

MGI

Master's Degree Program in
Information Management

How Efficient Data Governance and Systems Management Influences a Digitally Driven Company

A PSI 20 Telecommunications Company Use-Case

Tomás Barral Paiva

Dissertation

presented as a partial requirement for obtaining the Master's Degree Program in Information Management with a
specialization in Systems and Information Technology

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

How Efficient Data Governance and Systems Management Influences a Digitally Driven Company

A PSI 20 Telecommunications Company Use-Case

Tomás Barral Paiva

Dissertation presented as the partial requirement for obtaining a Master's degree in Information
Management, specialization in Information Systems and Technologies Management

Supervisor/Orientador(a): Prof. Doutor Pedro Manuel Carqueijeiro Espiga da Maia Malta

Novembro 2023

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledge the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Lisbon, 28th of November

DEDICATION

I dedicate this thesis to a distinguished member of our team whose relentless dedication and exceptional performance in managing client sales have been instrumental in our achievements. Your unwavering commitment to meeting and surpassing goals and objectives has been an inspiration to us all. May this dedication serve as a token of appreciation for your invaluable contributions, and I sincerely hope that this acknowledgment encapsulates the gratitude and respect we hold for your outstanding efforts.

ACKNOWLEDGEMENTS

I extend my heartfelt gratitude to those who have been instrumental in the completion of this thesis.

First and foremost, I am immensely grateful to my girlfriend, Catarina, for her unwavering patience, understanding, and support throughout this academic journey. Her enduring patience during my periods of absence has been invaluable and deeply appreciated.

I would like to mention my dear friend, André Lopes, that helped me in my most difficult times and has always been present in my life. Without him this thesis would have been submitted far later.

I owe a debt of gratitude to my colleagues and coworkers for their collaborative efforts, guidance, and encouragement. Their contributions have enriched this work immeasurably.

To my beloved family, I express my deepest appreciation for their unwavering support and encouragement. In particular, I am indebted to my grandmother, whose unwavering presence and support have been a constant throughout my life. Her availability and unwavering belief in me have been a pillar of strength.

I also wish to honor the memory of my deceased grandmother, whose aspiration for her only grandson to complete his Master's degree has been a driving force. Her enduring wish has been a guiding light, motivating me to persist and achieve this milestone.

Finally, but not less important, I would like to thank Prof.Dr. Pedro Malta to all the help and trust in me. His guidance is directly responsible for this deliverance.

I am profoundly thankful to all those mentioned above and countless others whose support and encouragement have been invaluable in this academic pursuit.

ABSTRACT

In an era of rapid technological advancements, the corporate landscape continually evolves, often driven by mergers and acquisitions designed to strengthen market positions and enhance business capabilities. The fusion of company X and company Y into company Z, a prominent player in the telecommunications sector present in the PSI 20, stands as a testament to this trend. From a business perspective, this fusion represents a strategic masterpiece, but it also poses formidable challenges within the realm of information systems governance and data management.

This master's thesis explores the intricacies surrounding information systems architecture and governance in the context of the companies X and Y merger. Our primary focus is on the challenges that arise in the handling of vast data repositories, particularly regarding client portfolios and personal information. Merging two distinct information ecosystems while ensuring data coherence is no small feat, and this research seeks to unravel the current status of the systems in question and hopefully provide valuable suggestions. The study encompasses a multi-faceted approach. Initially, the existing information systems architecture and governance practices within company Z will be assessed. This evaluation involves an in-depth analysis of data management protocols, system integration strategies, and the alignment of IT initiatives with overarching business goals.

In light of the findings from the assessment, we embark on a rigorous analysis of data inconsistencies and redundancies within the client portfolio and personal information databases. This process involves an exploration of various data management practices, ranging from data acquisition to storage and dissemination. The culmination of this study is the development of a proposal aimed at enhancing the existing information systems and rectifying data incoherencies. This proposal will encompass a detailed roadmap for reconfiguring data architectures and the implementation of robust governance frameworks. By applying industry best practices and leveraging cutting-edge technologies, we aim to provide actionable insights that will empower company Z to optimize data management and promote information systems cohesion. This thesis presents a unique opportunity to shed light on the intricate interplay between corporate fusion, information systems governance, and data management, using Company Z as a case study. Through meticulous evaluation and strategic recommendations, it aims to contribute to the broader discourse on managing information systems and data in the dynamic landscape of corporate transformations.

KEYWORDS

Information Technology; Information Systems; Information Management; Data Governance; Data Coherency; Information Systems Management.

Sustainable Development Goals (SGD):



INDEX

1. Introduction	1
2. Literature review	3
2.1. Information Technology and Information Systems.....	3
2.2. Information Management and Data Governance	4
2.3. Data Quality and Data Coherency	6
2.4. Information Dissemination.....	7
2.5. Data in Digital Strategy	8
3. Methodology	10
4. Available Data	12
4.1. Current Systems	17
5. Field Work.....	19
5.1. Data Analysis	19
5.2. Digital Campaign Use Case	22
5.3. Focus Group.....	22
6. Proposal.....	25
6.1. Understanding Data Mesh	25
6.2. Fundamentals of Data Mesh Architecture	25
6.3. Resolving Challenges Through Data Mesh	25
6.4. Use Cases of Data Mesh	26
6.5. Principles of Data Mesh Architecture	27
6.6. Implementation Strategies for Building Data Mesh.....	27
6.7. Conclusion and Disclaimer	27
7. Conclusions, future works and Limitations	29
8. Bibliographical References	30

LIST OF FIGURES

Figure 1 - The layered architecture of capabilities and roles in data mesh (Goedegebuure et al., 2023).....	8
Figure 2 - Examples of client Identifiers	12
Figure 3 - Examples of GDPR data	13
Figure 4 - Examples of Portfolio data	14
Figure 5 -Examples of Billing data	14
Figure 6 - Examples of Geographical data.....	15
Figure 7 - Examples of Utilization data	15
Figure 8 - Examples of Campaigns data	16
Figure 9 - Examples of Sales data	16
Figure 10 - Examples of Client Service data	17
Figure 11 - Examples of Retail data	17
Figure 12 - Current Systems by Domain.....	18
Figure 13 - SAs' by department.....	19
Figure 14 - Incoherencies between departments	20
Figure 15 - Data coherency between departments	21

LIST OF ABBREVIATIONS AND ACRONYMS

Arpu: Average Revenue Per User

B2C: Business-To-Consumer

E2E: End to End

GDPR: General Data Protection Regulation

IS: Information Systems

IT: Information Technology

IVR: Interactive Voice Response

NRT: Near Real Time

1. INTRODUCTION

In today's rapidly evolving digital landscape, the role of information systems architecture and management, alongside efficient data governance, has become instrumental in shaping the success trajectory of businesses. As the world becomes increasingly digitized, organizations across sectors are recognizing the paramount importance of leveraging data coherency, streamlined systems management, and robust data governance practices. This recognition stems from the realization that these factors are pivotal in ensuring operational efficiency, fostering innovation, and ultimately sustaining competitive advantage.

This thesis explores the nexus between efficient data governance, systems management, and their impact on a digitally driven company, focusing on the telecommunications sector in Portugal. With the pervasive integration of technology in every facet of modern society, the telecommunications industry stands as a pivotal player, acting as the conduit for the flow of data and information. Consequently, the quality and coherence of data within these systems become critical components determining the operational success and adaptability of companies within this sector.

Within the company studied, the current state of information systems architecture and management often grapples with inherent challenges related to data flux and coherency. Despite technological advancements, this company encounters difficulties in maintaining data coherence across departments, leading to inefficiencies, data loss, and fragmented systems. These challenges hinder seamless project development and execution, often resulting in delays, increased costs, and suboptimal outcomes.

To comprehensively understand the complexities and challenges faced in developing projects across departments due to data incoherencies and loss, this thesis employs a focused approach. A primary methodology involves conducting a rigorous focus group analysis within a PSI20 Telecom company. This approach aims to unearth the nuanced difficulties experienced by different teams within these organizations concerning data coherence and its impact on project development.

Furthermore, this research endeavors to provide a tangible understanding of the implications of data incoherency and loss. To quantify the impact, a specific use case of a digital campaign will be analyzed to ascertain the direct correlation between data coherence issues and revenue loss. This exercise is pivotal in elucidating the tangible financial repercussions of inefficient data and systems management, emphasizing the criticality of implementing robust data governance practices and streamlined systems management protocols.

In summary, this thesis seeks to shed light on the crucial interplay between efficient data governance, systems management, and the operational success of digitally driven companies, particularly within the telecommunications sector in Portugal. By delving into the intricacies of data coherence challenges and their quantifiable impact, this research aims to offer actionable insights that can guide these companies toward adopting measures that enhance operational efficiency and profitability in an increasingly digital world.

2. LITERATURE REVIEW

The main objective of this literature review is to understand the importance of Information Technology and Information Systems. Having accomplished that, it's important to understand information systems architecture, data governance, and information management. Finally, how that influences proper information dissemination and ensures data coherency. The focus is on client information from a digital strategy perspective.

