). This is particularly important given the recent advancements in experimentation and measurement technologies, which enable the collection of vast quantities of high-dimensional data (Igarashi et al., 2016). By addressing such challenges related to data quality, integration, and governance, businesses can achieve better results in terms of efficiency, effectiveness, and compliance. The relationship between EDM and operational performance is evident in the literature. For instance, Arlbjørn et al. (2007) found that operating with valid master data leads to improved company performance. This is supported by Otto et al. (2011) and Silvola et al. (2016), who emphasize the importance of data quality in master data management. Organizations that focus on improving data quality at different levels and stages can enhance their overall performance (Ofner et al., 2013; Shahin, 2014).

Data governance, being a company-specific activity (Brous et al., 2016), is another aspect that impacts operational performance. Aligning data governance with business needs and fostering a shared understanding of data within the organization are critical for successful data management (Waddington, 2008). However, extant literature tends to focus more on IT systems and algorithms, rather than data governance and analytics (Brous et al., 2016; Kwon et al., 2014). The integration of business processes through IT systems is essential for improving operational performance (Genaroro & Lourero, 2015; Tian & Quan, 2008). By effectively managing their data, organizations can better align their strategic goals with operational activities, leading to improved business performance.

Business Process Management (BPM) also plays a role in enhancing operational performance by capturing organizational targets and ensuring that employees at all levels understand the necessary changes to meet those targets (Doebeli et al., 2011). BPM can facilitate the alignment between company strategy and IT systems, thus contributing to improved operational performance (Wisner & Stanley, 2008).

## **2.16 EDM challenges**

One of the major components of enterprise data management is the PDM. The processes of PDM are crucial for maintaining high-quality product data and ensuring their availability across organizations. Larger organizations often struggle with keeping product data up-to-date, even when using automated processes (Kropsu-Vehkaperä 2012). Manual tasks such as data entry and content checking are susceptible to human error, which can lead to inaccuracies and inconsistencies in product data. These practical challenges can be particularly difficult to manage during product ramp-ups and ramp-downs. One approach to mitigating these challenges is establishing clear data ownership within organizations (Kropsu-Vehkaperä 2012). However, defining data ownership can be difficult in practice and in consequence, unclear data ownership can result in incoherent data management practices (Silvola et al., 2011). Furthermore, many companies lack data governance models or data quality controls (Silvola et al., 2016). This can lead to poor data quality, disparate global processes, and insufficient guidance by ERP systems (Ahola, 2022).

Another challenge in ensuring high-quality product-related master data is the absence of holistic practices for deriving master data from corporate strategies. Data integration, or consolidating data from diverse sources, is critical for business intelligence (Kadadi, 2014). Companies often invest in tools to facilitate data integration, but integrating large volumes of data from various sources can still be a complex and challenging task.

In the era of big data, another key issue is the lack of understanding and knowledge professionals who can effectively manage and analyze these vast datasets (Nambiar & Mundra, 2022). This may result from a shortage of skilled data professionals, unclear communication with employees, or improper use of data repositories, among other factors.

Data security is also a significant challenge, as mishandling sensitive customer information can lead to costly data breaches (Toshniwal et al., 2014; Demchenko et al., 2014). Additionally, implementing and integrating different IT systems requires careful planning and execution (Chen, 2012; Cuzzocrea et al., 2013). Maintaining data quality and consistency is crucial, as is reconciling differences in naming conventions, domain definitions, and identification numbers from heterogeneous sources (Singh, 2010; Nambiar & Mundra, 2022). Also, extracting value from data is challenging. Different data utilization methods such as data analytics can provide a competitive advantage for businesses (Giebler et al., 2019). To maximize data utilization and value extraction, data must be made available to a wide range of users. This necessitates enterprise-wide metadata management and data transparency (Zhu, 2002). In another view Redman (2008) explored that, many managers mistakenly assume that IT systems will automatically address data challenges, highlighting the importance of studying product data in the context of integrated business processes.

## **2.17 Theory Synthesis**

A combination of strategies and technology tools is required for good enterprise data management practices that promote data quality and business efficiency and effectiveness. Implementing data governance frameworks, according to Alhassan et al. (2019), can assist define policies, standards, and processes for managing data throughout its lifecycle, hence improving the efficacy and efficiency of business data management.

According to Nadeem et al. (2021), it is critical to use data analytics tools and methodologies to extract meaningful insights from massive datasets and make evidence-based judgments. Organizations can improve their performance and competitiveness in today's data-driven business environment by applying these strategies and best practices (Alhassan et al., 2019). Furthermore, using cloud computing and big data technologies can help businesses analyze and store huge amounts of data efficiently (Wamba et al., 2015).

Ensuring high-quality product data is essential in organizational data management for accurate and effective decision-making across various business functions (Choi & Luo, 2019). Abraham et al. (2019) suggest that organizations need to implement data quality management practices and data governance frameworks to maintain the quality of their product data. This can be achieved with data quality metrics and tools such as data profiling, data cleansing, and data enrichment (Brous et al., 2016).

Effective product master data management is critical for organizations to stay competitive in the market (Wang & Strong, 1996). According to Silvola (2018) and Harkonen et al. (2019), master data and business data form product data, which plays a significant role in supporting business transactions and various business processes. Ensuring the accuracy, completeness, and consistency of transactional data is crucial for effective demand forecasting and related operations such as inventory planning (Choi & Luo, 2019). Managing business related product data is another critical element in EDM that supports the availability of up-to-date, real-time data to various stakeholders throughout the product development process (Arnold et al., 2011; Lachmayer, 2015; Vajna et al., 2009; Feldhusen, 2013; VDI 2219, 2016). This also involves exchanging, creating, maintaining, and controlling product-related information throughout the product's lifecycle (Eynard et al., 2004; Gu & London, 2010). Product data management (PDM) ensures consistent and accurate information across different departments and functions within the organization (González et al., 2019; Lucas-Estañ et al., 2018). Especially, PDM systems crucially performs in streamlining processes, reducing costs, and improving overall efficiency, which is essential for businesses to remain competitive in today's rapidly evolving market (Bousdekis et al., 2021).

Data governance is another important aspect of good organizational data management, which involves establishing policies and procedures for managing data throughout the organization (Abraham et al., 2019). Data governance ensures that product data is used effectively for various purposes, such as data exchange, product design, manufacturing, and marketing (González et al., 2019; Lucas-Estañ et al., 2018). By implementing effective data governance, organizations can better align their decision-making processes with their overall governance policies (Dyché & Levy, 2006, pp. 150; Haggmann, 2013, pp. 234; Khatri & Brown, 2010, p. 148; Otto, 2013, p. 96). It directly affects the quality, integrity, security, and usability of data (Brous et al., 2016; van Gils 2020; Ashdown et al. 2021). Implementing strong data governance models and establishing clear data ownership within organizations can help mitigate challenges related to data quality (Kropsu-Vehkaperä 2012; Silvola et al., 2011).

Good data management also involves several key components, as identified in the literature review. These components include Business Process Management (BPM), data quality, data integration, and the use of advanced IT systems, such as Enterprise Resource Planning (ERP) and Customer Relationship Management (CRM). BPM is important for aligning organizational targets with operational activities, ensuring that employees understand the necessary changes to meet objectives (Doebeli et al., 2011). Whereas, data integration is essential for consolidating data from diverse sources, enabling effective business intelligence (Kadadi 2014). Successfully implementing data integration solutions can lead to improved efficiency through reduced manual labor and quicker access to critical information (Martijn et al., 2015). By effectively managing data, organizations can achieve better alignment between their strategic goals and operational activities, leading to improved business performance (Wisner & Stanley, 2008). Furthermore, advanced IT systems, such as ERP and CRM enhance data management by unifying various organizational functions, streamlining processes, and facilitating better resource utilization (Silvola 2018; Herman et al., 2020). The integration and orchestration of these systems are crucial for achieving efficient data management (Irani et al., 2003; Wixom et al., 2013). Data quality needs to be guaranteed for making better strategic and operational decisions for which data accuracy, reliability, and consistency is achieved through clear maintenance processes and well-defined data ownership responsibilities (Chakravorty 2020; Brous et al., 2016). Comprehensively all these key components form a holistic approach that leads to improved decision-making, enhanced operational

performance, and ultimately, better business outcomes (Chakravorty 2020; Steinfield et al., 2011; Bughin et al., 2010).

Especially in the context of Industry 4.0, good data management has become increasingly important due to the integration of advanced technologies like the Internet of Things (IoT) and Artificial Intelligence (AI) (Bousdekis et al., 2021). Advanced analytics and big data technologies are thus considered as key aspects in enterprise data management. It optimizes resource allocation, reduces operational costs, and improve overall business performance (Bughin et al., 2010; Giebler et al., 2019). Investing in skill development of data professionals and necessary training have become essential in good data management, as it addresses the shortage of skilled data professionals and promotes clear communication with employees to facilitate effective data management and analysis (Nambiar & Mundra, 2022). This is also related to the data security which nowadays is another essential perspective for protecting sensitive customer information and avoiding costly data breaches (Toshniwal et al., 2014; Demchenko et al., 2014). Figure 1 illustrates the synthesized framework for good enterprise data management practices.

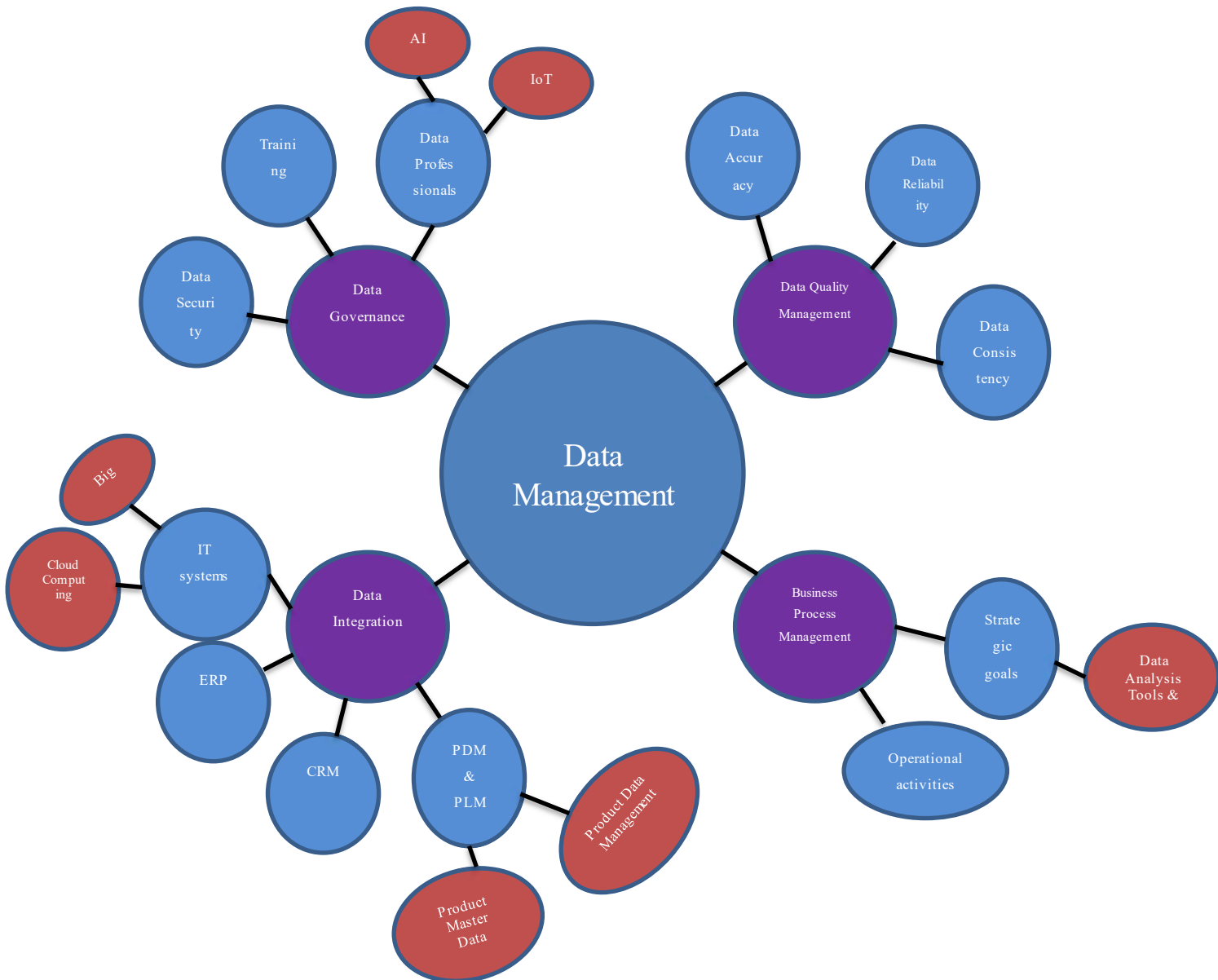


Figure 1. Framework for good enterprise data management practices

### Literature-Based EDM Practices

Several practices emerge in the literature review as crucial for enhancing the effectiveness and efficiency of Enterprise Data Management (EDM). In today's data-driven business climate, these strategies assist organizations in improving their data management capabilities, improving organizational performance, and gaining a competitive advantage. The following paragraphs outline some literature-based EDM practices that organizations can implement. Adopting these practices organizations can ensure the effective and efficient management of their data assets which eventually leads enterprises towards

better decision-making, improved customer experiences, streamlined business processes, and increased innovation, ultimately contributing to the overall success of the organization.

*Data Governance Frameworks:* Creating a data governance framework assists in the development of rules, standards, and processes for managing data throughout its lifecycle. It assures the consistency, correctness, and security of data across the organization's many sources and systems. Furthermore, good data governance promotes collaboration and coordination among the various parties in charge of creating, preserving, and utilizing data.

*Data Quality Management:* Implementing data quality management methods ensures that data is accurate, full, and consistent. This may entail the use of data quality measures and technologies such as data profiling, cleansing, and enrichment. Having high-quality data can lead to better decision-making, better customer experiences, and more operational efficiency.

