UNIVERSIDADE CATÓLICA PORTUGUESA

A Corporate Metadata Dashboard for the BI & Analytics Department of EDP Comercial

Curricular Internship in EDP Comercial

Gonçalo Meira da Cunha Pereira

Católica Porto Business School

2022

UNIVERSIDADE CATÓLICA PORTUGUESA

A Corporate Metadata Dashboard for the BI & Analytics Department of EDP Comercial

Curricular Internship in EDP Comercial

Presented to Universidade Católica Portuguesa

to Obtain Master Degree in Management

by

Gonçalo Meira da Cunha Pereira

Under orientation of

Prof. Drº Pedro Duarte Silva

Católica Porto Business School

2022

Acknowledgements

The first word of my thesis goes to the victims of the Ukrainian war and their families. Amazing how in 2022, two countries so close live such different realities: I have the opportunity to study and pursue my dreams while others like me, there, take up arms to fight for the freedom of their country.

My second acknowledgement goes to my parents, of course. Father and Mother, thank you for being my support not only these two years, but throughout my life. Always with a sweet word for me when everything becomes more difficult. To them, you deserve all the best!

To my sister Inês, for teaching me what resilience is through example, and never settling for less than I deserve. To my brother Sérgio, despite the distance, for being the most genuine, simple and pure person I know, and for guiding me through his advice.

To my supervisor, Pedro Duarte Silva, for his availability, wisdom and for all the help to carry out this project.

To my mentor and friend, Francesco Costigliola, for all the conversations during this year, that guided me to the business world, and in the culmination of the mentoring process, for the opportunity to work with him and with the team of EDP Comercial in these last 6 months, it was really enriching and an experience that I will keep for life.

Finally, I would like to thank my girlfriend, Débora, for all the support and love always. To my little Bali, who whenever I remember this master's degree, she is the one I will think of first and I will surely smile. To my best friends, Juca and Durães, for their constant friendship and for always accompanying and supporting me at all stages of my life. And to Zé and Pedro, who were the surprise of this master's degree, two people I met and who I will certainly take with me from now on.

Abstract

Companies that want to remain competitive in the market need to constantly update themselves. With the development of technology, it is now possible to collect more data about the business and companies have made large investments in this collection, processing and analysis of data.

One of the business difficulties is the organization of these data, which is now in increasing quantities. Thus, the challenge of this project arose: the organization of large amounts of data in a corporate metadata dashboard to improve the internal management of a department in relation to its developed projects.

The DOMO Business Intelligence software was used to facilitate the processing of information and allow the visualization of data in a more accessible and useful method.

The project results were positive, bringing to the company in question, EDP Comercial, an improvement in the department's Internal Management, with advantages in terms of data visualization and in economic terms.

Keywords: Data; Dashboard; Business Intelligence; Metadata; Data Visualization

Words: 7840

Resumo

As empresas que se querem manter competitivas no mercado precisam de se atualizar constantemente. Com o desenvolvimento da tecnologia, é agora possível recolher mais dados sobre o negócio e as empresas têm feito grandes investimentos nessa recolha, processamento e análise desses mesmos dados.

Uma das dificuldades do negócio é a organização desses dados, que agora são em cada vez maiores quantidades. Assim, surgiu o desafio deste projeto: a organização de grandes quantidades de dados em um dashboard de metadados corporativo para melhorar a gestão interna de um departamento em relação aos seus projetos já desenvolvidos.

O software DOMO Business Intelligence foi utilizado para facilitar o processamento das informações e permitir a visualização dos dados de forma mais acessível e útil.

Os resultados do projeto foram positivos, trazendo para a empresa em causa, a EDP Comercial, uma melhoria na Gestão Interna do departamento, com vantagens em termos de visualização de dados e em termos económicos.