2.1. INFORMATION TECHNOLOGY AND INFORMATION SYSTEMS

Information Technology (IT) refers to the comprehensive domain that encompasses the use of computer systems, software, networking infrastructure, and associated technologies to manage, process, and disseminate data and information. It is a dynamic field that has evolved over time and plays a pivotal role in various aspects of modern life, including business, education, communication, entertainment, and more (Moore, 1965; Gilder, 1990). There have been several pivotal moments in the history of information technology. Moore's law, for instance, illustrates the exponential increase in hardware efficiency and processing power over time (Moore, 1965). The present IT landscape was made possible by the development of microprocessors and the digitalization of information (Gilder, 1990), thus these developments have led to significant modifications in data processing capacities, allowing companies and institutions to manage ever-more complicated data sets.

Modern corporate processes are inextricably related to information technology. According to Laudon and Laudon (2021), IT systems improve organizational efficiency and competitiveness by facilitating essential tasks including supply chain management, customer relationship management, and enterprise resource planning. Organizations depend on IT applications and infrastructure to streamline operations and maintain flexibility in ever-changing marketplaces. A strong IT infrastructure is the cornerstone upon which modern enterprises base their operations, data management, and communication in the digital age. It is impossible to exaggerate how important this kind of infrastructure is for effective data processing, safe data storage, and smooth information sharing both inside and outside of organizations. With the continuous growth in data volume and complexity, a well-structured IT infrastructure allows firms to handle and analyze data at rapid speeds, resulting in better decision-making and increased operational efficiency. This is made possible by the continual development in data volume and complexity. It makes it possible to implement a variety of tools and applications that are essential to contemporary corporate operations, including business intelligence platforms, ERP software, and CRM (customer relationship management) systems (Laudon & Laudon, 2021).

Advancements in technology have provided many ways for information technology to be incorporated into the business world. One that gained immense popularity in recent years is Information Systems (IS). They are a great depiction of how humans and technology interact to provide a better output for business purposes. It's difficult to predict the future, but it is with great certainty that IS will be present in organizations for a long time. So, it's safe to conclude that it's a field that will be subject to great advancements, mainly in automation and the use of AI. Information Technology (IT) and Information Systems (IS) work together to help businesses gain strategic advantage, so both need to be properly defined. Information Technology refers to any computer-based tool that people use to work with

information and support an organization's information and information-processing needs (Rainer, R. K., & Prince, B., 2022).

IT has had a great effect on individuals, societies, economies, and in our planet. It is making our world smaller, driving globalization, and enabling more communication between people that otherwise would not be possible. IT is shorting Milgram's theory, which established that any two random people in the USA are only six people apart from each other. In addition to bringing people together, information technology since the beginning of its implementation has had a tremendous impact on organizations, in fact, if there is a fail it's common to cease operations until the problem is fixed. One example of that is the LAX 2014 incident, an airplane with a complex flight plan caused the ERAM (En Route Automation Modernization) system to fail due to lack of memory. This simple and common problem costed the company two billion dollars, and it took more than twenty-four hours to fix the problems. This shows how IT impacts businesses and consequently people's lives.

An Information System (IS) collects, processes, stores, analyzes, and disseminates information for a specific purpose (Rainer, R. K., & Prince, B., 2022). So, ISs collect (input) and manipulate data (process) and generate and distribute reports (output) based on the data-specific IT services, such as processing customer orders and generating payroll, which are delivered to the organization. Finally, the ISs save (store) the data for future use. These systems, comprising hardware, software, networks, and data resources, play a fundamental role in facilitating communication, streamlining operations, and driving decision-making processes across various industries. The importance of information systems cannot be overstated, as they are instrumental in enhancing efficiency, innovation, and competitiveness in both business and societal contexts. One of the most popular business strategies for achieving success is the development of a competitive advantage. Competitive advantage exists when a company has superior resources and capabilities than its competitors which allow it to achieve either a lower cost structure or a differentiated product. For long-term business success, companies strive to develop sustainable competitive advantages, or competitive advantages that cannot be easily copied by the competition (Porter, 1998). To stay ahead, corporate leaders must constantly seek new ways to grow their business in the face of rapid technology changes, increasingly empowered consumers and employees, and ongoing changes in government regulation.

2.2. INFORMATION MANAGEMENT AND DATA GOVERNANCE

With the development of IT and IS, data collection became a reality every company adopted, and transforming data into information is extremely important. That is why information management (IM) and data governance have become indispensable disciplines in modern businesses. Information management guides the acquisition, custodianship, and distribution of corporate data and involves the management of data systems, technology, processes, and corporate strategy. Data governance, or information governance, controls enterprise data through formal policies and procedures and provides employees and stakeholders with high-quality data they can trust and access on demand. It has to do with the procedures, guidelines, and standards that control data management, making sure that information is secure, correct, and useful for operational and decision-making processes. Additionally, there are some principles that ought to be considered such as: Accountability - clear accountability and ownership for data are necessary for data governance in order to guarantee that data is managed efficiently, roles and duties for data custodians and stewardship must be clearly defined (Gupta, 2014);

Transparency - organizations must establish transparent data quality and governance policies, making them accessible to all relevant stakeholders (Wang, 1998). Data Lifecycle Management - reliability, security, and regulatory compliance are all ensured by effective data governance, which covers the full data lifecycle from creation to archiving and deletion (Roberts & Reid, 2016).

In fact, bad decisions and results can result from the analysis of inaccurate data, which is widely referred to as dirty data, and lead to increased costs, decreased revenue, and legal, reputational, and performance-related consequences (Turban, E., Pollard, C., & Wood, G., 2018). For example, marketing campaigns based on the segmentation of wrong client information can result in poor performance, loss of revenue, misuse of company funds, and even legal troubles. Which in turn results in reduced client satisfaction and loyalty. Information management is the use of IT tools and methods to collect, process, consolidate, store, and secure data from sources that are often fragmented and inconsistent (Turban, E., Pollard, C., and Wood, G., 2018). Data is the central part of an information system, and most business initiatives fail or succeed based on the data available, so companies are dependent on their systems to deliver on-demand quality and usable data.

Thus, there are crucial components to data governance such as Data Stewardship – data governance hinges on assigned data stewards who bear the responsibility for upholding data quality, integrity, and adherence to regulations. These experts supervise data management, guaranteeing that data is precisely defined, properly documented, and in compliance with established norms, as asserted by Boeckhout, Zielhuis, and Bredenoord (2018); Data Standards – It encompasses data definitions, naming conventions, and formatting guidelines to ensure consistency and uniformity in data usage (Redman, 2008); and lastly Data Quality Frameworks - Implementing data quality frameworks, which include techniques, procedures, and metrics for evaluating and improving data quality, is a component of data governance (Wang and Strong, 1996). A methodical approach to data quality management is offered by frameworks such as the Total Data Quality Management (TDQM) framework.

In fact, when the previous components are not considered, organizations face information management challenges and most of the time data is scattered in many different systems, possibly with different privacy policies, so employees have limited access to them. Despite being the generation of companies with more data available than ever before, users can have a difficult time accessing it when needed if proper guidelines are not implemented. Easy access to large volumes of information, management must also deal with information degradation where users can't understand what the data means or what it is useful for. Information deficiencies are common, the book "Information technology for management: On-demand strategies for performance, growth and sustainability" (Turban, E., Pollard, C., & Wood, G., 2018) indicates some common reasons for information deficiencies:

1. Data Silos: Data silos are stand-alone data stores. Their data is not accessible by other ISs that need it or outside that department. This makes communication between departments difficult and can cause data incoherencies.
2. Lost or Bypassed Data: Data can simply get lost between systems, which can cause a decrease in information value that in turn has a great effect on companies' decisions.
3. Poorly Designed Interfaces: These interfaces require extra time and effort to use which increases the likelihood of misunderstanding or ignoring the data.

4. Nonstandardized data formats: When users are presented with inconsistent or nonstandardized formats, errors increase. Attempting to analyze or compare this data is more difficult. For example, if hours are in different time zones, the user cannot compare the values before converting the data to a unique format.
5. Difficult to hit moving targets: The information the company is after keeps changing faster than the ISs can operate, because of the four reasons mentioned above, it's impossible to have near real-time proper information.

This is why information is important: it increases data quality and provides trustworthy information for decision-making processes, helping businesses capitalize on the full value of their systems.

2.3. DATA QUALITY AND DATA COHERENCY

There is a strong connection between data quality and information technology. Information systems and technologies are essential for maintaining data quality since they manage large datasets. Data quality assessment and improvement approaches are provided by frameworks such as the Data Quality Assessment Framework (Lee et al., 2018). For the purpose of keeping timely, accurate, and reliable data, information technology procedures and tools are essential. Admittedly, reliable analytics, well-informed decisions, and efficient operations all depend on high-quality data that includes data that is timely, accurate, consistent, reliable, and full (Wang & Strong, 1996). Several fundamental ideas are involved in ensuring data quality:

1. Data Profiling: Improving the quality of data starts with an understanding of its composition and structure. Organizations can find anomalies and disparities by using data profiling tools (Couto et al., 2022).
2. Data Cleansing: The process of data cleansing entails finding and fixing mistakes, inconsistencies, and duplicates in datasets. By taking these steps, data integrity and accuracy are guaranteed (Kimball et al., 2008).
3. Data Validation: Data integrity and validity must be confirmed, and methods for doing so include cross-field and cross-record validation (Redman, 1996).
4. Data Standardization: Data quality depends on consistent coding structures, naming conventions, and data formatting, and those standards reduce errors and inconsistencies (Gupta, 2014).