*Master Data Management:* MDM is an important component of EDM because it provides a single source of truth for critical data items relevant to business transactions and analytics. MDM enables effective corporate processes and decision-making by ensuring the integrity, availability, and timeliness of master data, such as product, location, and customer. The MDM framework is concerned with maintaining master data in remote environments, which facilitates departmental communication and decision-making.

*Product Master Data Management:* For a company to succeed, its product master data must be accurate and consistent. Maintaining accurate and consistent information about items and making sure that customers and other stakeholders can easily access this information are both essential components of effective product master data management.

*PDM (Product Data Management) systems:* PDM systems ensure the availability of up-to-date, real-time data to various stakeholders throughout the product development process as part of the product lifecycle. Product models and information are saved, managed, and provided by PDM systems. As a result, good PDM facilitates collaborative engineering and management of product development projects, product structures,



documents, and quality. As a result, it contributes to process streamlining, cost reduction, and overall efficiency. Furthermore, in the context of Industry 4.0, the function of PDM becomes even more crucial for organizations to remain competitive in the market with the development of new technologies such as the Internet of Things (IoT) and Artificial Intelligence (AI). PDM greatly aids businesses in streamlining their processes, lowering costs, and increasing overall efficiency.

*Product Structure, Productization and variant logic:* Product data is managed and technical product upgrades are facilitated by effective productization and a clear product structure. It is also necessary for responding to, controlling, and meeting the needs of stakeholders. Organizations should prioritize the development of well-defined product structures that can be readily controlled and shared among many stakeholders. Productization and organizational data management require a clear product structure that describes the relationships between product components and their properties. Organizations can build new product variants that suit consumer expectations while increasing overall product family performance by incorporating variant logic and adhering to the Fit, Form, and Function (FFF) guideline in product development. This method enables businesses to better service their clients and enhance customer loyalty while also lowering manufacturing and operating costs.

*Business Data Management:* Effectively managing business data enables firms to capitalize on important information created by various business operations such as sales, customer contacts, and operational activities. This is useful for decision-making, marketing, and customer interaction.

*Transactional Data Management:* Organizations must manage transactional data in order to assess purchasing behaviors, consumer behavior, spending patterns, and other trends. This data can be utilized to personalize marketing campaigns, enhance the consumer experience, and boost customer loyalty.

*Business Process Management (BPM):* Through BPM, firms can optimize operations, eliminate bottlenecks, and improve overall performance by aligning business processes with strategic objectives. Process leadership, governance, performance, strategic

alignment, people competency, project execution, and technology should all be considered in BPM.

*Data Analytics Tools and Techniques:* Adopting data analytics tools and methodologies enables firms to gain valuable insights from large datasets and make evidence-based decisions. Organizations invest in data analytics capabilities in order to examine various sorts of data, such as business data, transactional data, and product data. Integrating cutting-edge analytics techniques such as machine learning, natural language processing, and artificial intelligence has become critical for improving an organization's ability to use data in strategic decision-making. These tools aid in the recognition of patterns and trends, the generation of predictions, and the simplification of company procedures. Furthermore, data analytics contributes to better consumer interactions, more successful marketing campaigns, and more refined product development processes.

*Using Cloud Computing and Big Data Technologies:* Use of cloud computing and big data technologies allows enterprises to efficiently process and store enormous amounts of data. This enables firms to get new insights and improve decision-making through more effective data management and analysis.

*Collaboration and communication:* It is critical to ensure good collaboration and communication across the various departments and stakeholders engaged in the creation, management, and use of data in order to improve the efficacy and efficiency of corporate data management.

*Continuous improvement, skill development, training, and monitoring:* Implementing continuous improvement and monitoring techniques enables organizations to discover areas for improvement in their data management processes and make the required changes to improve efficiency and effectiveness. Furthermore, investing in data professionals' skill development and giving appropriate training to promote clear communication with employees and assist successful data administration and analysis. Concurrent evaluation and training programs, regular audits, performance measurement, and feedback mechanisms are all part of this.

*Data integration and centralized data management:* Data integration is the process of merging data from diverse sources and systems in order to generate a single picture of an organization's data assets. Data consolidation, data federation, and data virtualization are all effective data integration practices. Organizations create centralized data management systems to assure consistency and accuracy of data across diverse departments and functions to facilitate data integration. This strategy encourages improved collaboration and aids in the preservation of a single source of truth for product data. Data integration and centralized data management solutions assist enterprises in breaking down data silos, improving data accessibility, and improving data quality.

*Data security and privacy:* As organizations handle massive volumes of sensitive data, ensuring data security and privacy is critical. To safeguard their data assets from illegal access and breaches, organizations use comprehensive data security methods such as data encryption, data masking, and data access controls. To safeguard the privacy of personal data, enterprises must also comply with relevant data protection rules, such as the General Data Protection Regulation (GDPR) in Europe.

*Data storage and management:* Adequate data storage and management methods ensure availability, dependability, and performance of data assets. Organizations assess various data storage alternatives, such as on-premises, cloud-based, and hybrid storage solutions, and then deploy the optimal solution for their needs. Furthermore, data management strategies including data archiving, backup, and lifecycle management assist firms in maximizing their data storage resources and lowering expenses.

*Data sharing and cooperation:* Data sharing and collaboration among stakeholders improves the potential value of an organization's data assets. Organizations create data-sharing rules and procedures, as well as collaboration tools and platforms, to allow employees to access and exchange data in a secure and efficient manner. This results in better decision-making, more innovation, and higher organizational performance.

## **3 DATA MANAGEMENT PRACTICES AS VIEWED BY AI**

### **3.1 Research method and collection of material**

The research methodology comprises with a systematic and comprehensive approach to understanding good data management practices, focusing on their impact on data quality, business efficiency, and effectiveness.

The study consists of three main stages.

1. The initial stage of this thesis involves a comprehensive literature review, where a thorough analysis of existing research is conducted to recognize and define good data management practices that support data quality and business efficiency and effectiveness. Through the analysis of the literature review, this study seeks to address RQ1, providing a well-founded understanding of good data management practices as demonstrated in the literature.
2. The second stage involves employing the Artificial Intelligence tools such as ChatGPT to examine the current data management approaches from an AI perspective, highlighting the key challenges confronted by enterprises and their methods for assessing data quality. This phase incorporates the use of ChatGPT to collect and analyze data from various sources, such as interviews, surveys, and case studies, in order to provide a comprehensive understanding of the AI viewpoint on data management practices. The analysis derived from this stage answers the RQ2.
3. In the third stage, a comparative analysis of the data management practices from both the literature review and AI perspective is conducted. This comparison highlights the differences, similarities, and potential areas for improvement, leading to evidence-based recommendations to enhance data management processes within enterprises. Consequently, the findings from this stage provides the answer to RQ3.

The overall nature of this research is qualitative. Therefore, the methodology has been designed to provide valuable insights to help companies optimize their data management

strategies and drive overall performance, ensuring the study's relevance and applicability to real-world business scenarios.

### **3.2 ChatGPT and its application in this study**

ChatGPT, a powerful artificial intelligence model (OpenAI, 2023), served as a vital instrument in the methodology of this study, making it capable of understanding prompts, generating text, and carrying out conversations much like a human would. Its ability to provide quick and comprehensive responses made it a reliable tool for data collection, particularly for generating AI-reviewed insights on the topic of effective Enterprise Data Management (EDM) practices. Earlier research has been calling for responsible and ethical ways to use ChatGPT to support professional work (Lund & Wang, 2023). Therefore, the tool was considered to support the research. ChatGPT having demonstrated abilities in natural language processing (Qin et al., 2023) makes it an interesting tool for the purpose of analyzing EDM practices.

In terms of application, the use of ChatGPT in this study allowed for a novel approach to data collection and empirical analysis. In essence, it has been viewed as a source of expert information where a series of questions related to the study was asked to gather AI perspectives for data management practices. Instead of treating the responses from ChatGPT as ready solutions, they were considered as raw data, for which a thorough thematic analysis (Herzog, et al., 2019) has been performed. This approach has made it possible to extract insights and patterns that complimented the outline of the conclusions. The use of AI perspectives compared with the traditional article sources represents an innovative move in academic research and broadens the horizons for future studies in the field of technology and data management.

### **3.3 Artificial Intelligence-reviewed EDM Practices**

This Master's thesis explores the best practices in Enterprise Data Management (EDM), a critical component of modern business operations that ensures data's reliability, accessibility, and security. Recognizing the dual importance of traditional literature and the evolving artificial intelligence perspectives in this field, a comprehensive exploration of these both have been set out to conduct.

In order to unravel the AI perspectives, the capabilities of ChatGPT have been engaged which embodies the practical applications of AI in data management. This unique AI tool served as a knowledge base to answer a set of targeted questions (Appendix 1) designed to elicit insights on proficient EDM practices.

The outcomes of the ChatGPT interaction provide a multidimensional view into EDM, encompassing its most significant components, the implications of machine learning, the significance of technical improvements, and more. This thesis attempts to detail these findings in order to contribute to the scholarly discourse on EDM while also providing practical recommendations for enterprises looking to optimize their data management methods.

The following section contains a detailed presentation of the questionnaire results, which were specifically developed to acquire insights on AI-assisted best practices for EDM from ChatGPT. The responses are organized in such a way that a thorough evaluation of each factor is possible, offering an in-depth grasp of AI's opinion on effective EDM techniques. They provide vital insights into the possible benefits and new tactics enabled by the integration of AI in EDM, making them an important component of this research.

The insights provided by ChatGPT, gathered in response to the questionnaire in Appendix 1 and Appendix 2 are compiled and presented as follows:

1. *The most critical aspects of an effective EDM system includes strong data governance to ensure accuracy, consistency, and security, and high-quality data to support informed decision-making. It must integrate data from various sources and provide easy access to users, ensuring user-friendly interaction. The system should be scalable, accommodating business growth and increasing data volume without compromising performance. Regulatory compliance, including data privacy laws, is a must, as is robust metadata management for an improved understanding of data assets. A well-structured data architecture facilitates efficient data retrieval and processing, while real-time processing capabilities enable prompt responses to changing business conditions.*

2. *Adopting a well-bounded data management strategy* has several advantages, including better decision-making due to access to high-quality, pertinent data, operational efficiency due to automation and optimized processes, regulatory compliance due to effective data governance, and better customer service due to a good understanding of customer requirement and their behaviours. Well-rounded data management strategy also enhances data accessibility by leading to actionable insights. It boosts efficiency and productivity by streamlining processes and minimizing errors, ensures regulatory compliance to avoid legal issues, and improves data quality. A strong data management strategy also enhances data security, provides cost savings by reducing errors and optimizing data utilization, and improves customer service with a comprehensive view of customer interactions. It prepares the enterprise for future growth by allowing seamless scalability of data infrastructure, provides a competitive edge by effectively utilizing data, and fosters a data-driven culture, enhancing data literacy among employees.
3. *Aligning business objectives with a data strategy* guides the relevance of the data collected, enhances informed decision-making, increase operational efficiency, and improves ROI on data investments. To achieve this alignment, enterprises should understand their business objectives, identify relevant data, and promote cross-team collaboration during strategy development. Clear, measurable data goals should be set those support business objectives, and the data strategy should be reviewed and adjusted regularly to keep it aligned with evolving business goals. Promoting a data-driven culture can ensure that the data strategy becomes an integral part of the business operations. This alignment is not a one-time process; it requires continuous communication, collaboration, and commitment to a data-driven approach.
4. *Data literacy*, an ability to interpret and communicate with data, is crucial for effective data-driven decision-making in organizations. To enhance data literacy, organizations should invest in training programs to educate employees about the value and usage of data. Promoting a data-driven culture, providing access to necessary data and tools, and encouraging cross-functional collaboration can further enhance understanding. Recruitment should also consider data literacy as

a valuable skill. Furthermore, fostering continuous learning can ensure employees keep up with evolving data science trends. Implementing mentorship programs can provide support for less experienced staff. By taking these steps, an organization can significantly improve its data literacy, leading to better utilization of data assets and improved decision-making processes.

5. *Maintaining high-quality data in an enterprise data management system* involves defining clear data quality standards, implementing a robust data governance framework, and validating data at entry points to prevent errors. Regular data audits and automation of data quality checks can identify and rectify issues, while a data cleansing process can correct identified errors. Integrating data from multiple sources ensures consistency, and training employees enhance their role in maintaining data quality. The quality of data should be tracked over time using relevant metrics to identify trends and demonstrate the effectiveness of data quality initiatives. Creating a data-centric culture that values high-quality data is essential, as maintaining data quality is an ongoing process that requires organization-wide involvement.
6. As a pivotal element, *data governance* ensures data quality, security, regulatory compliance, improved efficiency, and productivity, and informed decision-making. To implement data governance effectively, an organization should establish a Data Governance Council, comprising representatives from various areas of the organization, to develop and oversee the data governance strategy. Clear policies and procedures should be defined to regulate data handling, and data governance tools should be implemented to assist in tracking data assets, maintaining data quality, and protecting data. Assigning data stewards can help in overseeing the implementation of data governance policies across the organization. Employee training and education are essential to ensure everyone understands the policies and their roles in maintaining data integrity. Regular reviews and adjustments of policies and procedures are required to keep them current with changes in regulations, technology, and business objectives. This continuous focus on data governance provides a robust foundation for an organization's data management strategy.



7. A *well-structured data architecture* is important for efficient data storage, retrieval, and processing, maintaining data quality, enabling scalability, ensuring data security, and facilitating regulatory compliance. To achieve this, it's essential to understand the organization's data needs and develop a comprehensive data model that outlines how data will be organized, stored, and accessed. Selecting the right technology—be it databases, data warehouses, or data lakes—is important to meet specific data management requirements. Aligning data architecture with data governance strategy ensures consistency in data quality, security, and privacy. Additionally, planning for scalability and adopting solutions like cloud-based systems or distributed storage can ensure data architecture evolves with the growing needs. Testing and refining the data architecture is a continuous process that allows for optimization and adjustment over time.
8. Assessing an organization's *data management maturity* involves evaluating its data governance, data infrastructure, data culture, data quality, and data integration practices against established benchmarks or industry standards. Key elements include the presence of robust data governance policies, modern and efficient data storage and processing tools, a culture fostering data-driven decision-making, and processes ensuring data quality and effective data integration from multiple sources. Data management maturity models, like the one from CMMI Institute, can be used for a structured evaluation. Regular assessments allow an organization to identify areas for improvement, monitor progress, and ensure alignment with its goals.
9. *Data integration* aims to combine data from various sources into a unified view. Challenges in this process can be managed by standardizing data formats to ease the combination and usage of data and implementing data governance to regulate data handling, responsibility, and quality. Utilizing data integration tools can automate this process, ensuring consistency and reliability. Maintaining high data quality is critical, as is planning for scalability to accommodate future growth or changes. Addressing data security and privacy issues, considering a dedicated data integration specialist, and training the team on data management principles and specific tools are also important steps. Lastly, the data integration strategy should be regularly reviewed and adapted to respond to evolving data environments,

business needs, and technology advancements. This approach helps streamline data integration, supporting more effective enterprise data management.