Palavras-chave: Dados; *Dashboard*; *Business Intelligence*; Metadados; Visualização de dados

Palavras: 7840

Table of Contents

List of Figures

1. Introduction

We live in the information era. It is said that we never had so much data being created at each second. This data is and it will be a differentiating factor for every company, independently from the industry they are in. Business competitiveness means that the company must invest a lot in the organization, so that resources are not wasted. In this information age, where a large amount of data is created at every moment, there is a need by the business structures that an investment to process and analyse data occurs, which has effectively been happening. In fact, an increasing focus on data analysis departments and greater employability in these areas is a reality, with even the “data scientist”, “data analyst” or “data engineer” being some of the most recruited professions in recent years. These professionals can analyse and create insights through data, leading companies to more thoughtful, fact-based decisions. This ever-expanding sector is now critical to keeping company directors or managers abreast of what is going on in the company. Executive decisions are the main component affecting the growth of the organization. While one right decision can make the business to reach the sky, one wrong decision can bring down the business.

Thus, in this sense, in this project: "Creation of a Repository/Dashboard with all Metadata and Metainformation on the developments in the BI & Analytics department" the role of data in companies will be explained, specifically in the case of EDP Comercial. The project path will be concise, starting in Topic 2 with an explanation of the concepts with the help of literature review. First, in 2.1, we will address the meaning and contextualization of Data and the purpose of its use in the organizational environment. Then, in 2.2, we go through the concept of Data Visualization and the powerful tool, the Dashboard, revealing its meaning and how this type of tool evolved, which is increasingly used by companies. And succinctly, in 2.3, we will also cover important topics related to the project such as the purpose of Business Intelligence in organizations and, in 2.4, the hard to define Metadata concept,

revealing its meaning and how this type of data emerged. From this brief contextualization of the theme's concepts, we will go to the heart of the project, which is Topic 3. In 3.1, we start with a brief explanation of the company where the project was developed - EDP Comercial. After that, in 3.2, the problem encountered is exposed and what difficulties the company faced and currently found in internal management, that is, the reasons that led to the development of the project. After that moment, in 3.3, we will go through the stages of the project in detail, explaining how it was executed technically. Finally, in 3.4, we will see what results were found and if, through the project, it was possible to face some difficulties felt by the company before the development of the Metadata Repository/Dashboard. After explaining the Project, we move on to Topic 4, which will be the main conclusions of the project developed and in Topic 5 we will have the bibliographic references that support the document.

The overall objective of the project is generally to improve internal management through the preservation of digital records improving their location, understanding and management and in the end, to verify if this project was useful in practical and financial terms for the organization.

2. Literature review

2.1 Data Contextualization

Data is a collection of facts, observations, or even raw elements that are not organized or processed in any particular manner and that contain no discernible meaning. Data is the basic unit recording a thing, an event, an activity, or a transaction. A single piece of data is insignificant on its own. Data does not become significant until it is interpreted and thus comes to bear meaning. (Monino, 2021)

Despite that, knowledge or information about data are currently assumed as critical success factors for productivity, competitiveness and value creation. Data have always

had a strategic value, but today, with the magnitude of data available, and our ability to treat them, it has become a new form of asset class.

With technological advances, this modality of data collection and processing has received special attention. In fact, it is a reality that Data can come from a variety of sources (typically both internal and external to an organization) and in a variety of types. The capability to exploit data is not new, but nowadays, we have a reality of interconnection with the explosion of sensors, smart devices as well as social networking, and the emergence of the Internet of Things (IoT) concept, where these devices are all connected and constantly emitting information. For that reason, companies had to get up to speed and invest in new ways of dealing with not only structured traditional relational data but also semi-structured and unstructured data. (Monino, 2021). Even so, the upgrade in the data area does not arise only from the technological improvement, there was a change in data that is more accurate and in greater quantity in the present that allow us to draw better conclusions and make better decisions that will revolutionize the competitive business environment, according to Mayer-Schonberger and Cukier (Mayer-Schonberger and Cukier, 2013), the real revolution is not in the machines that calculate data but in data itself and how we use it.