The consistency and integrity of data across several locations or caches in distributed systems and caching is referred to as data coherency. In these situations, data may be present in multiple copies on different systems or caches, thus keeping coherency makes sure that all of these copies are accurate and consistent. Admittedly, it guarantees that any changes made to data in one location that is read or modified will also be reflected in all other pertinent locations where the data is cached or stored. In addition, this requires various synchronization mechanisms such as cache coherence protocols in multiprocessor systems (Tanenbaum & van Steen, 2006) which specify guidelines for the sharing and updating of data between various caches and memory locations. In turn, those mechanisms synchronize data across distributed systems through a variety of techniques, which could entail utilizing distributed databases, disseminating updates, or applying strategies like distributed locks and transactions. Indeed, when there is concurrent access to data, data coherency is essential, by making

sure that read and write operations done simultaneously don't produce contradictory or inconsistent data states.

2.4. INFORMATION DISSEMINATION

As established beforehand, ensuring that the right information is available to the right people and systems at the right time plays a crucial role, and that is linked to the information dissemination. This concept refers to the process of dispersing knowledge, data, or information to different users, systems, or stakeholders, and its importance is connected to the fact that it supports and encourages collaboration, communication, and well-informed decision-making. For businesses, government agencies, and academic institutions to operate effectively, timely access to accurate data is essential. There are several methods to diffuse it, those including: reports, data sharing platforms, dashboards, among others, however it presents several challenges such as data quality, accessibility and security, which can be mitigated, in part by a solid data governance (Boeckhout, Zielhuis, and Bredenoord, 2018).

In contemporary data management, an emerging domain-driven decentralized data architecture known as "Data Mesh" is gaining attention as a way to enhance information dissemination and minimize or avoid operational bottlenecks associated with centralized, monolithic data architectures in enterprises. Four guiding principles define the architecture of data mesh: data as products, domain ownership of data, self-serve data platform, and federated computational governance. Data assets within an organization are products that let users find and use domain data quickly and independently. Data products are created and managed by domain teams that are closest to the data source. When domains consume data products from one another, a mesh of data products is created. A federated governance team is in charge of creating and implementing product interoperability standards and compliance policies throughout the data mesh to guarantee the data products are compatible and adhere to internal and external laws and policies (Goedegebuure et al., 2023).

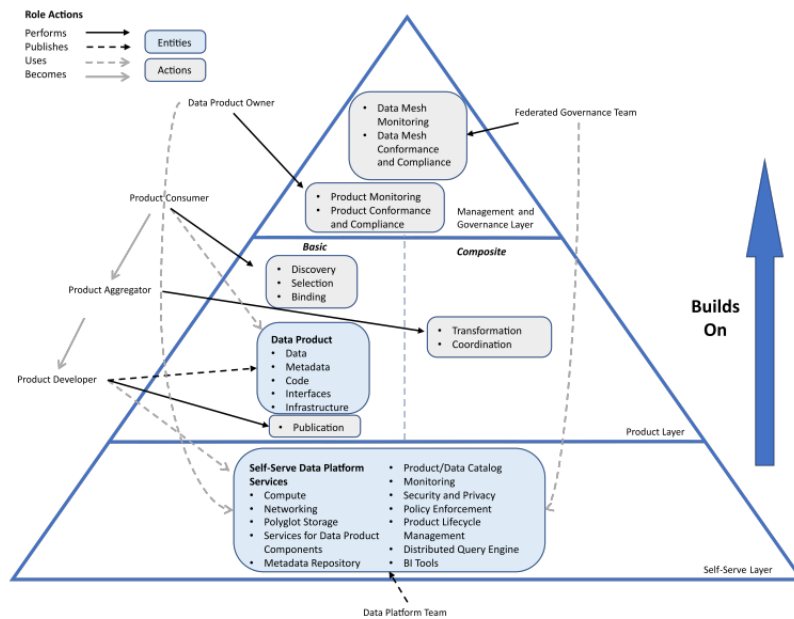


Figure 1 - The layered architecture of capabilities and roles in data mesh (Goedegebuure et al., 2023)

2.5. DATA IN DIGITAL STRATEGY

With today's digital environment, data plays an increasingly important role in determining organizational strategies. Incorporating data into digital strategies not only yields valuable insights but also poses distinct obstacles that enterprises must surmount to fully realize its possibilities. Therefore, it is paramount to shed some light on the complex relationship between data and digital strategy, highlighting the obstacles and game-changing implications.

"Competing on Analytics: The New Science of Winning" by Davenport and Harris (2007) makes a strong case for why businesses that use data analytics have an advantage over their competitors. In the digital sphere, data serves as the foundation for well-informed decision-making, allowing businesses to anticipate and influence market developments in addition to reacting to them. There is no question about the strategic imperative: companies that can effectively compete on analytics will be in a better position to handle the complexity of the digital landscape.

In addition, data driven strategies allow companies to be more precise in an ever-shifting setting as stated by Chaffey and Ellis-Chadwick (2019), who explore the complexities of digital marketing as well as the mutually beneficial relationship between data and digital marketing tactics. In an environment where consumer behavior is constantly changing, data-driven insights are essential for developing marketing strategies that are both focused and successful. One of the main tenets of contemporary marketing is personalization, which allows businesses to adjust their strategies to each customer's preferences and behaviors. However, personalization can only be achieved by careful data analysis.

Nevertheless, there are plenty of challenges whilst making use of data into data strategies. Concerns about data privacy and security, as brought up by Cavoukian and Castro (2010), highlight how important it is for businesses to manage complicated regulatory environments while putting strong security measures in place to safeguard client information. The need to approach innovation,

creativity, and competitiveness from a "design-thinking" standpoint—that is, an integrative, multidisciplinary, innovative, and inspiring way of looking at the world and overcoming obstacles—is becoming increasingly apparent. The same design-thinking approach needs to be used when dealing with privacy: networked data systems and technologies need to include privacy by default.

In addition, there is the challenge of data integration and quality, which requires organizations to address issues of accuracy and reliability since data inaccuracies have the potential to jeopardize the integrity of decision-making processes and reduce the efficacy of digital strategies (Inmon, 2002). Accurate data guarantees that the company strategy has a firm basis, reducing the possibility of unfavorable outcomes stemming from inadequate data quality. Businesses can make well-informed decisions based on factual information when they have accurate data. Better decision-making, more successful strategies, and increased operational efficiency result from this. On the other hand, inaccurate data may lead to poor choices, resource waste, and even reputational harm to a company. Throughout its lifecycle, data must remain dependable, accurate, and consistent and this is known as data integrity. Both regulatory compliance and preserving the trustworthiness of an organization's data depend on this. Information that is erroneous or lacking due to compromised data integrity can have a detrimental effect on operations, decision-making, and regulatory compliance (Inmon, 2002).

With that in mind, when data is used wisely in digital strategies, the results are revolutionary. Real-time insights-driven informed decision-making enables organizations to react quickly to changes in the market and long-term relationships can be fostered by customer-centric strategies that improve customer satisfaction and loyalty through an understanding of consumer behaviors. This allows organizations to lower their operating costs and streamline processes through operational optimization, which is a direct outcome of data-driven insights. Utilizing data to recognize new trends and act quickly on opportunities helps businesses stay innovative and competitive.

All things considered, data in digital strategy is a two-edged sword that can have both revolutionary effects and difficulties in which organizations have to deal with the challenges of using data effectively, protecting privacy, guaranteeing data quality, developing expertise, and navigating moral dilemmas. By overcoming these obstacles, data can be fully utilized, revolutionizing tactics for success in the digital age. Organizations hoping to prosper in the rapidly changing digital landscape must strike the correct balance between using data for strategic advantages and resolving related issues (Davenport and Harris, 2007)

3. METHODOLOGY

If you cook food for the first time, it is a good idea to have some guidelines or even a cookbook to provide you with a process or method to do it (Venable, J. R., Pries-Heje, J., & Baskerville, R. L., 2017). Design Science Research (DSR) is a new research approach and paradigm that helps users to apply the correct research methodology when researching Information Systems.

First, this master's thesis is based on the identification of a problem within a PSI 20 company and working towards its awareness quantifying use cases to show the magnitude of the research opportunity. After identifying the problem and its importance, a suggestion or proposal is presented in order to eliminate the obstacle, and with the help of a use case and a focus group this proposal will be developed and evaluated.