10. Ensuring *data security and privacy* in enterprise data management may involve several strategic components. These include establishing a robust data governance policy that defines data handling, access, and protection rules, and implementing strong access controls and encryption to safeguard data. Regular audits are necessary to check for vulnerabilities and ensure policy compliance, as is anonymization or pseudonymization of sensitive data. Organizations should also have an incident response plan for potential data breaches, train all staff members on data security practices, and ensure system and software updates and patches are regularly installed. Compliance with relevant data protection regulations is obligatory. Employing data security tools like firewalls, intrusion detection systems, and data loss prevention tools can further enhance data security. These responsibilities require continuous attention and frequent policy reviews.
11. *Emerging technologies* are significantly influencing the future of EDM practices by introducing *new capabilities and opportunities*. Key among these are AI and ML (machine learning), which can automate and enhance data management tasks, including data cleansing and integration. IoT increases data volumes but also allows real-time data analysis. Cloud computing provides scalable and flexible data storage and processing solutions. Blockchain offers enhanced security and traceability, while 5G technology accelerates data collection and processing, supporting real-time analytics. Edge computing improves real-time decision-making by processing data closer to its source, and architectural approaches like data fabric and data mesh ease the management and access of distributed data. Data privacy technologies respond to increasing demands for secure data handling. Furthermore, an augmented data management strategy utilizes AI to enhance tasks like data quality and privacy management. Understanding and incorporating these evolving technologies are vital for future-proofing enterprise data management strategies.
12. *Artificial Intelligence (AI) and Machine Learning (ML)* are importantly optimizing data management through various ways, such as, they automate data

cleaning and pre-processing tasks, thereby enhancing efficiency and accuracy. AI and ML can automatically categorize and classify data, making retrieval easier. They are instrumental in anomaly detection, identifying data quality issues or potential security threats. Predictive analytics through AI and ML aid in decision-making across various business sectors. AI also streamlines data integration and transformation, simplifies metadata management, and optimizes query performance. AI's role in enhancing data security and privacy through anomaly detection and data anonymization is significant. AI-powered chatbots improve user interaction with data, and automated decision-making using AI offers real-time, effective solutions. While the role of AI and ML is vital, it's crucial to address the ethical considerations and potential biases proactively.

13. *Data management* has an important role in digital transformation that enables informed decision-making through the provision of accurate and timely data. It allows businesses to leverage customer data effectively, resulting in personalized experiences, a crucial aspect of digital transformation. Data management also enhances operational efficiency through automation, saves resources, and reduces errors. It also promotes innovation by providing a solid framework for deploying emerging technologies like as AI and ML. Furthermore, data management plays a vital role in risk management where it ensures data security and regulatory compliance amidst the increased data sharing in digital transformation. It also supports scalability, enabling easy expansion of digital technologies across the organization. Effective data management aids in establishing a data-driven culture, a fundamental part of digital transformation.
14. *Unstructured data*, such as text, images, and videos, can contain valuable information but is more difficult to manage and analyze than structured data. Effective strategies include implementing data cataloguing and metadata creation for easy retrieval and indexing. Establishing data governance policies can standardize practices around storage, security, and usage. Specialized tools and technologies, such as natural language processing (NLP) for text data and image recognition for visual data, can facilitate the analysis of unstructured data. Appropriate data storage solutions, often cloud-based for scalability and flexibility, are crucial. AI and machine learning can provide meaningful insights

from unstructured data, while robust data security measures ensure the protection of sensitive information. Effective data integration processes are essential when combining structured and unstructured data.

## 4 COMPARATIVE ANALYSIS

### 4.1 Comparison of Literature-based and AI-reviewed Practices

Both the literature-based practices and the AI-reviewed practices identified in this thesis offer valuable insights into effective Enterprise Data Management (EDM). While the literature-based practices provide a solid foundation based on established principles and best practices, the AI-reviewed practices leverage advanced technologies and data-driven approaches to enhance EDM. By combining the principles from the literature with the capabilities of AI technologies, organizations can optimize their data management strategies and drive business efficiency and effectiveness in today's data-driven digital environment. Here are the key comparisons of these two sets of practices:

#### ***Data Governance Framework:***

**Literature-Based:** The literature emphasizes the importance of establishing a comprehensive data governance framework that defines policies, standards, and procedures for managing data throughout its lifecycle. This framework ensures consistency, accuracy, and security of data.

**AI-Reviewed:** The AI-reviewed practices also highlight the significance of data governance but emphasize the use of AI tools to automate and enhance governance processes. AI can assist in data classification, access control, and policy enforcement, thereby augmenting the effectiveness of the data governance framework.

#### ***Data Quality Management:***

**Literature-Based:** The literature suggests implementing data quality management practices, such as data profiling, cleansing, and validation, to ensure accuracy and reliability. It emphasizes the use of metrics and tools to monitor and improve data quality continuously.

**AI-Reviewed:** The AI-reviewed practices reinforce the importance of data quality management but highlight the role of AI in automating data cleansing, anomaly detection, and data enrichment. Large datasets can be analyzed by AI to spot trends and outliers, which improves the accuracy and consistency of the data.

***Master Data Management (MDM):***

Literature-Based: MDM is recognized as a crucial component of EDM, ensuring a single source of truth for core data elements. The literature emphasizes the need for MDM frameworks to manage master data efficiently.

AI-Reviewed: MDM is similarly emphasized by the AI-reviewed methods, however they advise using AI algorithms to improve MDM procedures. AI can assist in data matching, deduplication, and synchronization across systems, ensuring the integrity and availability of master data.

***Data Security and Privacy:***

Literature-Based: The literature stresses the importance of robust data security measures, including encryption, access controls, and compliance with data protection regulations.

AI-Reviewed: The AI-reviewed practices align with the literature and highlight the potential of AI in enhancing data security and privacy. AI can help detect and respond to security threats, identify sensitive data patterns, and automate privacy compliance processes.

***Data Analytics and Business Intelligence:***

Literature-Based: In the current data-centric landscape, Data Analytics and Business Intelligence are often at the forefront of considerations on data management strategies. The literature emphasizes the use of data analytics and business intelligence tools to derive valuable insights from data for decision-making and process optimization.

AI-Reviewed: The AI-reviewed practices reinforce the role of data analytics and business intelligence but highlight the potential of AI in advanced analytics techniques, such as machine learning and natural language processing. AI can enable more sophisticated data analysis, prediction, and automation of insights generation.

***Continuous Improvement and Adaptability:***

Literature-Based: The literature emphasizes the need for continuous improvement and adaptability in data management practices to keep up with evolving technologies and business needs.

AI-Reviewed: The AI-reviewed practices align with the literature, emphasizing the importance of staying informed about emerging technologies and industry trends, and continuously updating data management strategies.

## **4.2 Addressing key challenges in Effective EDM**

Implementing and maintaining an effective EDM strategy presents a number of problems that must be addressed in order to achieve successful data management. Key difficulties have been found by a comparative comparison of literature-based practices and AI-reviewed practices, providing insights into areas where traditional practices and AI technology can merge to optimize data management operations. Addressing these issues is critical for firms looking to improve their data management capabilities and overcome roadblocks to optimal EDM. Several specific difficulties and accompanying solutions uncovered through comparative study are described below:

1. Challenge: Effective data governance and integration are vital for ensuring data consistency, accuracy, and accessibility across the organization. However, organizations face challenges in defining and implementing comprehensive data governance frameworks and integrating data from diverse sources.

Solution: Creating a solid data governance framework with defined roles, responsibilities, and data management processes. This framework should take into account data quality, data integration, data security, and compliance. AI technology can be used to automate data governance activities, improve data governance frameworks, and facilitate data integration.

2. Challenge: Data Quality Management Maintaining high data quality is crucial for reliable decision-making and operational efficiency. However, organizations struggle with data quality issues, including data inconsistencies, inaccuracies, and completeness.

Solution: Implementing data quality management practices that include data profiling, data cleansing, validation, and enrichment. Organizations should establish data quality metrics, employ automated tools for data quality monitoring, and combine AI techniques to detect anomalies, identify patterns, and improve data accuracy.

3. Challenge: Data Security and Privacy Protecting sensitive data from unauthorized access, breaches, and ensuring compliance with data protection regulations is a significant challenge for organizations.

Solution: Implementing robust data security measures, including encryption, access controls, and anonymization techniques. Implications of AI technologies can enhance data security through AI-based threat detection, anomaly detection, and compliance automation.

4. Challenge: Data Analytics and Business Intelligence Deriving valuable insights from data through effective data analytics and business intelligence is a challenge due to the increasing volume, variety, and complexity of data.

Solution: Using artificial intelligence (AI) technology such as machine learning and natural language processing to improve data analytics capabilities. Organizations can adopt advanced analytics tools and techniques to uncover hidden patterns, gain predictive insights, and automate data-driven decision-making processes.

5. Challenge: Continuous Improvement and Adaptability Keeping up with evolving technologies, industry best practices, and changing business needs is an ongoing challenge in EDM.

Solution: Embracing a culture of continuous improvement and adaptability by regularly evaluating and updating data management strategies. Organizations should invest in employee training and development to enhance data literacy and promote awareness of emerging technologies and trends.



## 5 RECOMMENDATIONS AND DISCUSSION

### 5.1 Development proposals

This thesis presents a synthesis of classical concepts and AI capabilities through a careful comparison with literature-based techniques, providing a more comprehensive understanding of effective EDM. This broad perspective is useful in today's data-driven world, when the capacity to successfully manage and exploit data is critical to company efficiency and effectiveness. The following recommendations are made based on the data and analyses presented in this thesis to improve Enterprise Data Management (EDM) processes and promote corporate efficiency and effectiveness:

*Establish a comprehensive data governance framework:* Organizations should develop and implement a data governance framework that outlines clear roles, responsibilities, and processes for managing data. This framework should include policies, standards, and procedures for data quality, security, privacy, and compliance. It should also promote cross-functional collaboration and communication to ensure consistent and coordinated data management practices throughout the organization.

*Invest in data quality management:* Organizations should prioritize data quality management practices to ensure accurate, complete, and reliable data. This involves implementing data profiling, data cleansing, and data validation techniques. Additionally, organizations should establish data quality metrics and measurement processes to continuously monitor and improve data quality over time.

*Implement a master data management (MDM) system:* Adopting an MDM system is recommended to centralize and manage critical master data across the organization. This includes data related to customers, products, suppliers, and other key entities. An MDM system ensures data consistency, eliminates data redundancies, and provides a single source of truth for accurate and reliable data. It is also suggested to leverage AI algorithms in the MDM system. These advanced technologies can assist in data matching, deduplication, and synchronization across various systems, further strengthening the integrity and availability of master data. By employing AI in MDM processes,

organizations can automate data management tasks, enhance data quality, and derive more value from their master data.

*Enhance data security and privacy measures:* Organizations must prioritize data security and privacy by implementing robust security controls, access management mechanisms, and encryption techniques. Compliance with data protection regulations, such as the General Data Protection Regulation (GDPR) and industry-specific requirements, should be ensured. Regular security audits, vulnerability assessments, and staff training programs are also recommended to maintain a strong security posture.

*Utilize data analytics and business intelligence:* To obtain valuable insights from their data, organizations should invest in modern data analytics and business intelligence tools and methodologies. Using data visualization, data mining, and predictive analytics, organizations may make informed decisions, detect patterns, and uncover lucrative business opportunities. The application of machine learning and artificial intelligence technologies can improve data analytics capabilities even further.

*Develop a data-driven culture:* It is critical to develop a data-driven culture within the firm. This entails increasing data literacy and awareness among employees at all levels. Data literacy abilities, such as data interpretation, analysis, and decision-making, should be improved through training programs and workshops. Organizations may empower people to use data effectively in their daily work and create innovation by establishing a data-driven culture.

*Embrace emerging technologies:* To unleash new opportunities in data management and analysis, organizations need keep up with evolving technologies such as cloud computing, big data analytics, and the Internet of Things (IoT). Exploring the potential benefits of these technologies, such as scalability, cost-effectiveness, and real-time data processing capabilities, can give businesses a competitive advantage.

*Continuously evaluate and improve data management processes:* Regular evaluation and improvement of data management processes are crucial for maintaining data quality and adapting to evolving business needs. Organizations should establish an AI-based feedback loop and performance measurement mechanisms to monitor the effectiveness

of data management practices. This feedback can drive continuous improvement initiatives and ensure that data management processes remain aligned with organizational objectives.

*Encourage cross-functional teamwork and collaboration:* Effective data management necessitates collaboration and cooperation across departments and teams. Organizations should foster cross-functional collaboration and create avenues for successful communication and knowledge sharing. This will make it easier to share data-related insights, best practices, and lessons learned, resulting in more efficient and effective data management processes.

## **5.2 Scientific implications**

This study presents a new and comprehensive knowledge of enterprise data management strategies by blending artificial intelligence and literature viewpoints, filling a large vacuum in existing literature. This study's holistic approach provides a detailed view of the complicated business data management landscape.

The conclusions of this thesis emphasize the importance of strong data management techniques in today's data-driven corporate environment. They emphasize the importance of strong data governance, good data quality, and effective data integration for business efficiency and effectiveness. This has theoretical implications for information systems, data science, and business management, highlighting the importance of data management in these fields.