The acceleration of the digital era and technological progress has transformed the lack of information to information overload. The development of large amounts of data and their access represents now which we called Big Data. The term describes innovative techniques and technologies to capture, store, distribute, manage, and analyse larger sized datasets with high velocity and diverse structures that conventional data management methods are incapable of handling.

There were also changes in the technicians who work with Big Data. So, to be prepared to work with this recent evolution required an adaptation of human capital. Next generation of data analysts and extant supply chain personnel should not show proof of their skills only in the technical and analytical aspect of their functions, as they are

expected to also be capable of transferring their analytics competences, understanding the big data ecosystem and applying such knowledge to business problems-(Brinch et al., 2021). Of course, this should also require a subset of skills, including data management, analytical models and tools, and business and decision-making processes. In fact, skills needed most for data scientists besides the technical ones come from subjects like forecasting (qualitative and quantitative), statistics (estimation and sampling) and economics (to understand the core business and the cost of opportunity in every decision). (Schoenherr & Speier-Pero, 2015).

Data is, therefore, a form of wealth, and exploiting it results in an essential competitive advantage for an ever-tougher market. The capability to exploit data is not new. What is new is the capability to do something meaningful with that data, rapidly and cost effectively. In a study, even though investments in big data analytics solutions are still very risky, researchers assessed four forms of business value - transactional, strategic, transformational, and informational - and concluded that successful big data investments generated profit from improved customer satisfaction and higher market performance. (Raguseo & Vitari, 2018)

However, it is necessary to pay attention to the quality of the data. The quality of the data is not always excellent, it can have wrong information, which means that if the data is not good, all the inferences and conclusions that are drawn from these datasets will also not be of good quality. At a minimum, data quality should be acknowledged, measured, monitored, and controlled. Ideally, the goal of any measurement or monitoring activity would be to improve the quality of the process and product. A lack of data maturity is likely to lead to a low level of big data value creation. This includes confining data within functional silos, not sharing data across the organization for analytic purposes, and the lack of governance procedures to secure adequate data quality and system standardization, all of which are critical for big data use. (Hazen et al., 2014)

2.2 Data Visualization & Dashboards

With the appearance of more data and to better understand it, other areas, such as data visualization, emerged and evolved. In the business world today, we need to paint a picture of the data so we can interpret it. Thus, data visualization gives us a clear idea of what information means by giving it visual context through maps or graphs. It is the last stage of the data flow process from the moment they are collected by the company until the moment they are presented to the data consumer, whether internal or external.

Data visualization is indeed important because having a huge amount of information at hand, the best way to create a simplification of that knowledge is through easy-to-understand visualizations. Although the analytics behind these views may be extremely complicated, the goal is to make it easier for any member of the organization to use. For these reasons, visualization techniques are tools to increase decision quality in decision making because they are used to not lose insights from information that can be crucial in decision making (Toasa et al., 2018). We cannot refer to data visualization without referring to the importance of design and the elements that make it up. Tufte, in "The Visual Display of Quantitative Information" (Tufte, 2001), claims that: “Graphical elegance is often found in simplicity of design and complexity of data”. In this sense, the goal of a good data visualization is always to facilitate reading through a simple and easily understandable design, which has behind it a disorganized and difficult to understand database.

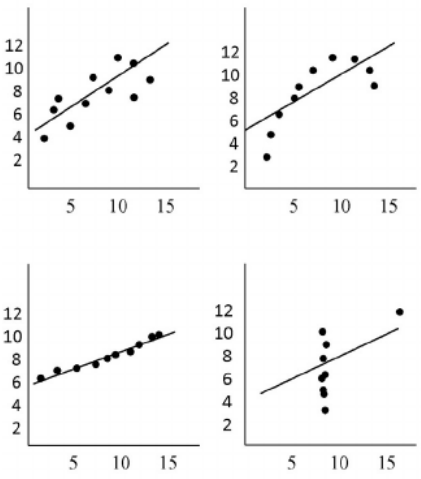
Despite being an area that has had a strong investment by companies, there is still a lot to improve as they are sometimes being used incorrectly. In practice, this area is often used only in the presentation of results and not in their exploration. According to Nestorov (Nestorov et al., 2019): There is a tendency to use flashy graphs and charts that do not allow the user to draw the intended conclusions, such as pattern and trend identification. Thus, with the diversity of data and respective types, there is a need for

improvement by the analysis and presentations to describe the relationships between elements and summarize complexity in a simple and understandable way.