Given this information, the methodology chosen in DSR Process Model (DSRPM) is comprised of five steps, according to Vaishnavi & Kuechler (2004, 2007, 2015):

1. Awareness of the problem: The problem consists in unquantifiable data incoherencies between departments and data sources. This was first noticed when crossing data between departments to create digital campaigns. What was evaluated was that was very difficult for the different teams to understand what the correct information was about client's data. This caused unnecessary filters and exclusion that obligated impacting less clients in the digital campaigns which in turn causes a loss of revenue and most important an ignorance about the current situation of the company's clients.
2. Suggestion: Due to the requirements each team demands about the information that is needed for the well-functioning of their tasks it is difficult for the information systems to accommodate all data demands. So, what is happening is a great development of ad-hoc processes built by each team to get the data necessary and a lack of ownership about the different data available in the company. Given this, it is important to create a new framework that solves these incoherencies and lack of ownership. A framework with decentralized ownership that names a clear responsible for each available data and it's the only source of information for all departments. On top of that, it is important to create a robust data dissemination system to enable easy access to the data and a clear data privacy policy to prevent breaches of client information.
3. Development: In this step, we asked for the help of three different departments at the company: CRM (Customer Relation Management), Commercial and, Experience and Value. From these departments, we asked for the same data and a simple functional description of what the data is. The objective is to quantify the data incoherencies in a random sample of a hundred thousand clients that bought a service from the company in the year 2023, this includes new clients and clients that upgraded their service with the company. In addition to this, more analysis were done to further understand the complexity and seriousness of the problem.
4. Evaluation: In the evaluation step, a simulation is going to quantify what impact a single source of this particular data would have on the revenue of the company using past results from the main monthly digital company of the company that focuses on service upgrade and extension of the bidding period.

5. Conclusion: In the last step, an observation was made to further understand how effective the solution is and what limitations can come from this solution.

In addition to this, a qualitative approach was chosen to conduct this thesis. Conboy, K., Fitzgerald, G., & Mathiassen, L. (2012) state that IS qualitative research aims to empirically investigate a variety of phenomena concerning IS through qualitative data from a variety of sources, such as interviews, observations, design efforts, interventions, and archival materials. In this case data collection and a focus group is used to develop the investigation. According to Castellan, C. M. (2010) a qualitative study is usually conducted to gain understanding of a situation and the purpose of this thesis is to create awareness on the impact that data governance and IS management have on a digitally driven company. This dissertation sought to collect data from different teams and then interrogate different employees on the current state of the data delivered and the IS in a holistic manner. This method will then assist on the creation on a proposal to resolve the problems identified.

4. AVAILABLE DATA

Within company Z, the B2C (Business-to-Consumer) segment is endowed with a wealth of diverse and comprehensive customer data, essential for driving operational efficiency, enhancing customer experiences, and supporting marketing campaigns. The company adeptly manages an array of data, including client identifiers, portfolio details, RGPD-compliant information, billing specifics, geographical data, mobile usage statistics, multi-channel campaign insights (SMS, telemarketing, email, push notification, IVR), sales metrics, customer service interactions, and retail data related to company Z stores. Overall, the information available can be divided into 10 different domains or groups:

1. **Client Identifiers:** This group includes information that is able to identify individually each client. PersonId and Customer Account (CA) are a combination of numbers that identifies one single person, but a person can have more than one service with Company Z, then it's important to differentiate them with the Service Account (SA) that identifies each service. APPID is a key that is related to the CA, but it identifies clients that have an active account in the APP of the company. Finally, other common identifiers are used, like phone number, email address, fiscal number, and name. Client identifiers serve as fundamental markers crucial for businesses, particularly in the realm of digital strategy. These unique identifiers, such as account numbers or email addresses, play a pivotal role in accurately recognizing and understanding individual clients within a company's system. In essence, the accurate identification of clients is vital for the success of a digital strategy, enabling personalized experiences, targeted marketing, efficient customer service, and insightful data analysis—all essential components for achieving business goals and establishing robust customer relationships in the digital landscape.

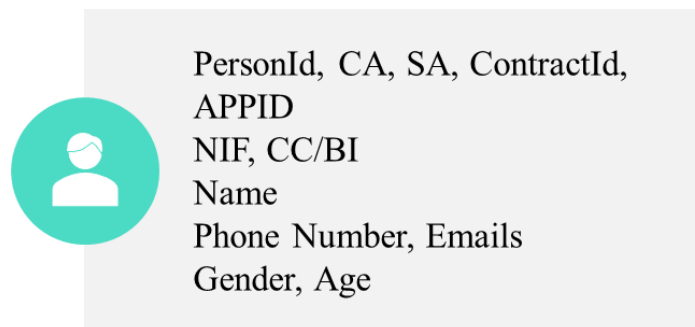


Figure 2 - Examples of client Identifiers

2. **GDPR:** This group refers to data necessary to comply with the General Data Protection Regulation, in order to respect client's privacy. Marketing Flags tell the organization if the client accepted or refused to receive marketing campaigns from different communication channels (SMS, MMS, Email, Telemarketing, and Push Notification). Company Z also needs to inform if it is keeping track of mobile data, that can be used for personalized advertising and needs to know if the client authorized the usage of their personal information. Respecting GDPR data, including marketing flags and authorization parameters, is not only essential for legal compliance but also critical for building trust and maintaining positive relationships with

customers. By aligning digital strategies with these parameters, businesses can deliver personalized, targeted communications and services while ensuring full compliance with data protection laws, thereby fostering stronger and more respectful connections with their clientele.



Figure 3 - Examples of GDPR data

3. Portfolio Data: This group encompasses all information about the services that clients have with the company. Understanding portfolio data is crucial not only for providing personalized and tailored services but also for driving revenue growth. By leveraging this information, businesses can enhance customer satisfaction, foster engagement, and drive incremental revenue by suggesting additional services or equipment that complement a client's existing portfolio. Due to the large volume of different services, three sub-groups were formed:
 - a. Essential services and premiums: All the information about the client's packages (Mobile, Internet, TV) is in this sub-group. Includes add-ons like paid channels, increments in internet velocity, etc.
 - b. Profile: Data used for profiling. This includes information about the segment, type of package, revenue, binding period, type of TV, nationality, etc.
 - c. Equipment: Contains information about the equipment of the client. What TV box does the client use? What is his internet router? Does he have a smartphone? If so, what smartphone is?

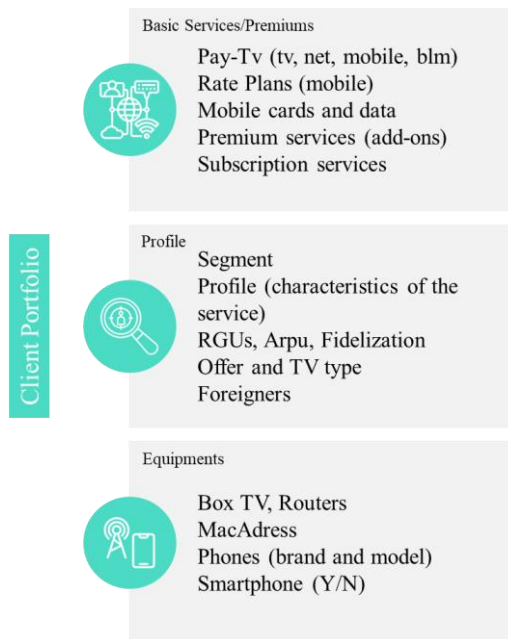


Figure 4 - Examples of Portfolio data

4. **Billing Data:** This group contains the data necessary for billing purposes. All clients have a BillingAccountID, an identifier of their invoices, belong to a billing cycle, have the choice to receive their invoice via email or letter, and form of payment. Also contains debt and payment accords control, in order to monitor clients who don't pay on time or don't pay at all. Understanding billing data is crucial for maintaining a healthy financial relationship with clients while also ensuring the company's financial stability. By leveraging this information, businesses can better manage financial interactions, support debt resolution, and maintain positive customer relationships, ultimately leading to a more sustainable and trustworthy partnership with their clientele.

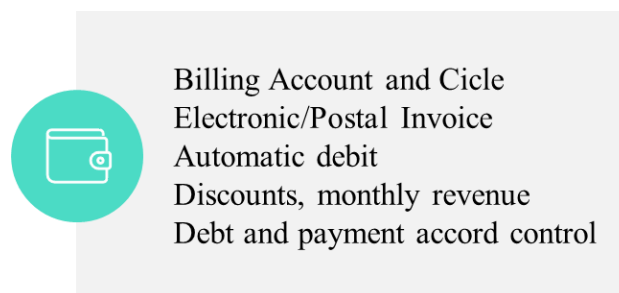


Figure 5 -Examples of Billing data

5. **Geographical Data:** Geographical information includes the client's addresses and coordinates, which district and municipality they live and what service is available in their location (fiber or satellite). Also contains information regarding where competition is fiercer, and where the organization's stores and cinemas are situated. Understanding geographical data is crucial for

businesses to optimize their market presence, identify growth opportunities, and tailor services to meet the diverse needs of customers in different regions. By leveraging this information, businesses can strategize their expansion plans, fine-tune marketing efforts, and identify areas where a more significant presence is needed, ultimately ensuring a more strategic and efficient approach to serving a diverse customer base.