Furthermore, the incorporation of AI tools in the research process, such as ChatGPT, reveals a fresh methodological approach for data collecting and analysis. This marks a novel shift in research approach that may stimulate future research into the potential of AI in qualitative research. Furthermore, this research has broader implications for AI, machine learning, and data analytics. The research contributes to the expanding conversation on the role of AI in business and decision-making processes by demonstrating how AI and advanced analytics may improve enterprise data management.

The thesis also engages with timely issues around data security and privacy, contributing to the ongoing scientific discourse in the field of cybersecurity and data protection. The exploration of the interplay between data management practices and compliance with data protection regulations provides valuable insights that advance our understanding of this complex relationship.

The study's findings could help shape the creation of new data management frameworks and tools, contributing to developments in business data management technologies. The study establishes a solid foundation for future advancements in the sector by identifying the critical variables for good data management.

Based on these findings, it is expected that this thesis will motivate additional study on enterprise data management, particularly the role of AI and the significance of continual development in response to the continuously expanding technological landscape. This will help to push the boundaries of present scientific understanding and shape the future trajectory of research in this topic.

### **5.3 Managerial Implications for Business Efficiency and Effectiveness**

The findings of this thesis offer valuable insights into good enterprise data management practices from both literature and artificial intelligence perspectives. These insights have significant managerial implications for organizations striving to improve their business efficiency and effectiveness. Organizations may ensure that their decision-makers have access to accurate, timely, and trustworthy information by implementing strong data management processes. This enables them to make informed decisions that match with the strategic goals of the firm, resulting in better resource allocation and overall performance. Improved data quality and accessibility also facilitate enhanced communication and collaboration among various departments, fostering a more streamlined and efficient organizational structure.

Moreover, implementing the recommended data management practices helps businesses optimize their operations and processes. This can lead to reduced operational costs, minimized data redundancies, and increased responsiveness to market trends and

customer needs. In turn, these improvements can result in higher customer satisfaction, increased market share, and ultimately, a more competitive position in the marketplace.

Another managerial implication of this thesis is the potential for improved risk management and regulatory compliance. Organizations can reduce the possibility of data breaches, fines, and reputational harm by implementing a solid data governance framework and complying with data protection rules and industry standards. This fosters trust among stakeholders and customers, leading to stronger relationships and long-term loyalty.

Furthermore, by utilizing advanced analytics tools and techniques, firms can reveal hidden patterns and insights from their data, thereby fostering innovation and identifying new chances for growth. These insights can assist firms in making market strategic moves, differentiating themselves from competitors, and remaining adaptive in the face of changing market conditions.

This thesis highlights the importance of continuous improvement and adaptability in data management practices. By regularly evaluating and updating their strategies, organizations can stay ahead of emerging trends, technologies, and industry best practices. This proactive approach ensures that businesses can maintain their competitive edge and continue to drive efficiency and effectiveness in the long term.

## 6 CONCLUSIONS

### 6.1 Key findings

Focusing on a combination of literature review and AI-reviewed practices, this study has emerged with several key findings. These findings shed light on the critical aspects of EDM, the role of AI in enhancing data management processes. Therefore, by embracing these findings, organizations can optimize their EDM strategies, drive business efficiency and effectiveness, and leverage the transformative potential of AI in data management processes. The key findings include:

*Importance of Data Governance:* The thesis highlights the significance of establishing a robust data governance framework. It emphasizes that organizations should define clear policies, standards, and procedures for managing data throughout its lifecycle. Effective data governance ensures data consistency, accuracy, security, and collaboration among stakeholders.

*Data Quality Management as a Cornerstone:* The findings underscore the importance of data quality management practices in EDM. Implementing processes such as data profiling, cleansing, validation, and enrichment is crucial to ensure data accuracy, completeness, and consistency. Continuous monitoring and improvement of data quality are essential for informed decision-making and improved operational efficiency.

*Use of AI for Data Management:* The thesis identifies the potential of AI in revolutionizing data management practices. Machine learning and natural language processing are two AI tools and approaches that may automate manual processes, spot patterns, and provide valuable insights from enormous datasets. AI-powered approaches improve data integration, data security, data analytics, and decision-making.

*Role of Data Architecture and Integration:* The research emphasizes the importance of a well-designed data architecture and effective data integration strategies. A robust data architecture accommodates diverse data sources and types, enables seamless data flow, and supports efficient data storage, retrieval, and processing. Integration strategies, such as ETL processes and APIs, facilitate data interoperability and accessibility.

*Data Security and Privacy Considerations:* The thesis underscores the criticality of data security and privacy in EDM. Organizations must implement comprehensive measures to protect sensitive information from unauthorized access and ensure compliance with data protection regulations. Encryption, access controls, data masking, and privacy-aware practices are essential components of a robust data security framework.

*Business Efficiency and Effectiveness Outcomes:* The conclusions of this study emphasize the importance of good EDM in increasing corporate efficiency and effectiveness. Well-managed data assets, accurate data, and improved data accessibility all contribute to better decision-making, streamlined operations, cost savings, increased customer satisfaction, and a competitive advantage in the market.

*Need for Continuous Improvement:* The thesis emphasizes the importance of continuous improvement and adaptability in data management practices. Organizations must stay abreast of emerging technologies, industry trends, and evolving regulatory landscapes. Regular evaluation and updating of data management strategies ensure that they remain effective and aligned with changing business needs.

## **6.2 Validity and reliability of the study**

This study explores the crucial intersection of data management practices and artificial intelligence (AI) in strengthening business efficiency and effectiveness. It aims to provide a comprehensive understanding of optimal data management practices, identify key challenges, and propose AI-based solutions. With the rapid growth of digital technology and the increasing complexity of data, organizations need an integrated approach that leverages artificial intelligence to fully exploit data potential. This research fills a gap in the existing literature and addresses the evolving nature of data management and the potential of AI where the provided perspectives can be applied to real business scenarios. The results and recommendations of the study would inspire further research in the fields of IT systems, data science, and business management. The credibility and acceptance of any of such qualitative research, are reinforced by a thorough evaluation of its validity and reliability. Bryman & Bell (2003) propose that the validity and reliability of qualitative research can be assessed through the following four observational lenses:

1. Trustworthiness of the achieved results
2. Validity of the results in different environments
3. Repeatability of the observations
4. Impact of the researcher's experience and overall value of the results

The trustworthiness of the achieved results can be demonstrated by the integration of two perspectives such as literature-based practices and AI-reviewed practices. The use of multiple sources of the related articles enhances the credibility of the findings and provides a comprehensive understanding of effective EDM practices. The integration of AI perspectives adds a simultaneous dimension to the research, expanding the scope of insights and enriching the findings. The rigorous comparison and analysis contribute to the trustworthiness of the achieved results.

When assessing the validity of the results in different environments, this criteria can be addressed by considering the perspectives of literature and AI. By comparing these two viewpoints, the study acknowledges the potential variations and nuances in data management practices across different environments. The literature-based practices provide a foundation based on established principles and best practices, while the AI-reviewed practices leverage advanced technologies to enhance data management. This consideration of different perspectives enhances the validity of the research results in diverse environments.

The repeatability of the observations of this study can be assessed based on the clarity and comprehensiveness of the research methodology. The systematic and comprehensive approach has been ensured in this study involving a literature review and the use of AI tools like ChatGPT. The replication and repetition of this study to other researchers have been committed by the detailed explanation of the research process. Moreover, the methodology includes well-documented data collection and analysis techniques, and also enhances the repeatability of the observations.

There might be challenges in achieving objectiveness of the qualitative research work respective to the Researcher's interests, experience, and competencies (Yin, 2003). Thus, it impacts on the overall value of the qualitative research. However, a thorough



understanding has been displayed regarding the research domain through the integration of literature and AI perspectives. Appropriate steps have been taken to mitigate personal biases to ensure the objectivity of the research that can be proved by the comprehensive literature review and the use of AI tools. The overall value of the results is evident in the practical implications and recommendations provided for organizations aiming to enhance their data management strategies and drive business efficiency and effectiveness.

### **6.3 Limitations and future research**

The primary limitation of this thesis lies in its limited scope, as it focuses on comparing good enterprise data management practices from literature and artificial intelligence perspectives. As a result, the study may not comprehensively cover all aspects of data management or address the unique challenges faced by different industries. Moreover, while the research attempts to provide general recommendations, organizations with specific needs or contexts may require more tailored solutions.

Another limitation is the rapidly evolving nature of the technology landscape. New advancements are continuously emerging, and some of the findings and recommendations presented in this thesis may become outdated or less relevant in the future. The dynamic nature of the field poses a challenge for researchers aiming to provide long-lasting and universally applicable solutions.

Subjectivity in the comparison between literature and AI perspectives may also be a limitation. Author's personal biases could influence the analysis which potentially can affect the accuracy and impartiality of the results. Additionally, access to sufficient case study could be a limiting factor, especially when obtaining real-world examples from enterprises to study their data management practices.

Generalizability of the results and recommendations may be another limitation, as data management practices can be highly dependent on the specific context, industry, and organizational structure. Therefore, the findings may not be applicable to all organizations.

It has been observed that the responses generated by ChatGPT may exhibit slight variances, irrespective of the geographical location associated with the IP addresses of the devices, or the personal background information that ChatGPT tracks for the user signed into the account. This was observed by both the author of this study and the research supervisor, who, despite posing identical questions to ChatGPT, received slightly differing responses. Nevertheless, the key content appeared to remain very much the same.

Future research could explore the unique data management challenges and best practices within specific industries, such as manufacturing, healthcare, finance, or service industry. This would help to address the limitations of this thesis in terms of scope and applicability to different contexts.

A thorough case study of an organization that has effectively implemented data management best practices could offer insightful information about the variables that contributed to their success and the difficulties they encountered. This strategy might also reveal other best practices and tactics that the current research might not have explored.

Longitudinal studies assessing the long-term effects of implementing the suggested best practices on business efficiency and effectiveness could provide valuable insights into the sustainability and impact of these practices. Such studies would help understand how well the recommendations withstand the test of time and adapt to evolving business needs.

As new technologies and methods continue to emerge, future research could focus on their potential impact on enterprise data management practices. This would ensure that the recommendations remain up-to-date and relevant in the ever-changing technology landscape.

Investigating the influence of data protection and privacy regulations on data management practices and how organizations can adapt to changing regulatory landscapes could be an interesting research direction. This would help organizations to discover complex legal requirements while maintaining efficient data management practices.

Future research could look into how artificial intelligence can be used to automate and improve many areas of organizational data management, such as data quality management, metadata management, and data analytics. This would help organizations leverage the power of AI to optimize their data management practices and drive business efficiency and effectiveness.

## REFERENCES

- Abraham R., Schneider J., Brocke J. V., (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49, p.424-438.
- Abu-Elkheir M., Hayajneh M., Abu Ali N., (2013). Data Management for the Internet of Things: Design Primitives and Solution. *Sensors*, ISSN 1424-8220, 13, p.15582-15612.
- Ahola S., (2022). Developing a customer master data management model. Master's thesis, Tampere University., Finland, November 2022.
- Al-Debei, M.M. (2011). Data Warehouse as a Backbone for Business Intelligence: Issues and Challenges. *Eur. J. Econ. Financ. Adm. Sci.*, 33, p.153–166.
- Alhassan, I., Sammon, D., & Daly, M. (2018). Data governance activities: a comparison between scientific and practice-oriented literature, *Journal of Enterprise Information Management*, 31(2), p.300-316.
- Allcott, H., Diamond, R., Dube, J.-P., Handbury, J., Rahkovsky, I., Schnell, M., (2019). Food deserts and the causes of nutritional inequality. *Q. J. Econ.* 134(4), p.1793-1844.
- Allen, M. & Cervo, D. (2015). *Multi-Domain Master Data Management. Advanced MDM and Data Governance*, Elsevier, Waltham, MA.
- Al-Mashari, M., & Zairi, M. (2000). The effective application of SAP R/3: A proposed model of best practice. *Logistics Information Management*, 13(3), p.156-166.
- Arlbjørn, J.S., Wong, C.Y. and Seerup, S. (2007) 'Achieving competitiveness through supply chain integration', *International Journal of Integrated Supply Management*, 3(1), p.4–24.
- Arnold V., Detterming H., Engel T., Karcher A. (2011). *Product Lifecycle Management beherrschen – Ein Anwendungsbuch für den Mittelstand*. Springer Verlag, Berlin Heidelberg.
- Ashdown, J., Eryurek, E., Gilad, U., Kibunguchy-Grant, A. & Lakshmanan, V. (2021). *Data Governance: The Definitive Guide*. O'Reilly Media, Inc.
- Awwad M., Kulkarni P., Bapna R., and Marathe A., (2018). Big Data Analytics in Supply Chain: A Literature Review. *Proceedings of the International Conference on Industrial Engineering and Operations Management*. p.418-425.

- Azad N., Ahmadi F. (2015). The customer relationship management process: Its measurement and impact on performance. *Journal of Uncertain Supply Chain Management*, 3(1), p.43-50.
- Bertino E., Ferrari E. (2018). Big Data Security and Privacy. Part of the Studies in Big Data book series (SBD, volume 31). p.425-439.
- Bidgoli, H. (2011). Successful Introduction of Cloud Computing into your Organization: A Six-Step Conceptual Model, *Journal of International Technology and Information Management*. 20(1-2), p.20-38.
- Borgia E. (2014). The internet of things vision: key features, applications and open issues. *Comp Comm.*, 54, p.1–31.
- Bousdekis A., Lepenioti K., Apostolou D., Mentzas G., (2021). A Review of Data-Driven Decision-Making Methods for Industry 4.0 Maintenance Applications. *Electronics*, 10, p.1-20.
- Brous, P., Herder, P., & Janssen, M. (2016). Governing asset management data infrastructures. *Procedia Computer Science*, 95, p.303–310.
- Brous, P., Janssen, M. & Vilminko-Heikkinen, R. (2016). Coordinating decision-making in data management activities: A systematic review of data governance principles. *Proceedings of the Electronic Government and Electronic Participation*, 2016.
- Brous, P., Janssen, M., & Vilminko-Heikkinen, R. (2016). Coordinating Decision-Making in Data Management Activities: A Systematic Review of Data Governance Principles, 5th International Conference on Electronic Government and the Information Systems Perspective (EGOV), Sep 2016, Porto, Portugal. in *Lecture Notes in Computer Science*, Vol.9820, Springer International Publishing, p.115-125.
- Bryman, A., & Bell, E. (2003). *Business research methods*. Oxford, UK. Oxford University Press.
- Brynjolfsson E., Hitt L., Kim H., (2011). Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? SSRN.
- Bughin, J., Chui, M., Manyika, J. (2010). Clouds, big data, and smart assets: ten tech-enabled business trends to watch. *McKinsey Q.* 56(1), p.75–86.
- Chaki, S. (2015). *The expert's voice in information management*. Enterprise Information Management in Practice, Springer, New York, NY.