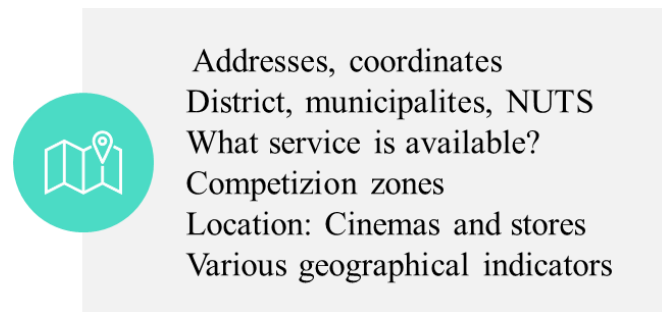


Figure 6 - Examples of Geographical data

6. Utilization Data: This information gives us insights into what our clients use and if what they are using is functioning well. For example, audiometry is used to understand what our clients are viewing. This is very useful for marketing campaigns, trying to sell a product that the client is probably interested in. Another example is mobile data usage, because if a client is constantly using the limit of mobile data, then it's useful for us as a company to sell him more. This is a view from a commercial standpoint, but there are other perspectives, like service quality. Our products automatically produce a report to alarm malfunctions to the company in order to contact those clients and propose a solution.

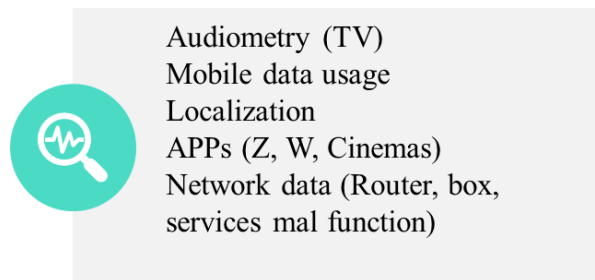


Figure 7 - Examples of Utilization data

7. Campaigns Data: This group gathers all information about commercial campaigns in the company (Digital, Outbound, Push Notification, etc). Retaining data from marketing campaigns is fundamental for telecommunications companies. It not only helps in evaluating performance, understanding audiences, and improving ROI but also supports informed decision-making and compliance with regulations. The insights gained from this data form the

bedrock for evolving strategies and ensuring a more targeted and effective approach to engaging customers and achieving business objectives.

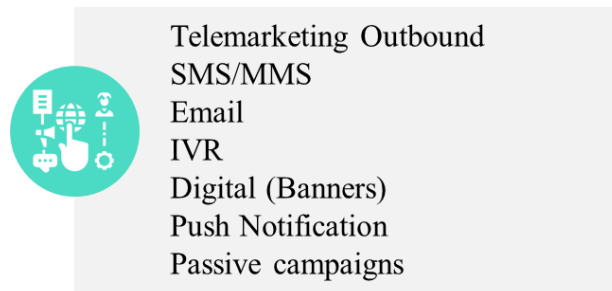


Figure 8 - Examples of Campaigns data

8. Sales Data: Sales data is the lifeblood of telecommunications companies, providing a comprehensive understanding of business performance, customer behavior, and logistical operations. Within this realm, the amalgamation of logistics, seller information, orders, and equipment details forms a foundation crucial for strategic decision-making and operational efficiency. It facilitates crucial decision-making, drives operational efficiency, and empowers the delivery of personalized services, ensuring sustained growth and customer satisfaction within the competitive telecommunications landscape.



Figure 9 - Examples of Sales data

9. Client Service Data: This department's work is based on metrics, so they gather a lot of data from their teams and the clients. Client service data, encompassing complaints, unresolved issues, call statistics (repetitions, durations, subject matter), and the need for technical issue resolution, forms an invaluable repository for telecommunications companies. The aggregation and analysis of this client service data are pivotal for telecommunications companies. It forms the basis for targeted improvements in service quality, resource allocation, and employee training, ultimately leading to a higher standard of customer service and increased customer satisfaction. By addressing complaints, resolving issues efficiently, and leveraging insights from call data, telecommunications companies can continuously evolve their services to better cater to the needs and expectations of their customer base.



Clients complaints
Unresolved clients complaints
Calls data (repetitions, duration, subject)
Ots'

Figure 10 - Examples of Client Service data

10. Retail Data: Company Z also deals with physical stores and cinemas. This data, encompassing interactions with clients, time metrics, resolution rates, tickets sold, movies played, sessions, location, and service quality, is pivotal for assessing and enhancing the customer experience in both these domains. Insights derived from this data support the identification of customer preferences, operational improvements, and strategic decision-making. By leveraging this information, telecommunications companies and cinema operators can adapt their services, optimize operations, and continuously refine their offerings to meet the evolving needs and expectations of their customer base.



Stores: Interactions, time metrics, resolution rates
Cinemas: Dashboard (movies, cinemas, sessions)

Figure 11 - Examples of Retail data

4.1. CURRENT SYSTEMS

There currently more than three hundred active information systems in company Z, which is indicative of the challenges that it faces. Operating and extensive portfolio of information systems can be challenging due to the complexity and scale of its technology infrastructure. Given this, the systems were organized into thirteen domains, as showed in Figure 12, in order to facilitate interpretation.

Having 300 plus IS, creates enormous challenges, such as:

1. Integration Complexity: Managing integration among 300 systems can be a significant challenge. Ensuring seamless communication and data flow between these systems requires complex middleware or integration platforms.
2. Data Silos and Duplication: With numerous systems, the risk of data silos (isolated data in various systems) increases, leading to duplication, inconsistencies, and difficulties in obtaining a single source of truth. Even more in a case where the architecture was formed based on a merger of two large companies.

3. **Resource Drain:** Maintaining and managing such a vast number of systems can be resource-intensive in terms of human resources, time, and budget.
4. **Interoperability Issues:** Different systems might use diverse technologies, making interoperability a major issue. Standardization and ensuring compatibility among systems become challenging.
5. **Security and Compliance:** Managing security across many systems poses a significant challenge. Ensuring compliance with data protection regulations and implementing robust security measures across all systems becomes more complex.
6. **User Adoption and Training:** Introducing and training employees on multiple systems might lead to resistance, lower adoption rates, and increased training costs.
7. **Decision-making Complexity:** Extracting meaningful insights from disparate systems for decision-making becomes intricate, potentially affecting the agility and responsiveness of the organization.

While having numerous information systems offers opportunities for specialization, departmental efficiency, and innovation, managing such a diverse landscape of systems requires careful planning, strategic investment, and continuous optimization to mitigate challenges and leverage the advantages they bring.

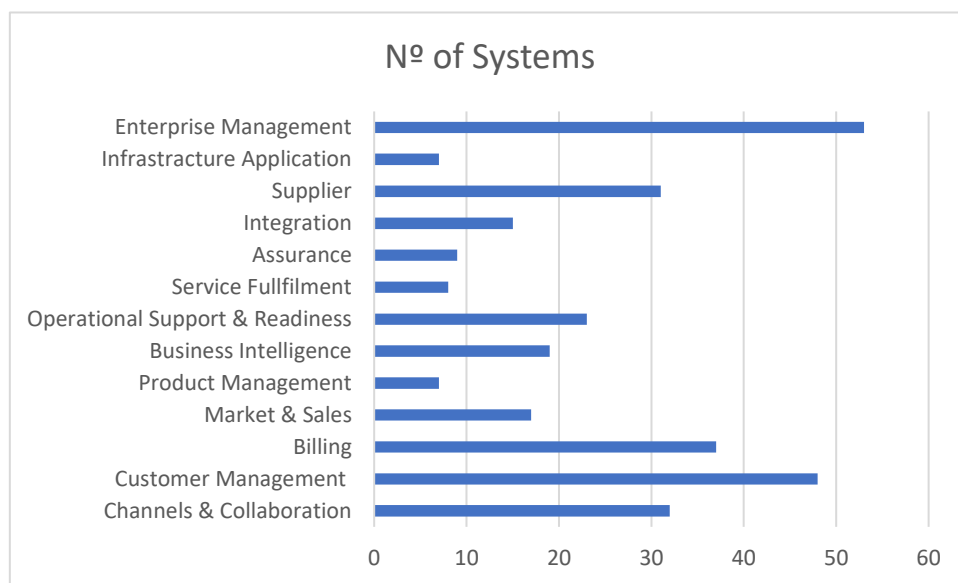


Figure 12 - Current Systems by Domain

5. FIELD WORK

5.1. DATA ANALYSIS

As mentioned in the methodology section, hundred thousand serviceaccounts were random picked from the company's database. After that, all serviceaccounts were sent to three different departments and was asked to retrieve the annual revenue per user (arpu) from each serviceaccocunt to understand if the data is coherent.

Firstly, is imperial to understand what *arpu* means. In simple terms, it means how much the client pays for the service. However, the devil is in the details and the first question received from the departments was: "What *arpu* do you want?". This creates additional questions: Is this term broader than it should be? Is it making more difficult the communication between each department? Probably yes because everyone seems to have a different way of defining what *arpu* is. Despite this difficulty, a consensus was reached and *arpu* was defined as what clients paid each month from the package they had, excluding premiums, exceeded limits in data usage and calls made, etc. Given this definition is expected the value to be static each month unless the client alters its package. This is important because, for example, the billing department works with a dynamic *arpu*, meaning that for them what matters is the total paid from the client each month.

After receiving the data the first problem arose. With the help of Figure 13, it is clear that the commercial department was the only one capable of accessing the data of the hundred thousand serviceaccounts, meaning that the departments Experience and Value and Customer Relationship Management couldn't access vital information about their clients. As we can see in Figure 13, CRM lacked the data from 12 clients and EV lacked the data of 217 clients. Despite representing less than 1% in both cases it is still an important issue to consider because it is a piece of vital information for business decisions.