- Chakravorty, R. (2020). Common challenges of data governance. *Journal of Securities Operations & Custody*, 13(1), p.23–43.
- Chalupa S., (2017). Transaction Data Based Hotel Customer Segmentation. *Proceedings of the 10th International scientific conference: Karvina Ph.D. conference on business and economics*. p.74-82.
- Chen, E.T. (2012). Implementation issues of enterprise data warehousing and business intelligence in the healthcare industry. *Commun. IIMA*, 12, 3.
- Chen, H. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), p.1165-1188.
- Chen, M., Han, J., & Yu, P.S. (2000). Data mining: An overview from a database perspective. *IEEE Transactions on Knowledge and Data Engineering*, 8(6), p.866-883.
- Choi T., Luo S., (2019). Data quality challenges for sustainable fashion supply chain operations in emerging markets: Roles of blockchain, government sponsors and environment taxes. *Transportation Research. Part E*, p.139-152.
- Cleven, A., & Wortmann, F. (2010). Uncovering four strategies to approach master data management. *Proceedings of the 43rd Hawaii International Conference on System Sciences*.
- Coltman, T. (2007). Why build a customer relationship management capability? *Journal of Strategic Information Systems*, 16(3), p.301-320.
- Cuzzocrea, A., Bellatreche, L., Song, I.Y. (2013). Data warehousing and OLAP over big data: Current challenges and future research directions. In *Proceedings of the Sixteenth International Workshop on Data Warehousing and OLAP, DOLAP '13*, San Francisco, CA, USA, 28 October 2013; Association for Computing Machinery: New York, NY, USA, p.67–70.
- Dahlberg, T., Heikkilä, J., & Heikkilä, M. (2011). Framework and Research Agenda for Master Data Management in Distributed Environments. In T. Leino (Ed.), *Proceedings of IRIS 2011*. TUCS Lecture Notes No. 15, October 2011, p.82-90.
- Dallas I. & Wynn M. T., (2014). Business Process Management in Small Business: A Case Study. *Information Systems for Small and Medium-sized Enterprises*, p.25-46.
- Das T. K., Mishra M. R., (2011). A study on Challenges and Opportunities in Master Data Management. *International Journal of Database Management Systems*. 3(2). p.129-139.

- Demchenko, Y., Ngo, C., de Laat, C., Membrey, P., Gordijenko, D. (2014). Big Security for Big Data: Addressing Security Challenges for the Big Data Infrastructure. In *Secure Data Management*; Jonker, W., Petkovi'c, M., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, p.76–94.
- Doebeli, G., Fisher, R., Gapp, R., & Sanzogni, L. (2011). Using BPM governance to align systems and practice. *Business process management Journal* 17(2), p.184–202.
- Dreibelbis, A., Hechler, E., Milman, I., Oberhofer, M., van Run, P., & Wolfson, D. (2008). *Enterprise master data management: An SOA approach to managing core information*. Westford, Massachusetts: IBM Press.
- Dyché, J., & Levy, E. (2006). *Customer data integration: Reaching a single version of the truth*. John Wiley & Sons.
- Eigner M., Stelzer R. (2009). *Product Lifecycle Management – Ein Leitfaden für Product Development und Life Cycle Management*. Springer Verlag, Berlin.
- Elfarmawi W., (2019). Correlation Between Customer Relationship Management System Usage, Product Innovation, and Customer Satisfaction. *Foundations of Management*, 11, p.23-32.
- ElMaraghy E., Schuh G., ElMaraghy W., Piller F., Schonsleben P., Tseng M., Bernard A., (2013). Product Variety Management. *CIRP Annals-Manufacturing Technology*, 62, p.629-652.
- Etwaroo, S.R. (2019). *Data Quality Impact on Data Frameworks within Business Organizations*. ProQuest Dissertations Publishing
- Eynard B., Gallet T., Nowak P., Roucoules L., (2004). UML based specifications of PDM product structure and workflow. *Computers in industry*, 55(3), p.301-16.
- Farshad F., Bahar F., Mojtaba B., Mahmoud D., (2020). AI-Driven Data Monetization: The other Face of Data in IoT-based Smart and Connected Health. *IEEE Xplore*, 9(8), p.5581-5599.
- Feldhusen J., Grote K.-H., (2013). *Pahl/Beitz Konstruktionslehre – Methoden und Anwendungen erfolgreicher Produktentwicklung*. Springer Verlag, Berlin Heidelberg.
- Flamholtz E., (1995), “Managing organizational transitions: Implications for corporate and human resource management,” *European Management Journal*, 13(1), p.39–51.

- Gartner. (2022). Master Data Management (MDM), Information Technology, Gartner Glossary, Gartner website. Available at: <https://www.gartner.com/en/information-technology/glossary/master-data-management-mdm> (Accessed: 5 July 2022).
- Genaroro, A. F. S., & Lourero, G. (2015). Stakeholder management as an approach to integrated management system, transdisciplinary lifecycle analysis of systems. The 22nd International Conference on Concurrent Engineering. ISPE Inc.
- Ghinita G., Kalnis P., and Tao Y. (2011). Anonymous Publication of Sensitive Transactional Data. *IEEE Transactions on Knowledge and Data Engineering*, 23(2), p.161-174.
- Ghods, A., Al-Fuqaha, A., & Guizani, M. (2017). Toward better horizontal integration among IoT services. *IEEE Communications Magazine*, 55(2), p.72-79.
- Giebler C., Gröger C., Hoos E., Schwarz H., and Mitschang B., (2019). Leveraging the Data Lake: Current State and Challenges. *Proc. of the 21st International Conference on Big Data Analytics and Knowledge Discovery (DaWaK)*, Springer, Cham.
- Gomez, J.M., Bedolla, J.S., Ricci, F. and Chiabert, P. (2014) 'Validation process model for product life-cycle management', *International Journal of Product Lifecycle Management*, 7(2/3), p.230–246.
- González, I., Calderón, A.J., Figueiredo, J., Sousa, J. (2019). A literature survey on open platform communications (OPC) applied to advanced industrial environments. *Electronics*, 8, 510.
- Gu N. and London K. (2010). Understanding and facilitating BIM adoption in the AEC industry. *Automation in construction*, 19(8), p.988-999.
- Hagmann, J. (2013). Information governance – Beyond the buzz. *Records Management Journal*, 23(3), p.228–240.
- Haines, S. (2014). *The product manager's desk reference*, second edition (2nd ed. ed.). New York: McGraw-Hill Education.
- Hakim Y. R. A., (2022). Product Variance and Brand Image Determining Customer Loyalty: Purchase Behavior of Sosro Tea Products. *Studi Ilmu Sosial Indonesia*, 2(1), p.13-24.



- Hammer, M. (2010). What is business process management? In J. vom Brocke & M. Rosemann (Eds.), *Handbook on business process management 1. International handbooks on information systems* (p.3–16). Berlin Heidelberg: Springer.
- Hannila H., (2019). *Towards Data-Driven Decision making in Product Portfolio Management- From company-level to product-level analysis*, ISBN 978-952-62-2442-8 (PDF), University of Oulu.
- Harding, M., Leibtag, E., Lovenheim, M.F., (2012). The heterogeneous geographic and socioeconomic incidence of cigarette taxes: evidence from nielsen homescan data. *Am. Econ. J. Econ. Pol.* 4(4), p.169-198.
- Harkonen J., Haapasalo H., and Hamminen K., (2015), “Productisation: A review and research agenda,” *International Journal Production Economics*, 164, p.65–82.
- Harkonen J., Mustonen E., Haapasalo H., (2019),” *Construction-Related Data Management: Classification and Description of Data from Different Perspectives*”, *International Journal of Management, Knowledge and Learning*, 8(2), p.195-220.
- Harkonen J., Tolonen A., and Haapasalo H., (2017), “Service productisation: Systematising and defining an offering,” *Journal of Service Management*, 28(5), p.936–971.
- Harkonen, J., Tolonen, A., & Haapasalo, H. (2018a). *Modelling of manufacturing services and processes for effective productization*, the 20th International Working Seminar on Production Economics, Innsbruck, Austria, February 19-23, 2018.
- Harkonen, J., Tolonen, A., & Haapasalo, H. (2018b). *Modelling of Construction Products and Services for Effective Productisation*. MakeLearn & TIIM 2018 conference, 16-18 May 2018, Naples, Italy.
- Harkonen, J., Tolonen, A., & Haapasalo, H. (2018c). *Modelling of Construction Products and Services for Effective Productisation*. *Management*, 13(4), p.335-353.
- Harmon, P. (2014). *Business process change: A guide for business managers and BPM and Six Sigma professionals*. Morgan Kaufmann.
- Haug, A. & Arlbjørn, J.S. (2011). Barriers to Master Data Quality. *Journal of Enterprise Information Management*, 24(3), p.288–303.

- Herman, L. E., Sulhaini, S., & Farida, N. (2020). Electronic customer relationship management and company performance: Exploring the product innovativeness development. *Journal of Relationship Marketing*, p.1–19.
- Herzog, C., Handke, C. Hitters E. (2019). *Analyzing talk and text: thematic analysis*. The Palgrave Handbook of Methods for Media Policy Research, Palgrave Macmillan, pp. 385-401
- Igarashi, Y., Nagata, K., Kuwatani, T., Omori, T., Nakanishi-Ohno, Y., & Okada, M. (2016). Three levels of data-driven science. *Journal of Physics: Conference Series*, 699, 1.
- Irani Z., Themistocleous M., Love P.E. (2003). The impact of enterprise application integration on information system lifecycles. *Information & Management*, 41(2), p.177-187.
- Jain, V. (2016) ‘Conceptualising ERP systems’ value as a multidimensional formative construct’, *International Journal of Business Information Systems*, 21(4), p.439–461.
- Janardanan, V. K., Adithan, M., & Radhakrishnan, P. (2008). Collaborative product structure management for assembly modeling. *Computers in Industry*, 59(8), p.820-832.
- Jaravel, X., (2019). The unequal gains from product innovations: evidence from the U.S. Retail sector. *Q. J. Econ.* 134(2), p.715-783.
- Jeston, J., & Nelis, J. (2008). *Management by process: A roadmap to sustainable business process management*. Routledge.
- Kadadi, A., Agrawal, R., Nyamful, C., Atiq, R. (2014). Challenges of data integration and interoperability in big data. In *Proceedings of the 2014 IEEE International Conference on Big Data (Big Data)*, Washington, DC, USA, 27–30 October 2014; p.38–40.
- Kemppainen, T. (2010) *Product Data Ownership Network*, Master’s Thesis, Department of Process Engineering, University of Oulu.
- Khatri, V. (2016). Managerial work in the realm of the digital universe: The role of the data triad. *Business Horizons*, 59, p.673–688.
- Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, 53(1), p.148–152.

- Khayru, R.K. & F. Issalillah. (2021). Study on Consumer Behavior and Purchase of Herbal Medicine Based on The Marketing Mix, *Journal of Marketing and Business Research*, 1(1), p.1-14.
- Knolmayer, G. and Rothlin, M. (2006), "Quality of material master data and its effect on the usefulness of distributed ERP systems", *Lecture Notes in Computer Science*, 4231, p.362-371.
- Kropsu-Vehkapera H., and Haapasalo H., (2011), "Defining product data views for different stakeholders", *Journal of Computer Information Systems*, 52(2), p.61-72.
- Kropsu-Vehkapera, H. (2012) *Enhancing Understanding of Company-Wide Product Data Management in ICT Companies*, Doctoral Dissertation, *Acta Universitatis Ouluensis C Technica* 418, University of Oulu, Finland.
- Kropsu-Vehkapera, H. and Haapasalo, H. (2011) 'Defining product data views for different stakeholders', *Journal of Computer Information Systems*, 52(2), p.61-72.
- Kropsu-Vehkapera, H., Haapasalo, H., Jaaskelainen, O., & Phusavat, K. (2011b). *Product Configuration Management in ICT Companies: The Practitioners' Perspective*. *Technology and Investment*, 2(4), p.273-285.
- Kumar Das T, Mishra MR. (2011). A study on challenges and opportunities in master data management. *Int J Database Manage Syst.*, 3(2), p.129-39.
- Kwarteng M. A., Pilik M., Jurickova E., (2017). Mining Interest In Online Shoppers' Data: An Association Rule Mining Approach. *Acta Polytechnica Hungarica*, 14(7), p.143-160.
- Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics, *International Journal of Information Management*. 34, p.387-394.
- Lachmayer R., Mozgova I., Gottwald P., (2015). Formulation of Paradigm of Technical Inheritance. In *Proceeding of the 20th International Conference on Engineering Design (ICED15)*, 27.-30.07.2015, Milan, Italy.
- Lahtinen N.; Mustonen E.; Harkonen J., (2021). Commercial and Technical Productization for Fact-Based Product Portfolio Management Over Lifecycle. *IEEE Transactions on Engineering Management*, 68(6), p.1826-1838.
- Lake, P., Crowther, P. (2013). *Concise Guide to Databases*. Springer, London.