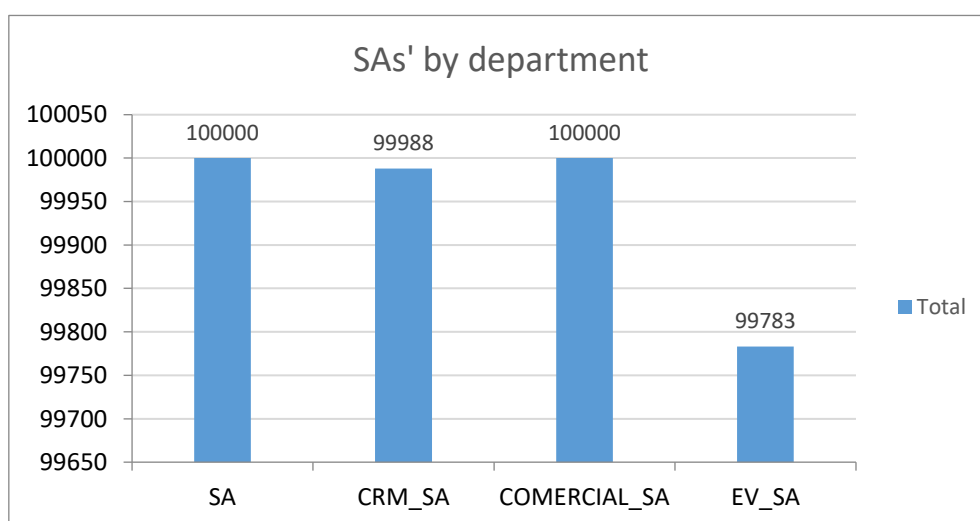


Figure 13 - SAs' by department

Having Figure 13 in consideration, it illustrates discrepancies in ARPU data among different departments—EV, CRM, and Commercial—pertaining to client payments at the end of the month. Among the 100,000 clients analyzed:

In 714 cases, all three departments showcased distinct ARPU values for the same clients, signifying a complete lack of consistency across the board.

Furthermore, there were 18,032 instances where the ARPU values between the EV and Commercial departments didn't align. This inconsistency might arise from differing methodologies or data sources employed by these departments to assess client revenue.

Similarly, the ARPU values between the EV and CRM departments didn't match in 4,955 cases, indicating a disparity in perspectives regarding client revenue, possibly due to using dissimilar metrics or data sources.

Moreover, 18,036 instances showed discrepancies between the ARPU values of the Commercial and CRM departments, highlighting conflicting assessments of client revenue between these two departments.

These inconsistencies underscore potential issues in data reconciliation, methodological differences, or variations in data sources. Addressing these discrepancies is imperative to ensure data accuracy, align departmental assessments, and establish a unified understanding of client revenue across the organization. Investigating the root causes comprehensively and implementing standardized data collection and reporting practices will be pivotal in mitigating such discrepancies in the future.

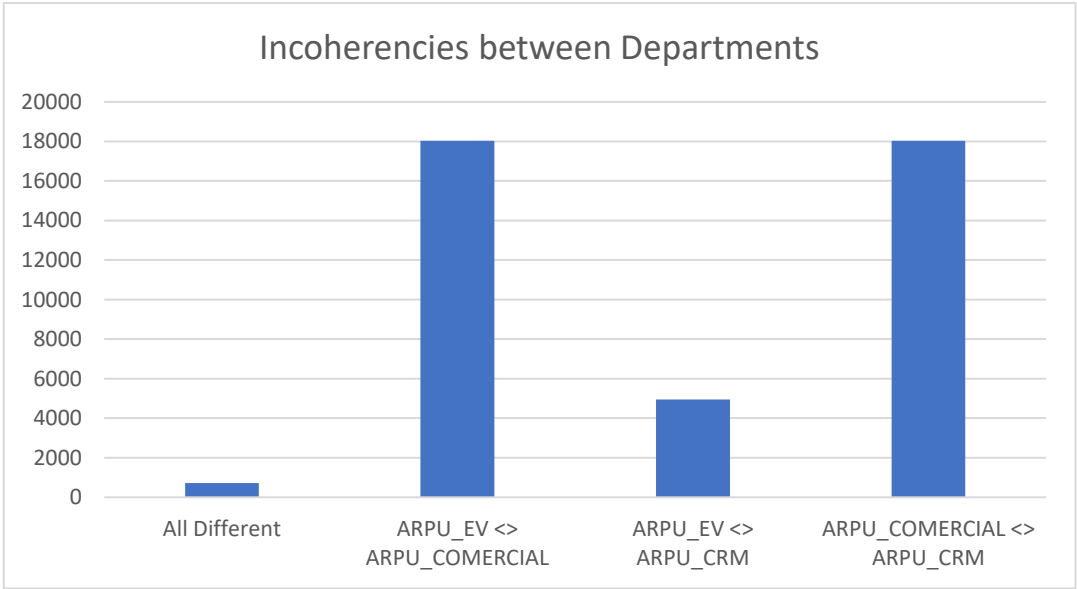


Figure 14 - Incoherencies between departments

With the help of Figure 15 we noticed that we have 20628 instances of incoherent data, this translates to more than 20.6% of values that we are not sure of its veracity. Having this value in mind it is safe to conclude that for every hundred thousand clients that we communicate with digitally to try to sell an upgrade of the service we do not actually know how much they pay. This creates a great obstacle on digital sales teams because it limits them on their personalization, since they cannot communicate the actual price the client pays. So, it is impossible to create a personalized message in that area, even in the cases of a down sell. Furthermore, this creates a risk for the company in term of service quality and its prejudicial for the image the company wants to transmit.

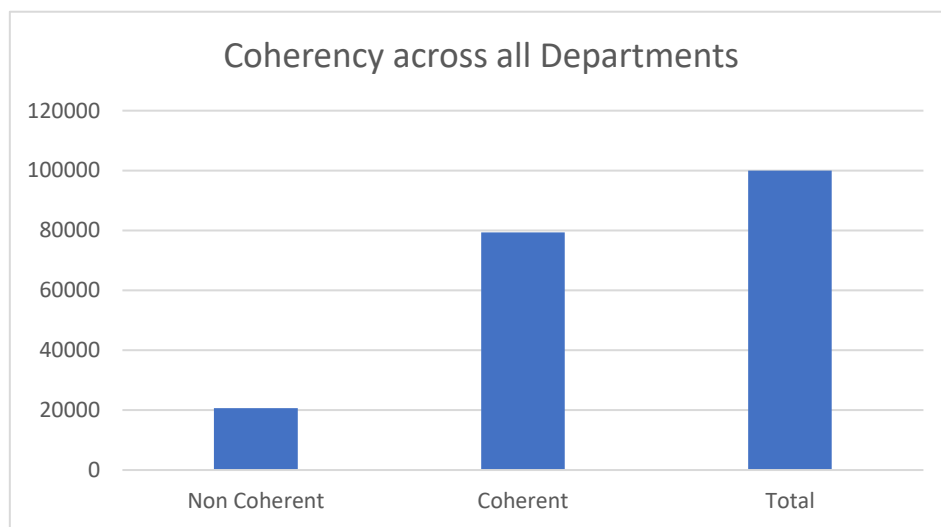


Figure 15 - Data coherency between departments

5.2. DIGITAL CAMPAIGN USE CASE

One of the most important digital campaigns at company Z, is the one that has the objective of increasing the bidding period of the clients while selling an upgrade of the service. This campaign is highly personalized with more than 30 customized parameters, but for the sake of the analysis only the *arpu* is considered.

This campaign shows the client a description of their current package with the total amount of euros he pays each month for the service (*arpu*), then shows what package the company Z is trying to sell, and finally the amount he is going to pay less or more by making a simple mathematical operation (sale *arpu* – current *arpu*). This simple personalization depends completely on the quality of the data in order to give the client the accurate values and to maximize the conversion rate.

Taking chapter 5.1 in consideration, it is expected that for every hundred thousand clients, roughly twenty thousand have inaccurate data regarding the *arpu*, representing twenty thousand clients or 20%, assuming that the veracity of data can only be guaranteed if there are no differences across departments.

Taking in account that this campaign affects 200.000 clients each month, it is safe to assume that 40.000 clients have inaccurate *arpu*. Knowing that the conversion rate E2E is approximately 0.9%, we can predict how many sales each month company Z is losing due to data incoherencies.

$200.000 \times 0,2 = 40.000$ and $40.000 \times 0,009 = 360$ sales.

The value of each sale depends on from client to client, but the average *arpu* for each client is roughly 40 euros. When the company makes a sale, it automatically restarts the bidding period of 24 months, meaning that is certain that client will stay with the company for the next 2 years. Having this data into account, it is possible to calculate the potential loss of revenue that the company is facing in this campaign:

$40 \times 24 \times 360 = 207.360$ euros each month.

This Use Case helps understand the importance of data coherency and quality. Only one parameter regarding client information, can have a tremendous impact in the revenue of the company. For example, this campaign is not making the company lose money directly but is certainly far from its true potential.