- Lee, Y., Madnick, S., Wang, R., Wang, F., & Zhang, H. (2014). A cubic framework for the chief data officer: Succeeding in a world of big data. *MIS Quarterly Executive*, 13(1), p.1–13.
- Lestari, U. P. & A. R. Putra. (2022). Brand Switching Behavior on Smartphone Product Purchases. *Journal of Science, Technology and Society*, 3(2), p.23–31.
- Lichtenthaler U., (2022). Data management efficiency: major opportunities for shared value innovation. *Management Research Review*, 45(2), p.156-172.
- Lindner A., (2015). *Methodik zur Verbesserung standardisierter Produkte auf der Basis der Auswertung von Nutzungsdaten deren Vorgängergenerationen*. Shaker Verlag, Bochum.
- Lindstrom, M. (2020), “Don’t let digital transformation make you less human”, *Harvard Business Review Digital Articles*, available at: <https://hbr.org/2020/05/dont-let-digital-transformationmake-you-less-human>
- Lucas, A. (2010). Corporate quality management. *Proceedings of the Iberian Conference on Information Systems and Technologies*. p.524–548.
- Lucas-Estañ, M.C., Sepulcre, M., Raptis, T.P., Passarella, A., Conti, M. (2018). Emerging trends in hybrid wireless communication and data management for the industry 4.0. *Electronics*, 7, 400.
- Lund, B.D., & Wang, T. (2023). Chatting about ChatGPT: how may AI and GPT impact academia and libraries? *Library Hi Tech News*. DOI: <https://doi.org/10.1108/LHTN-01-2023-0009>
- Macedoni, L., Xu, M., Feenstra, R.C., (2020). *Large Firms in Retail Markets: Multiple Products for Heterogeneous Consumers*. Working paper.
- Madera C., (2020). *Master Data and Reference Data in Data Lake Ecosystems*. Data Lakes, First Edition. Published by ISTE Ltd and John Wiley & Sons, Inc.
- Mansar, S. L., & Reijers, H. A. (2005). Best practices in business process redesign: an overview and qualitative evaluation of successful redesign heuristics. *Omega*, 33(4), p.283-306.
- Martijn N., Hulstijn J., Bruijne de M., Tan Y., (2015). Determining the Effects of Data Governance on the Performance and Compliance of Enterprises in the Logistics and Retail Sector. *International Federation for Information Processing*, p.454-466.

- McGilvray, D. (2006). Data governance: A necessity in an integrated information world, *DM Review*. December, 1–3.
- Meyliana, Hidayanto, A.N. and Budiardjo, E.K. (2016) ‘The critical success factors for customer relationship management implementation: a systematic literature review’, *International Journal of Business Information Systems*, 22(2), p.131–174.
- Mosley, M. & Brackett, M. (2010). *The DAMA guide to the data management body of knowledge (DAMA-DMBOK guide)*, first edition. Bradley Beach, N.J: Technics Publications.
- Moss, L. (2007). Critical success factors for master data management. *Cutter IT Journal*, 20(9), p.7-12.
- Mustonen, E., Harkonen, J., & Haapasalo, H. (2019a). From Product to Service Business: Productization of Product-oriented, Use-oriented, and Result-oriented Business. *Proceedings of 2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM2019)*, 15-18 Dec 2019, Macau.
- Nadeem U. Shahid, Nasir J. Sheikh, (2021). Impact of Big Data on Innovation, Competitive Advantage, Productivity, and Decision Making: Literature Review, *Journal of Business and Management*, 9, p.586-617.
- Nambiar A., & Mundra D., (2022). An Overview of Data Warehouse and Data Lake in Modern Enterprise Data Management. *Big data and cognitive computing*, 6(132), p.1-24.
- Ofner, M.H., Straub, K., Otto, B. and Oesterle, H. (2013) ‘Management of the master data lifecycle: a framework for analysis’, *Journal of Enterprise Information Management*, 26(4), p.472–491.
- Oliveira, T., Thomas, M., & Espadanal, M. (2014). Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors. *Information & Management*, 51(5), p.497-510.
- OpenAI (2023). ChatGPT. <https://chat.openai.com/chat>. Accessed June 2023.
- Otto, B. (2012). Managing the business benefits of product data management: The case of Festo. *Journal of Enterprise Information Management*, 25(3), p.272–297.
- Otto, B. (2013). On the evolution of data governance in firms: The case of Johnson & Johnson consumer products North America. In S. S (Ed.). *Handbook of data quality* (p.93–118). Berlin, Heidelberg: Springer.

- Otto, B., Lee, Y.W. and Caballero, I. (2011) 'Information and data quality in networked business', *Electronic Markets*, 21(2), p.79–81.
- Papazoglou, M.P., & van den Heuvel, W.J. (2006). Service-oriented design and development methodology. *International Journal of Web Engineering and Technology*, 2(4), p.412-442.
- Payne, A., & Frow, P. (2005). A strategic framework for customer relationship management. *Journal of Marketing*, 69(4), p.167-176.
- Petnik J., Lhotska L., Dolezal J., and Adolf J., (2019). Behavioural Data Modeling: A Case Study in IoT. *Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2019)*, p.259-264
- Porter ME, Heppelmann JE. (2014). How smart, connected products are transforming competition. *Harv Bus Rev.*, 92, p.64–88.
- Prayitno E., & Astuty NA., (2017). Positive Impact of Customer Relationship Management (CRM) Implementaion to Improving the Services of Animal Polyclinics Customers. *International Conference on Sustainable Information Engineering and Technology (SIET)*. p.246-250.
- Qin, C., Zhang, A., Zhang, Z., Chen, J., Yasunaga, M. & Yang, D. (2023). Is ChatGPT a general-purpose natural language processing task solver? *arXiv preprint arXiv:2302.06476*.
- Ram, J., Corkindale, D., & Wu, M.L. (2014). ERP adoption and the value creation: Examining the contributions of antecedents. *Journal of Engineering and Technology Management*, 33, p.113-133.
- Ranganathan, C., & Brown, C.V. (2006). ERP investments and the market value of firms: Toward an understanding of influential ERP project variables. *Information Systems Research*, 17(2), p.145-161.
- Redman, T. C. (2008). *Data driven: profiting from your most important business asset*. Harvard Business Press.
- Retnowati, E., U. P. Lestari, R. Mardikaningsih, E. A. Sinambela, D. Darmawan, A. R Putra, & S. Arifin., (2021). The Effect of Packaging, Product Variance, and Brand Equity on Nutella Consumer Trust. *Asian Journal of Management, Entrepreneurship and Social Science*, 1(1), p.169-180.

- Rosemann, M., & vom Brocke, J. (2015). The six core elements of business process management. In *Handbook on business process management*, 1, p.105-122.
- Rouibah, K. and Ould-Ali, S. (2007). Dynamic data sharing and security in a collaborative product definition management system. *Robotics & Computer-Integrated Manufacturing*, 23(2), p.217-33.
- Sääksvuori, A, Immonen, A, (2008). *Product Lifecycle management*, Berlin, Springer, ISBN: 978-3-540-78173.
- Salvador F., (2007), "Toward a product system modularity construct: Literature review and reconceptualization," *IEEE Transactions on engineering management*, 54(2), p.219–240.
- Schaffer T, Leyh C. (2017). Master data quality in the era of digitalization – toward inter-organizational master data quality in value networks: a problem identification. In: Piazzolo F, Geist V, Brehm L, Schmidt R. (Eds.) *Innovations in Enterprise Information Systems Management and Engineering. ERP Future 2016. Lecture Notes in Bus. Inf. Proc.*, Hagenberg, Austria, 285, p.99–113.
- Scheidel W., Mozgova I., Lachmayer R., (2017). Structuring Information in Technical Inheritance by PDM Systems. In *Proceeding of ICED 17-21st International Conference on Engineering Design*, Vancouver, Canada, 21-25 August 2017.
- Shahin, A., Pool, J.K. and Poormostafa, M. (2014) 'Evaluating and ranking hotels offering e-service by integrated approach of Webqual and fuzzy AHP', *International Journal of Business Information Systems*, 15(1), p.84–104.
- Silvola R. (2018). *One Product Data for Integrated Business Processes*. ISBN 978-952-62-2114-4 (PDF), University of Oulu.
- Silvola R., Tolonen A., Harkonen J., Haapasalo H. and Mannisto T., (2019). Defining one product data for a product. *Business Information Systems*, 30(4), p.489-520.
- Silvola, J. (2018). Data management approach for multi-domain product data. *International Journal of Product Lifecycle Management*, 11(2), p.160-175.
- Silvola, R, Jaaskelainen, O, Kropsu-Vehkaperä, H, Haapasalo, H, 2011, Managing one master data – challenges and predictions, *Industrial Management & Data Systems*, 111(1), p.146-162.

- Silvola, R., Harkonen, J., Vilppola, O., Kropsu-Vehkapera, H. and Haapasalo, H. (2016) 'Data quality assessment and improvement', *International Journal of Business Information Systems*, 22(1), p.62–81.
- Silvola, R., Kauhanen, J., Collin, J., Haapasalo, H. & Kropsu-Vehkapera, H. (2011). A framework for improving the launch concept of new services. *Proceedings of TIIM2011 Conference*, 28–30 June, 2011 Oulu, Finland, p.692–706
- Silvola, R., Kemppainen, T., Haapasalo, H., Kropsu-Vehkapera, H. & Jaaskelainen, O. (2011). Elements and implementation of product data ownership network. *Proceedings of TIIM2011 Conference*, 28–30 June, 2011 Oulu, Finland, p.735–752.
- Silvola, R., Tolonen, A., Harkonen, J., Haapasalo, H. & Männistö, T. (In press). Defining one product data for a Product. *International Journal of Business Information Systems*.
- Simões, D.M. (2010). Enterprise Data Warehouses: A conceptual framework for a successful implementation. In *Proceedings of the Canadian Council for Small Business & Entrepreneurship Annual Conference*, Calgary, AL, Canada, 28–30 October 2010.
- Singh, R.; Singh, K. (2010). A Descriptive Classification of Causes of Data Quality Problems in Data Warehousing. *Int. J. Comput. Sci. Issues*, 7(41).
- Solomon, M. R. (2015). *Customer Behavior: Buying, Having and Being*, 11th Edition, Prentice-Hall, New Jersey.
- Sonzini, M.S., Vegetti, M. and Leone, H. (2015) 'Towards ontology for product version management', *International Journal of Product Lifecycle Management*, 8(1), p.90–97.
- Sriti, M.F., Assouroko, I., Ducellier, G., Boutunaud, P. and Eynard, B. (2015) 'Ontology-based approach for product information exchange', *International Journal of Product Lifecycle Management*, 8(1), p.1–23.
- Steinfeld, C., Markus, M.L., Wigand, R.T. (2011). Through a glass clearly: standards, architecture, and process transparency in global supply chains. *J. Manage. Inf. Syst.* 28(2), p.75–108.
- Suominen A., Kantola J., and Tuominen A., (2009). Reviewing and defining productization. In *Proc. 20th. Annual Conference of the International Society. Professional Innovation Management*, Vienna, Austria, p.63–85.
- Terzi S, Bouras A, Dutta D, Garetti M, Kiritsis D. (2010). Product lifecycle management – from its history to its new role. *Int J Prod Lifecycle Manag.*, 4(4), p.360–389.



Thorntwaite W. (2015). Dealing with data quality: don't just sit there, do something. In: Kimball R, Ross M, Becker B, Mundy J, Thorntwaite W. (Eds). *The Kimball group reader: relentlessly practical tools for data warehousing and business intelligence*. Second ed. John Wiley & Son. Indianapolis (IN); p.2009.

Thusoo, A. & Sarma, J. S. (2017). *Creating a Data-Driven Enterprise with DataOps. Insights from Facebook, Uber, LinkedIn, Twitter, and eBay*. O'Reilly Media, Inc., Sebastopol, CA. ISBN: 978-1-491-97781-1

Tian, G., & Quan, L. (2008). An improved framework of business process management system with integrating the strategy management. *International Conference on Management Science and Engineering, 15th Annual Conference Proceedings*. Long Beach, USA.

Tidstam, A. and Malmqvist, J. (2015) 'A systematic process for developing configuration rules', *International Journal of Product Lifecycle Management*, 8(1), p.46–64.

Tolonen A., Harkonen J., and Haapasalo H., (2014). Product portfolio management—governance for commercial and technical portfolios over life-cycle. *Technol. Investment*, 5(4), p.173–183.

Toshniwal, R., Dastidar, K.G., Nath, A. (2014). Big Data Security Issues and Challenges. *Int. J. Innov. Res. Adv. Eng.* 2, p.15–20.

Vajna S., Weber C., Bley H., Zeman K., (2009). *CAX für Ingenieure*. Springer Verlag, Berlin Heidelberg.

van der Aalst, W. M. (2013). *Business process management: A comprehensive survey*. *ISRN Software Engineering*, 2013, p.1-37.

Van Gils, B. (2020). *Data Management*. Zaltbommel: Van Haren Publishing.

VDI 2219., (2016). *Information Technology in product development*. Beuth Verlag, Berlin.

Vilminko-Heikkinen, R. & Pekkola, S. (2017). Master Data Management and its Organizational Implementation: An ethnographical study within the public sector. *Journal of Enterprise Information Management*, 30(3), p.454–475.

Waddington, D. (2008), Adoption of Data Governance by Business, *DM review*, 18(12), p.1- 5.

- Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165(C), p.234-246.
- Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, 12(4), p.5-33.
- Wirtz J., Fritze M. P., Jaakkola E., Gelbrich K., Hartley N. (2021). Service products and Productization. *Journal of Business Research*, 137, p.411-421.
- Wisner, J. D., & Stanley, L. L. (2008). *Process Management, Creating Value along the Supply Chain*. Mason: Thomson South Western.
- Wixom, B.H., Yen, B., & Relich, M. (2013). Maximizing value from business analytics. *MIS Quarterly Executive*, 12(2), p.111-123.
- Yang, X., Moore, P.R., Wong, C.-B., Pu, J.-S. and Chong, S.K. (2007). Product lifecycle information acquisition and management for consumer products. *Industrial Management & Data Systems*, 107(7), p.936-953.
- Yin, R. K. (2003). *Case Study Research: Design and Methods* (3rd ed). Sage Publications Inc, Beverly Hills, CA.
- Zhang Y., Ren S., Liu Y., Sakao T., Huisingh D., (2017). A framework for Big Data driven product lifecycle management. *Journal of Cleaner Production*, 159, p.229-240.
- Zhu K. (2002). Information Transparency in Electronic Marketplaces: Why Data Transparency May Hinder the Adoption of B2B Exchanges. *Electron. Mark.*, 12(2), p.92-99.