5.3. FOCUS GROUP

A focus group is a research method that brings together a small group of people to answer question about a specific problem. Three different managers were gathered in a casual conversation to address the problem identified by this thesis and five questions were asked:

Question: What is your opinion on the current status of the information systems architecture? Knowing that exist more than 300 active systems.

Answer: Overall the answer was positive. Since they know other realities, it is considered to be a good architecture. Despite this, three main factors were mentioned:

1. Duplicate Systems: Since this company was the result of a merger between two large companies, a great number of information systems were duplicated since they do the same thing but for different realities. It's imperative to eliminate obsolete systems and to create more adequate systems for the new reality of the company.
2. Out of Date Systems: During a rapidly evolving world, the needs of businesses that want to continue to be competitive logically increase. Unfortunately, some IS are not capable of satisfying all the needs required.
3. Follow Tendencies: With the necessity of following more current technology tendencies, such as Cloud Solutions, it is important to make sure that the IS can stay in the race.
4. Lack of NRT systems: Globalization drives digitalization, every interaction the client has with the company should trigger an immediate reaction. From a marketing perspective it is very important to react on time to customer actions, so information fluidity and dissemination it's crucial.

Question: Have teams been suffering from data incoherencies when projects involve different departments?

Yes, three projects were especially mentioned:

1. Prices Increase: This project, as the name suggests, deals with increasing the prices of the company's services. An extremely complex problem that works with thousands of different combinations of products and dozens of requirements. One problem that teams often encounter is deciding which Service Account is the master of the Client Account (meaning, is the main service of the client), since a CA can have multiple SAs'. This happens because different systems have different master SAs', which transforms this in an intense manual work.
2. Cece Best Offer: This was a legal project that communicated with clients according to their bidding period. Since this was a legal requirement, the fines were high if the company didn't entirely comply. The problem identified was that the bidding periods were different across departments, so the company was not sure with which to communicate. The decision was made to communicate with all of them, in order to comply with the law. But without a doubt there were excessive messages sent to the clients.
3. Sales Indicators across departments: Another use case found was the sales indicator across departments. What was found is that different departments had different number of total sales per month, which reveals another case of data incoherency.

In conclusion, there is a lot of time wasted, money spent, and poor decision made due to data incoherencies that are linked with a poor governance of data and IS.

Question: What are the main impediments to reach an excellent data quality in the company?

Answer: In order to answer this question, four main points were mentioned:

1. Dimension: Since we are talking about one of the biggest companies in Portugal, there is a lot more to manage. The probability of having data problems, it is directly proportional to size of the company. There are more processes, more clients to manage and more systems. This causes a lot of pressure to ensure an efficient management.

2. Antiquity: It's a company with decades of history, and if the systems are not updated the methodologies have less data validations. More recent systems have a higher chance of satisfying the needs of the company.
3. Budget: Not everything that is idealized from the IT and Business departments are necessarily a priority for the company. Since the budget is logically limited, decisions need to be made.
4. Quick Actions: The implementation of new systems and new methodologies is a slow process, it needs to be well thought and carefully planned. Despite that, sometimes the deadlines are very demanding, which prevents that.

Question: Ideally, what can be made to resolve the problems identified in your teams?

Answer: Three main points were mentioned in order to ideally solve the existing problems:

1. Having a master system for each type of solution, organizing in domains. Then it's important to have a structured interconnection between systems and subsystems, planned from zero in order to have a personalized plan.
2. A concise data architecture with good data flows to prevent data loss between systems. Additionally, data entry policy and rules should be precise and well known to increase data quality (for example, do not permit fiscal numbers that don't comply with a certain algorithm).
3. Clear governance in each domain of systems, to have a clear owner and have a clear responsibility of the data.
4. An evolution policy is extremely important, to maintain the systems updated. For this is important that IT and Business departments work as one, to understand the needs and difficulties of each other.

Question: Data Mesh is an innovative architectural framework, do you think it's appropriate for the problems mentioned above?

Answer: Data Mesh can be explained in a very simple way. Since early age, we are taught to group objects, we look at them and see similarities, this makes us believe that they belong to the same group. This is the great purpose of Data Mesh, is to group similar data in domains and have a clear responsible for the quality and evolution of each domain. Additionally, other thing we are taught as children, sharing the objects (in this case, data) it is important and with Data Mesh we ideally know where to look for the data and information needed.

In the past years, the necessity for data products and agility of processes exploded, mainly due to new advanced analytics use cases. It was noticed by the company that the old school model where there was a great dependence of the IT department to inject and share data was slowing the company which creates a risk of not reaching the company's goals. Having this in mind, a new framework needs to be implemented and Data Mesh is for sure an answer to our problems.

In conclusion, the main aspect that it's important to update is the agility, because agility means a greater efficiency. This will release the burden of the IT department and is going to decentralized data ownership.

6. PROPOSAL

In this chapter the objective is to propose a solution for the problem identified in this dissertation. Despite using only the example of “arpu”, there are thousands of other domains that are equally affected from a poor data and information systems governance. Having this in mind, a proposal is presented that would mitigate or eliminate the identified problem.

6.1. UNDERSTANDING DATA MESH

In the realm of contemporary data management, the concept of Data Mesh has emerged as an innovative architectural framework. Derived from insights and principles espoused by technology leaders like IBM and Amazon, Data Mesh redefines conventional approaches to data organization, governance, and accessibility within complex organizational structures.

At its core, a Data Mesh embodies a decentralized approach to data management, challenging the traditional centralized data infrastructures prevalent in many organizations. This tackles directly a problem at Company Z, where the ownership is not clear and falls mainly on the IT department. It's a structural paradigm that acknowledges the distributed nature of data sources across diverse business domains within an enterprise.

6.2. FUNDAMENTALS OF DATA MESH ARCHITECTURE

A Data Mesh addresses intricate data security challenges by advocating distributed, decentralized ownership across an organization's myriad data sources. These sources, often stemming from disparate lines of business, necessitate integration for effective analytics. The architectural framework of a Data Mesh strives to unite these dispersed data sources through a system of centrally managed data sharing protocols and governance guidelines.

One of the foundational pillars of a Data Mesh is its emphasis on empowering individual business functions with control over data access, usage protocols, and formats. This strategic shift ensures that domain experts retain autonomy over their data domains, enhancing their ability to curate meaningful data products within a decentralized governance framework. With this, the right teams will have ownership of their data and will be responsible on who accesses the data, preventing data breaches.

6.3. RESOLVING CHALLENGES THROUGH DATA MESH

Organizations grapple with various challenges in managing vast volumes of data. Traditional approaches relying on central data teams often encounter obstacles that impede agility and accuracy in data processing. Data Mesh intervenes to alleviate these challenges by addressing key pain points:

1. **Siloed Data Teams:** Traditional central data teams, albeit proficient in technical aspects, often lack domain-specific knowledge. This impedes their ability to cater to a diverse set of operational and analytical needs within different business units. This point addresses the

ideally solution mention in the focus group, where there should be a proximity between the IT and the business department, enhancing the quality of the shared data.

2. **Slow Responsiveness to Change:** Centralized data pipelines demand extensive modifications for even minor alterations, leading to sluggish responsiveness to evolving business requirements. Impediments like this is what causes teams to not reach their full potential, mainly in digital strategy where data quality is so important. This rigidity inhibits adaptability and agility.
3. **Reduced Accuracy:** Disconnected data flows between business units and central data teams often result in a lack of motivation to provide accurate and relevant data. This disconnect diminishes the quality and trustworthiness of available data assets. An example of that was seen in the analysis on 5.2. chapter, where various departments had different values for the same data.

Having these challenges in mind, implementing a Data Mesh architecture creates many benefits:

1. **Democratic Data Processing:** Empowering domain experts to create and manage data products within a decentralized governance framework accelerates data access, thereby enhancing overall business agility.
2. **Increased Flexibility:** By decentralizing data infrastructure, a Data Mesh mitigates operational bottlenecks and technical constraints, promoting flexibility and adaptability in data processing.
3. **Cost Efficiency:** Transitioning away from batch processing toward real-time data streaming optimizes resource allocation and storage costs, fostering improved budgeting and reduced operational expenses.
4. **Improved Data Discovery:** Preventing the formation of data silos and ensuring central governance of data assets enhances discoverability and accessibility across the organization.
5. **Strengthened Security and Compliance:** Enforcing data security policies and centralized monitoring mechanisms bolster data governance, ensuring adherence to compliance standards and enhancing overall data security.

6.4. USE CASES OF DATA MESH

The versatility of a Data Mesh extends across diverse big data use cases:

1. **Data Analytics:** Provisioning high-quality data from multiple business functions enables robust analytics, empowering teams to derive actionable insights and accelerate machine learning endeavors. Helping directly digital strategy and linked to that are digital campaigns that depend enormously on data quality to analyze performance.
2. **Customer Care:** Comprehensive customer data amalgamation facilitates efficient customer support and targeted marketing strategies, enhancing customer engagement and satisfaction.
3. **Regulatory Reporting:** Meeting regulatory requirements through centralized governance ensures timely and accurate reporting, benefiting both organizations and regulators alike.
4. **Third-party Data Integration:** Integrating external datasets within the Data Mesh architecture enables seamless blending with internal data, fostering consistency and reliability in analytics.