Questionnaire posed to ChatGPT to ascertain AI-reviewed EDM best practices.

1. What are the most critical aspects of an effective enterprise data management system?
2. What are the key benefits of adopting a well-rounded data management strategy for an enterprise?
3. How important is it to align business objectives with data strategy, and what are the best ways to do this?
4. How should organizations approach data literacy to ensure the efficient use of their data assets?
5. What are some best practices for ensuring data quality in an enterprise data management system?
6. What role does data governance play in an enterprise data management strategy, and how can it be implemented effectively?
7. How important is having a well-structured data architecture, and what are some strategies for achieving this?
8. How can an organization assess the maturity of its data management?
9. How can companies manage the challenges of data integration in the context of enterprise data management?
10. How can data security and privacy be ensured in enterprise data management?
11. How do emerging technologies impact the future of enterprise data management?
12. How can artificial intelligence and machine learning contribute to optimizing data management?
13. What role does data management play in supporting digital transformation?
14. How should organizations deal with the challenges of managing unstructured data?

## **1. What are the most critical aspects of an effective enterprise data management system?**

ChatGPT:-

Effective enterprise data management systems are crucial for businesses to operate efficiently and compete in today's data-driven world. Here are some of the most critical aspects of such systems:

1. **Data Governance:** A successful enterprise data management system needs clear and effective data governance policies to ensure data accuracy, consistency, and security. This includes clear definitions of who owns what data and how data should be used and protected.
2. **Data Quality:** It's critical that the data within the system is of high quality. This includes ensuring data is accurate, consistent, and up-to-date. Poor data quality can lead to inaccurate analysis and poor business decisions.
3. **Data Security:** With increasing concerns over data breaches and privacy issues, a robust data management system must have strong security protocols in place. This includes both physical and digital security measures.
4. **Data Integration:** The ability to integrate data from various sources is a critical feature. This includes data from internal sources, such as different departments or business units, as well as external sources. Integration is necessary for a complete and accurate view of the business.
5. **Data Accessibility:** Data should be easily accessible to those who need it to make decisions. This means the system should have a user-friendly interface and allow for easy data retrieval and analysis.
6. **Scalability:** As the business grows, the amount of data it generates will also grow. An effective data management system needs to be able to scale to accommodate this growth without sacrificing performance.
7. **Compliance:** Enterprises operate under many regulations, many of which have implications for data management. The system must be able to ensure and demonstrate compliance with these regulations, including data privacy laws.
8. **Metadata Management:** Metadata, data about data, is a crucial aspect of managing data. It can help users understand the source, usage, and relevance of data, thus improving the overall utility of the data assets.

9. **Data Architecture:** A well-structured data architecture is essential to ensure data is correctly organized, stored, and can be efficiently retrieved and processed. This should include data modeling and design that facilitate analysis and insights.
10. **Real-Time Processing:** In today's fast-paced business world, the ability to process and analyze data in real time is becoming increasingly important. This can provide businesses with a competitive advantage by allowing them to respond to changes and make decisions more quickly.

Remember, implementing an effective enterprise data management system is a significant undertaking that involves a range of factors beyond just the technical aspects. It also requires a cultural shift within the organization towards valuing, managing, and using data effectively.

## **2. What are the key benefits of adopting a well-rounded data management strategy for an enterprise?**

ChatGPT:-

A well-rounded data management strategy provides numerous benefits to an enterprise, including but not limited to the following:

1. **Improved Decision-Making:** By ensuring data is accurate, consistent, and easily accessible, a data management strategy enables better decision-making. It allows organizations to derive actionable insights from their data, leading to more informed and effective decisions.
2. **Enhanced Efficiency and Productivity:** Good data management can streamline business processes, reduce redundancy, and minimize errors. This can lead to significant improvements in operational efficiency and productivity.
3. **Regulatory Compliance:** A well-rounded data management strategy will consider regulatory requirements and ensure that data handling processes are compliant with relevant laws and regulations, such as GDPR or CCPA. This can prevent legal issues and potential fines.
4. **Improved Data Quality:** A data management strategy that includes data quality control can greatly improve the reliability of your data, reducing inaccuracies and inconsistencies.
5. **Increased Security:** A strong data management strategy includes robust data security measures, helping to protect sensitive data from breaches and cyber attacks.

6. **Cost Savings:** Better data management can lead to significant cost savings. By reducing errors, improving efficiency, and avoiding regulatory fines, organizations can save money. Additionally, identifying redundant data storage and improving data utilization can also lead to cost reduction.
7. **Better Customer Service:** Accurate and easily accessible data can enable better customer service by providing a complete, real-time view of customer interactions. This can lead to improved customer satisfaction and loyalty.
8. **Scalability:** A good data management strategy will plan for future growth, allowing your data infrastructure to scale smoothly as your business grows.
9. **Competitive Advantage:** In today's data-driven world, businesses that can effectively manage and utilize their data have a significant advantage over their competitors. This can lead to improved products, services, and business strategies.
10. **Data Literacy:** With a solid data management strategy, businesses can foster a data-driven culture and improve data literacy among employees, encouraging more data-informed decision making across all levels of the organization.

By combining these benefits, a well-rounded data management strategy can greatly enhance the overall success and profitability of an enterprise.

### **3. How important is it to align business objectives with data strategy, and what are the best ways to do this?**

ChatGPT:-

Aligning business objectives with a data strategy is crucial for several reasons:

1. **Ensures Relevance:** Aligning the data strategy with business objectives ensures that the data collected and analyzed is relevant to the needs of the business. This helps avoid wasting resources on irrelevant data.
2. **Guides Decision-Making:** A well-aligned data strategy can guide decision-making throughout the organization, providing insights that are directly relevant to the organization's goals.
3. **Fosters Efficiency:** When data strategies align with business objectives, the business can operate more efficiently. Decisions can be made faster and with more confidence, as they're based on solid, relevant data.
4. **Improves ROI:** Investments in data infrastructure, tools, and personnel can be significant. Alignment with business objectives helps ensure a positive return on these investments by focusing resources on areas that will generate the most value.

Here are some ways to align business objectives with a data strategy:

1. **Understand Business Objectives:** The first step is to have a clear understanding of the business's objectives. These could range from increasing revenue or market share to improving customer satisfaction or operational efficiency.
2. **Identify Relevant Data:** Once the objectives are understood, identify what data is most relevant to these goals. This includes not just what data to collect, but also how to analyze and use it.
3. **Collaborate Across Teams:** Involve stakeholders from across the organization in the development of the data strategy. This will help ensure the strategy meets the needs of different departments and functions and is aligned with overall business objectives.
4. **Set Data Goals:** Establish clear, measurable goals for the data strategy that support the business objectives. This could include targets for data collection, analysis, or usage.
5. **Regular Review and Adjustments:** Business objectives and data landscapes evolve over time. Regularly review and adjust the data strategy as needed to ensure it remains aligned with business goals.
6. **Promote a Data-Driven Culture:** Foster a culture that understands the value of data and uses it in decision-making. This will help ensure that data strategy is not just a side project, but a core part of how the business operates.

Aligning a data strategy with business objectives is not a one-time task, but an ongoing process that requires communication, collaboration, and a strong commitment to being data-driven.

#### **4. How should organizations approach data literacy to ensure the efficient use of their data assets?**

ChatGPT:-

Data literacy, the ability to read, understand, analyze and communicate with data, is becoming increasingly important as more organizations leverage data-driven decision making. Here are several strategies an organization can employ to improve data literacy:

1. **Training and Education:** Implement comprehensive training programs to educate employees on the importance of data, how to interpret it, and how to use it in their roles. This can be through workshops, online courses, or regular training sessions.

Ensure the training is appropriate for the various levels of data literacy in your organization.

2. **Promote a Data-Driven Culture:** Encourage a culture where decisions are made based on data, not just intuition. Make it clear that everyone has a role to play in using data effectively.
3. **Access to Data:** Provide employees with access to the data they need to make decisions. This includes not just raw data, but also data analysis tools and dashboards that make it easier to understand and use the data.
4. **Cross-Functional Collaboration:** Encourage collaboration between data experts and other teams. This can help non-data experts learn more about using data effectively, and can help data teams understand the needs of the rest of the organization.
5. **Mentorship and Support:** Create a system where employees who are more data literate can support those who are less experienced. This can be through formal mentorship programs or more informal support networks.
6. **Recruit for Data Literacy:** When hiring, consider data literacy as a valuable skill. This doesn't mean every hire needs to be a data expert, but a basic understanding of data can be a valuable asset.
7. **Continuous Learning:** The field of data science and analytics is rapidly evolving, so it's important to provide opportunities for ongoing learning. This can help your team stay up-to-date with the latest tools and techniques.
8. **Communicate the Value of Data:** Make sure all employees understand how using data effectively can benefit them in their roles, and the organization as a whole. When people understand the value of data, they're more likely to invest the time and effort needed to improve their data literacy.

By improving data literacy, an organization can make more effective use of its data assets, leading to better decision making and improved business outcomes.

## **5. What are some best practices for ensuring data quality in an enterprise data management system?**

ChatGPT:-

Ensuring data quality is a vital component of an effective enterprise data management strategy. Here are some best practices for maintaining high-quality data:



1. **Define Quality:** First, you need to define what data quality means for your organization. This typically includes accuracy, completeness, consistency, timeliness, and relevance. Establishing clear data quality standards helps guide your data collection, storage, and use.
2. **Implement Data Governance:** Establish a data governance framework to oversee data management activities. This framework should include data stewardship responsibilities, data quality metrics, and procedures for resolving data issues.
3. **Validate Data at Entry Points:** To prevent errors, it's crucial to validate data at the point of entry. This might include automated checks for formatting, completeness, or consistency, and could be enforced through data entry forms or applications.
4. **Regular Auditing:** Regular audits can help identify and rectify data quality issues. This could involve checking for duplicate records, outdated information, or inconsistencies in data sets.
5. **Automate Where Possible:** Automating data quality checks can help ensure consistency and reduce the risk of human error. This can involve using data quality software or tools that check for common issues.
6. **Data Cleansing:** Implement a data cleansing process to rectify identified errors. This may involve removing, correcting, or standardizing data.
7. **Integrate Data Sources:** When data is collected from multiple sources, inconsistencies and discrepancies can arise. Using data integration tools can help bring these data sets together and ensure consistency.
8. **Train Employees:** Employees play a vital role in maintaining data quality. Regular training can help ensure they understand the importance of data quality and how to maintain it.
9. **Track Data Quality Over Time:** Use data quality metrics to track the quality of your data over time. This can help identify trends or issues and demonstrate the effectiveness of your data quality initiatives.
10. **Create a Data Culture:** Foster a culture that values high-quality data. This involves demonstrating the importance of data quality and its impact on decision-making and business outcomes.

Remember, maintaining data quality is not a one-time task, but an ongoing process that requires the involvement of the entire organization.

## **6. What role does data governance play in an enterprise data management strategy, and how can it be implemented effectively?**

ChatGPT:-

Data governance is a crucial component of an enterprise data management strategy. It is the overarching policy and procedure that regulates the availability, usability, integrity, and security of the data employed in an enterprise. Here's how data governance plays a crucial role:

1. **Data Quality:** Data governance sets the standards for data quality and ensures these standards are met. This includes setting guidelines for data collection, storage, and use to ensure data is accurate, consistent, and reliable.
2. **Data Security:** Data governance policies outline how data should be protected and who has access to it. This includes physical and digital security measures, as well as procedures for data sharing and disclosure.
3. **Compliance:** Data governance is essential for ensuring compliance with regulations. This includes laws related to data protection, privacy, and industry-specific regulations.
4. **Efficiency and Productivity:** Data governance can streamline data management processes and reduce errors, leading to improved efficiency and productivity.
5. **Decision-Making:** By ensuring data is reliable and easily accessible, data governance supports data-driven decision making.

For effective implementation of data governance, consider the following steps:

1. **Establish a Data Governance Council:** This should be a cross-functional team that includes representatives from different areas of the organization. This council is responsible for developing and overseeing the data governance strategy.
2. **Define Policies and Procedures:** The council should establish clear policies and procedures that outline how data should be handled. This includes rules for data collection, storage, use, and sharing.
3. **Implement Data Governance Tools:** There are many tools available that can help implement data governance. This includes data catalogs to track data assets, data quality tools to monitor and maintain data quality, and data security tools to protect data.
4. **Assign Data Stewards:** Data stewards are responsible for overseeing the implementation of data governance policies in their area of the organization. They

serve as a bridge between the data governance council and the rest of the organization.

5. Training and Education: All employees should be trained on the data governance policies and understand their role in ensuring data is handled correctly.
6. Regular Review and Adjustment: Data governance should not be static. Regularly review and adjust your policies and procedures to reflect changes in regulations, technology, and business objectives.

By implementing effective data governance, organizations can ensure their data is reliable, secure, and used effectively, providing a solid foundation for their data management strategy.

## **7. How important is having a well-structured data architecture, and what are some strategies for achieving this?**

ChatGPT:-

A well-structured data architecture is of paramount importance in an enterprise data management strategy for several reasons:

1. Efficiency: A well-designed data architecture promotes efficient storage, retrieval, and processing of data. It minimizes redundancy and ensures data can be accessed and used efficiently.
2. Data Quality: Good data architecture supports data quality by ensuring consistency, integrity, and accuracy of data.
3. Scalability: As an organization grows and its data needs evolve, a well-structured data architecture can scale to accommodate these changes without compromising performance.
4. Security: A well-planned data architecture helps implement data security and privacy measures effectively, ensuring the protection of sensitive data.
5. Compliance: Data architecture can facilitate compliance with regulations related to data handling and storage.

Here are some strategies for achieving a well-structured data architecture:

1. Understand Your Needs: Before designing your data architecture, you need to understand your data needs. What data will you collect? How will it be used? Who needs access to it? How fast does it need to be processed?

2. **Develop a Data Model:** This outlines how data will be organized, stored, and accessed. It should take into account the types of data you have, the relationships between different data sets, and how data will flow through your systems.
3. **Choose the Right Technology:** Select the data storage and management systems that best fit your needs. This might include databases, data warehouses, data lakes, or other data storage technologies. Also, consider the software for data processing and analysis.
4. **Implement Data Governance:** A good data architecture should align with your data governance strategy. This includes rules for data quality, security, and privacy.
5. **Plan for Scalability:** Your data architecture should be designed to scale as your organization grows. This might involve using cloud-based systems or implementing distributed storage and processing solutions.
6. **Test and Refine:** Once you've implemented your data architecture, test it thoroughly to identify any issues or bottlenecks. Be prepared to refine your architecture as you gain more experience with your data needs.

Remember, creating a well-structured data architecture is an ongoing process that requires regular review and adjustment as your organization's data needs evolve.

## **8. How can an organization assess the maturity of its data management?**

ChatGPT:-

Assessing the maturity of an organization's data management involves evaluating its processes, systems, and culture around data. A mature data management framework typically has well-defined, repeatable, and efficient processes, and a strong culture of data-driven decision making. Here are steps an organization can take to assess its data management maturity:

1. **Establish a Benchmark:** Start by defining what mature data management looks like for your organization. This could be based on industry standards, best practices, or your organization's specific goals. This sets a benchmark against which you can assess your current state.
2. **Review Data Governance:** Examine your data governance policies and practices. This includes data quality standards, data security measures, compliance procedures, and data stewardship roles.

3. Evaluate Data Infrastructure and Tools: Assess the tools and systems you have in place for data storage, processing, and analysis. Are they up-to-date? Do they meet your current and anticipated future needs?
4. Assess Data Culture: Look at how data is used within your organization. Is there a culture of data-driven decision making? Are employees trained and comfortable using data?
5. Check Data Quality: Evaluate the quality of your data. Are there processes in place to ensure data is accurate, consistent, and reliable?
6. Examine Data Integration: If your organization uses data from multiple sources, assess how well this data is integrated. Are there systems in place to ensure data is consistent and accessible across the organization?

To help with the assessment, you can use a data management maturity model. These models, such as the one developed by the CMMI Institute, provide a structured approach to assessing data management maturity. They typically involve a series of stages, from initial (chaotic, ad hoc, individual heroics) to managed (processes are monitored, controlled, and consistently improved).

By regularly assessing your data management maturity, you can identify areas for improvement, track your progress over time, and ensure your data management practices continue to support your organization's goals.

## **9. How can companies manage the challenges of data integration in the context of enterprise data management?**

ChatGPT:-

Data integration - the process of combining data from different sources into a unified view - is a crucial part of enterprise data management. However, it can pose significant challenges, particularly for large organizations or those with complex data environments.

Here are some strategies for managing these challenges:

1. Standardize Data Formats: When data comes from different sources, it's often in different formats. Standardizing data formats can make it easier to combine and use data. This may involve creating a common data model that all data must conform to.

2. **Implement Data Governance:** Clear data governance policies can guide data integration efforts. This includes defining how data should be handled, who is responsible for it, and how data quality will be maintained.
3. **Use Data Integration Tools:** There are many tools available that can automate the process of data integration. These tools can help manage complex data environments, perform transformations, and ensure data is consistent and reliable.
4. **Ensure Data Quality:** Poor data quality can make data integration more difficult and can lead to unreliable results. Regularly check and improve the quality of your data to support integration efforts.
5. **Plan for Scalability:** Your data integration strategy should be scalable to accommodate future growth or changes in your data environment. This may involve using cloud-based systems or distributed computing architectures.
6. **Handle Data Security and Privacy:** Data integration can raise issues of data security and privacy, particularly when integrating data from external sources. Ensure you have robust data security measures in place, and that all data use complies with relevant privacy regulations.
7. **Consider a Data Integration Specialist:** Depending on the size and complexity of your data environment, it may be worthwhile to have a data integration specialist on your team. This person or team can focus on the technical and strategic aspects of data integration.
8. **Train Your Team:** Make sure your team has the skills and knowledge they need to manage data integration. This may involve training in specific tools, as well as broader education about data management principles.
9. **Regular Review and Adaptation:** Regularly review your data integration strategy and adapt it as needed. This will help you respond to changes in your data environment, new business needs, or advances in data integration technology.

By addressing these challenges, companies can make data integration a smoother, more efficient process, which in turn supports more effective enterprise data management.

## **10. How can data security and privacy be ensured in enterprise data management?**

ChatGPT:-

Ensuring data security and privacy is a fundamental aspect of enterprise data management. Here are several strategies that organizations can implement to secure their data and protect privacy:

1. **Data Governance:** Establish a data governance policy that outlines the rules for handling and protecting data. This includes defining who can access data, how it can be used, and what security measures should be in place.
2. **Access Controls:** Implement strong access controls to ensure that only authorized individuals can access your data. This could involve password protection, two-factor authentication, or more advanced security measures.
3. **Encryption:** Use encryption to protect data both when it is in transit (being transferred over a network) and at rest (stored in databases, hard drives, etc.). Encryption converts data into a format that can only be read by someone with the decryption key.
4. **Regular Audits:** Conduct regular audits to identify potential vulnerabilities and ensure compliance with your data security policies and procedures.
5. **Data Anonymization:** If you are working with sensitive data, consider using techniques such as anonymization or pseudonymization to protect privacy. These methods involve removing or changing identifying information so that it cannot be linked back to individuals.
6. **Incident Response Plan:** Have a plan in place for responding to data breaches or other security incidents. This should include steps for identifying and containing the breach, investigating what happened, notifying affected parties, and preventing future incidents.
7. **Staff Training:** All staff members should be trained on data security practices. This includes understanding the risks, recognizing potential threats (such as phishing attacks), and following proper procedures for handling and storing data.
8. **Updates and Patches:** Ensure all systems and software are kept up to date with the latest security patches and updates. This can help protect against known vulnerabilities.
9. **Compliance with Regulations:** Ensure your data management practices comply with relevant regulations. This might include the General Data Protection Regulation (GDPR) in the EU, the California Consumer Privacy Act (CCPA) in the U.S., or other local or industry-specific regulations.
10. **Data Security Tools:** Use tools designed to protect data security. This can include firewalls, intrusion detection systems, anti-virus software, and data loss prevention (DLP) tools.

Remember, data security and privacy are ongoing responsibilities that require constant vigilance and regular review of policies and procedures.

## **11. How do emerging technologies impact the future of enterprise data management?**

ChatGPT:-

Emerging technologies can significantly impact enterprise data management by offering new capabilities, changing the way data is collected and processed, and creating new opportunities and challenges. Here are a few key technologies to consider:

1. **Artificial Intelligence (AI) and Machine Learning (ML):** These technologies can automate and enhance many aspects of data management. For instance, machine learning algorithms can help with data cleansing by detecting anomalies and errors. They can also improve data integration and enable advanced analytics.
2. **Internet of Things (IoT):** IoT devices generate vast amounts of data that need to be managed. This can increase the complexity of data management but also provides opportunities for real-time data analysis and decision making.
3. **Cloud Computing:** Cloud platforms have revolutionized data storage and processing. They offer scalable, flexible, and often cost-effective solutions for data management, allowing for distributed data architecture and advanced processing capabilities.
4. **Blockchain:** While commonly associated with cryptocurrencies, blockchain has implications for data management as well. Its decentralized, transparent, and immutable nature can provide enhanced security and traceability for data transactions.
5. **5G Technology:** The increased speed and capacity of 5G networks can accelerate data collection, transmission, and processing. This can support real-time data analytics and the management of large data volumes, particularly from IoT devices.
6. **Edge Computing:** This involves processing data closer to its source rather than in a centralized data-processing warehouse. It's particularly relevant for IoT data, reducing latency, and improving real-time decision making.
7. **Data Fabric and Data Mesh:** These are architectural approaches that provide a unified data access layer over distributed data sources, making it easier to manage and access data across a large enterprise.



8. **Data Privacy Technologies:** In response to increasing regulatory scrutiny and public concern over data privacy, new technologies for data anonymization, pseudonymization, and secure multi-party computation are gaining importance.
9. **Augmented Data Management:** Gartner has identified this as a trend, where machine learning and AI are used to enhance data management tasks like data quality, metadata management, privacy management, and more.

These technologies are rapidly evolving, and their implications can vary depending on the specifics of an organization's data environment and business objectives. Staying current with these developments and understanding how they can be applied to your enterprise data management efforts is crucial for future-proofing your data strategy.

## **12. How can artificial intelligence and machine learning contribute to optimizing data management?**

ChatGPT:-

Artificial Intelligence (AI) and Machine Learning (ML) have the potential to revolutionize data management by automating processes, improving accuracy, and deriving insights. Here are several ways AI and ML can contribute to optimizing data management:

1. **Data Cleaning and Preprocessing:** AI and ML algorithms can automate the cleaning and preprocessing of data. They can detect and correct errors, handle missing data, and standardize formats, which are traditionally time-consuming tasks.
2. **Data Classification and Categorization:** ML models can automatically classify and categorize data based on predefined criteria or learned patterns. This helps in organizing data more effectively and making it easier to retrieve and use.
3. **Anomaly Detection:** ML models can be trained to recognize patterns in data and subsequently identify anomalies or outliers that may indicate data quality issues or potential security threats.
4. **Predictive Analytics:** AI and ML can analyze historical data to make predictions about future events. This can be invaluable for decision-making in various business domains like sales forecasting, customer churn prediction, and inventory management.

5. Natural Language Processing (NLP): AI-powered NLP can be used to extract information from unstructured text data, making it possible to analyze and utilize data from sources like emails, customer reviews, or social media.
6. Data Integration and Transformation: AI can help automate the process of integrating data from different sources, identifying relationships and mapping data from one format to another, thus streamlining the ETL (Extract, Transform, Load) process.
7. Metadata Management: AI can be used for generating and managing metadata, which helps in understanding and using data more effectively. For example, AI can auto-tag data or provide semantic search capabilities based on metadata.
8. Optimized Query Performance: Machine learning can be used to optimize query performance by predicting the best execution plans or indexing strategies, leading to faster data retrieval.
9. Data Security and Privacy: AI and ML can enhance data security by detecting unusual access patterns or breaches, and also by helping in the anonymization of data to preserve privacy.
10. Chatbots and Virtual Assistants: AI-powered chatbots and virtual assistants can help users access and interact with data more easily, without requiring advanced technical skills.
11. Automated Decision-Making: AI systems can help automate certain decision-making processes by analyzing data in real-time and applying predefined criteria or learned patterns to make decisions.

As organizations continue to collect and rely on larger amounts of data, the application of AI and ML in data management will likely become even more critical. It's important, however, to also be aware of the ethical considerations and potential biases that can arise from the use of AI and ML and address these proactively.

### **13. What role does data management play in supporting digital transformation?**

ChatGPT:-

Data management plays a pivotal role in supporting digital transformation. Digital transformation involves the use of digital technologies to create new or modify existing business processes, culture, and customer experiences to meet changing business and market requirements. Data is at the heart of this process. Here's how data management supports digital transformation:

1. **Informed Decision-Making:** Effective data management ensures the availability of high-quality, accurate, and timely data. This empowers businesses to make informed, data-driven decisions, which is critical to successfully navigating a digital transformation.
2. **Enhanced Customer Experience:** Data management allows businesses to leverage customer data effectively, leading to a more personalized and enhanced customer experience. Personalization is a key aspect of digital transformation, enabling businesses to offer tailored products, services, and interactions based on individual customer preferences and behaviors.
3. **Operational Efficiency:** Through the automation of data collection, storage, processing, and analysis, businesses can streamline their operations, eliminate manual errors, and save time and resources. This improved efficiency is a key objective of digital transformation.
4. **Innovation:** With a solid data management foundation, businesses can leverage advanced technologies like AI, machine learning, and predictive analytics to drive innovation, create new business models, and stay ahead of the competition.
5. **Risk Management:** Digital transformation often involves increased data collection and sharing, which can pose risks in terms of data security and privacy. Effective data management includes robust data security measures and ensures compliance with data protection regulations, helping to manage these risks.
6. **Scalability:** Good data management practices, such as the use of cloud-based storage and processing solutions, can make it easier to scale digital technologies across the organization.
7. **Data Culture:** A strong data management strategy can help foster a data-driven culture, where data is recognized as a valuable asset and is used consistently to drive decision-making. This cultural shift is often a key part of digital transformation.

By ensuring the availability, quality, security, and effective use of data, data management can provide the foundation needed for a successful digital transformation.

#### **14. How should organizations deal with the challenges of managing unstructured data?**

ChatGPT:-

Unstructured data – data that does not have a predefined model or format, such as text documents, emails, videos, audio files, social media posts, and more – constitutes a significant portion of the data generated today. Managing unstructured data poses certain challenges due to its size, lack of standardization, and the complexity of extracting meaningful information from it. Here are some strategies for dealing with these challenges:

1. **Data Cataloguing and Metadata:** Use data cataloguing tools to classify and index unstructured data. Create metadata that describes the content, source, and context of the data. This will make it easier to manage, search, and retrieve the data when needed.
2. **Implement Data Governance:** Establish data governance policies that cover unstructured data as well. This includes policies on data storage, security, privacy, and usage.
3. **Use Specialized Tools and Technologies:** There are tools and technologies designed specifically for managing and analyzing unstructured data. For instance, natural language processing (NLP) can be used to analyze text data, while image recognition and computer vision can be used for image data.
4. **Data Storage:** Choose appropriate data storage solutions that can handle the volume and variety of unstructured data. Cloud-based storage solutions can be a good choice due to their scalability and flexibility.
5. **AI and Machine Learning:** AI and machine learning algorithms can help extract meaningful insights from unstructured data. For example, sentiment analysis can be used to understand customer opinions from social media posts or reviews.
6. **Data Security and Privacy:** Unstructured data can contain sensitive information. Ensure you have robust data security measures in place, and that your handling of unstructured data complies with relevant privacy regulations.
7. **Data Integration:** If you're combining unstructured data with structured data, ensure you have processes in place to integrate these different types of data effectively.
8. **Data Lake Architecture:** Consider using a data lake architecture for storing unstructured data. Data lakes are designed to store large volumes of raw data in its native format until it is needed.

By implementing these strategies, organizations can better manage their unstructured data, making it easier to store, analyze, and extract valuable insights from this type of data.