6.5. PRINCIPLES OF DATA MESH ARCHITECTURE

To successfully adopt a Data Mesh paradigm, organizations must adhere to the following principles:

1. **Distributed Domain-Driven Architecture:** Organizing data management around business domains ensures domain teams take ownership of collecting, transforming, and providing data relevant to their functions. Enhancing domain to belong to the right teams, that are familiarized with the data in question.
2. **Data as a Product:** Treating data assets as products and adhering to principles of discoverability, addressability, trustworthiness, and self-description enhances data usability and relevance.
3. **Self-Serve Data Infrastructure:** Empowering business units to manage their data pipelines while leveraging self-serve infrastructure streamlines data processing and reduces dependencies on centralized teams.
4. **Federated Data Governance:** Balancing centralized governance standards with domain-level autonomy ensures adherence to global policies while allowing flexibility in implementation within domains.

6.6. IMPLEMENTATION STRATEGIES FOR BUILDING DATA MESH

The journey toward embracing a Data Mesh involves meticulous planning and strategic execution:

1. **Catalog Existing Data:** Conduct an exhaustive assessment of existing data, identifying relevant business domains and establishing harmonization rules for effective data correlation. This was done on chapter 4, where it is possible to realize the dimension of Company Z and the current groups that data can be stored in.
2. **Implement Governance Policies:** Define robust governance policies encompassing reporting, authentication, access controls, and compliance standards aligned with a federated data governance model.
3. **Build Self-Serve Infrastructure:** Develop a versatile self-serve data platform that simplifies data processing, storage, encryption, governance, and access control across domains.
4. **Choose Apt Technologies:** Leverage existing storage systems or embrace cloud-native technologies, ensuring they align with the scalability and interoperability requirements of a Data Mesh architecture.
5. **Cultural Shift:** Drive organizational change by fostering a culture that prioritizes data discovery, real-time processing, distributed data ownership, and domain-centric decision-making.

6.7. CONCLUSION AND DISCLAIMER

The advent of Data Mesh signifies a pivotal shift in the landscape of data management strategies. By prioritizing decentralized ownership, democratized data processing, and enhanced governance, organizations can leverage this architectural framework to unravel the complexities of data management, fostering agility, efficiency, and innovation.

The information presented above draws insights from industry leaders like IBM and Amazon and aims to provide an informative overview of Data Mesh architecture. The information is based on information present on their websites.

In conclusion there are five main principles that helps companies to adopt Data Mesh:

1. Ownership: Decentralization and distribution of responsibility of data to business departments.
2. Autonomy: Autonomous E2E developments made by each data teams.
3. Governance: Transversal governance with standards and tools that guarantee a consistent way of working.
4. Data as a product: Data is seen as a strategic asset with a focus on value creation to the clients.
5. Consistent Data: Marketplace of self-service data that encompasses sharing and reutilization of data products with common language.

Following these five principles, will help the company to adopt Data mesh and mitigate or eliminate data incoherencies or data loss between departments, like is displayed on chapter 5.1. Taking as an example the digital campaign present on chapter 5.1. it would help the responsible team to reach the full potential of the campaign, since the data “arpu” would have a single source of truth with a clear owner.

7. CONCLUSIONS, FUTURE WORKS AND LIMITATIONS

The meticulous examination conducted within this specific telecommunications company has unveiled a profound concern: the severe implications of data incoherency on the efficacy of digital campaigns and decision-making processes. An in-depth analysis of a dedicated use case unequivocally exposes a staggering reality—nearly 20% of critical data, the ARPU, suffers from glaring inconsistencies across various departments, compromising its reliability. This revelation serves as a poignant illustration of the detrimental repercussions resulting from the deeply entrenched data incoherency within the company's existing information systems framework.

Crucially, this company heavily relies on a wealth of customer data, forming the cornerstone of strategic decision-making processes. The reliability and coherency of this data directly influence the efficacy of marketing campaigns, service offerings, and overall customer experiences. The pervasive nature of data incoherency not only impedes the performance of digital campaigns but also casts shadows of doubt upon the validity of critical decisions made based on this data repository.

The prevailing information systems architecture within this company stands as a relic of a bygone era, inhibiting the seamless exchange and uniformity of data across its multifaceted departments. The urgent need for a transformative overhaul in this framework is undeniably palpable, and the company stands on the cusp of embracing a revolutionary shift. The proposition of adopting a Data Mesh architecture emerges as a beacon of hope, promising democratization of data access, improved domain-based organization, and the establishment of accountable domain owners vested with ensuring data integrity within their respective spheres.

Looking ahead, the quantification of potential revenue loss stemming from data-related issues within this specific company becomes imperative. This thesis underlines the indispensable nature of this quantification process in grasping the gravity of data incoherency. As this company's decisions heavily rely on customer data, future studies must consistently integrate methodologies to quantify revenue loss. This approach will serve as a stark reminder of the indispensability of maintaining high-quality, coherent data, forming the bedrock of informed decision-making within this company's operations.

While this study delved deep into the intricate workings of this telecommunications company, several limitations hindered the research process. The reluctance of certain teams to divulge specific information presented a barrier, limiting a comprehensive understanding of certain data processes. Moreover, the swift pace of the company's operations posed challenges, hampering the allocation of sufficient time for comprehensive discussions and data sharing, thereby constraining the depth of insights gleaned.

In summary, this investigation illuminates the profound impacts of data incoherency on digital campaign performance and decision-making within this specific telecommunications company. It underscores the urgency for transformative change within the company's information systems framework. Addressing these limitations and emphasizing revenue loss quantification in future research endeavors will continue propelling the imperative shift towards efficient data governance and systems management within this company's unique context.

8. BIBLIOGRAPHICAL REFERENCES

Alotaibi, M. Z., Alotibi, M. F., & Zraqat, O. M. (2021). The impact of information technology governance in reducing cloud accounting information systems risks in telecommunications companies in the state of Kuwait. *Modern Applied Science*, 15(1), 143-151.

Beynon-Davies, P. (2019). *Business information systems*. Bloomsbury Publishing.

Boeckhout, M., Zielhuis, G. A., & Bredenoord, A. L. (2018). The FAIR guiding principles for data stewardship: fair enough?. *European journal of human genetics*, 26(7), 931-936.

Couto, Julia & Damasio, Juliana & Bordini, Rafael & Ruiz, Duncan. (2022). New Trends in Big Data Profiling. 10.1007/978-3-031-10461-9_55

Gilder, G. (1990). "Microcosm: The Quantum Revolution in Economics and Technology." Simon and Schuster.

Goedegebuure, A., Kumara, I., Driessen, S., Di Nucci, D., Monsieur, G., Heuvel, W. J. V. D., & Tamburri, D. A. (2023). Data Mesh: a Systematic Gray Literature Review. arXiv preprint arXiv:2304.01062.

Gupta, B. (2014). "Data Governance: Creating Value from Information Assets." Technics Publications.

Kimball, R., Ross, M., & Thornthwaite, W. (2008). "The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data." Wiley.

Laudon, K. C., & Laudon, J. P. (2021). "Management Information Systems: Managing the Digital Firm." Pearson.

Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2018). "AIMQ: A methodology for information quality assessment." *Information & Management*, 45(2), 103-114.

Moore, G. E. (1965). "Cramming more components onto integrated circuits." *Electronics*, 38(8).

Rainer, R. K., & Prince, B. (2022). *Introduction to information systems: Supporting and transforming business*. John Wiley & Sons.

Redman, T. C. (2008). "Data Driven: Profiting from Your Most Important Business Asset." Harvard Business Press.

Roberts, M., & Reid, G. C. (2016). "Data Governance: How to Design, Deploy and Sustain an Effective Data Governance Program." Technics Publications.

Tanenbaum, A. S., & van Steen, M. (2006). "Distributed Systems: Principles and Paradigms." Prentice-Hall.

Turban, E., Pollard, C., & Wood, G. (2018). *Information technology for management: On-demand strategies for performance, growth and sustainability*. John Wiley & Sons.

Wang, R. Y., & Strong, D. M. (1996). "Beyond accuracy: What data quality means to data consumers." *Journal of Management Information Systems*, 12(4), 5-33.

ExoPlatform. 5 of the Biggest Information Technology Failures and Scares. (2017).
<https://www.exoplatform.com/blog/5-of-the-biggest-information-technology-failures-and-scares/>

Venable, J. R., Pries-Heje, J., & Baskerville, R. L. (2017). Choosing a design science research methodology.

Castellan, C. M. (2010). Quantitative and qualitative research: A view for clarity. *International journal of education*, 2(2), 1.

Conboy, K., Fitzgerald, G., & Mathiassen, L. (2012). Qualitative methods research in information systems: motivations, themes, and contributions. *European Journal of Information Systems*, 21(2), 113-118.

Dehgani, Z. (2022). *Data Mesh*. O'Reilly Media, Inc., 386 pp.

Amazon AWS. What is Data Mesh? <https://aws.amazon.com/what-is/data-mesh/>

IBM. What is Data Mesh? <https://www.ibm.com/topics/data-mesh>