There are some characteristics of a good design that serve as a guide for its good use, so it is important to emphasize that attractive displays of statistical information: have a properly chosen format and design; use words, numbers, and drawing together; reflect a balance, a proportion, a sense of relevant scale; display an accessible complexity of detail; often have a narrative quality, a story to tell about the data; are drawn in a professional manner, with the technical details of production done with care and avoid content-free decoration, including chart junk. (Tufte, 2001)

One example of a case which the data visualization can bring value is the Anscombe's quartet, see Figure 1. In this case, we have the scenario of four indistinguishable datasets using statistical methods: such as means, regression coefficient and residual sum of squares. However, the four datasets visualized as graphs, led to different conclusions. In fact, when visually analysed, the four graphs can be distinguished easily by a human, despite having similar values in statistical terms, which ultimately results in very different interpretations and conclusions. (Nestorov et al., 2019)

Fig. 1 – Anstcombe's Quartet



Property	Value
Mean of X (average)	9 in all 4 XY plots
Sample variance of X	11 in all four XY plots
Mean of Y	7.50 in all 4 XY plots
Sample variance of Y	4.122 or 4.127 in all 4 XY plots
Correlation (r)	0.816 in all 4 XY plots
Linear regression	y = 3.00 + (0.500 x) in all 4 XY plots

Data sets for the 4 XY plots

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	5.76
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	8.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	7.26	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

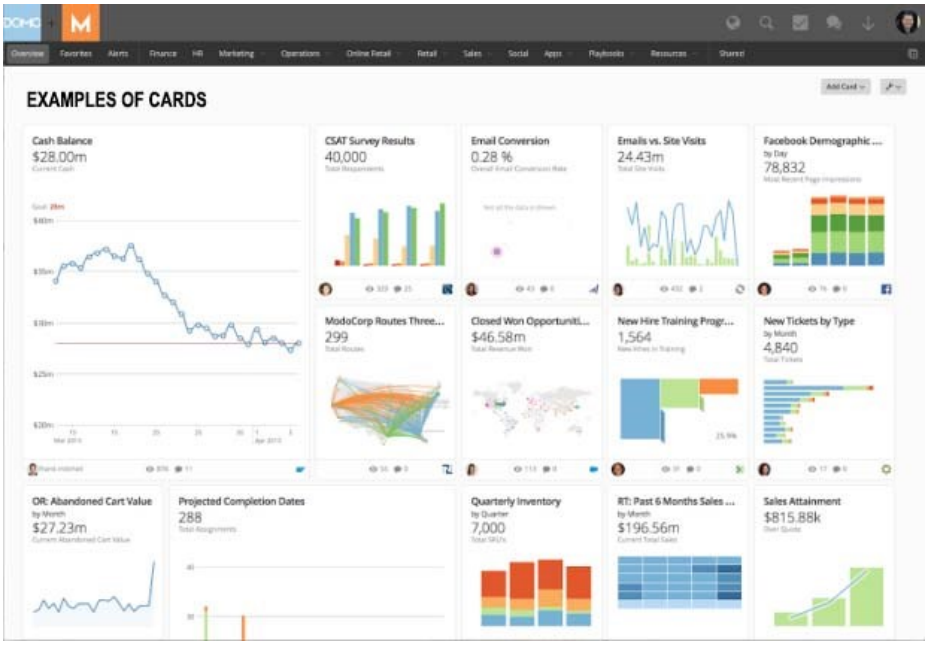
In advanced analytics, data scientists are creating machine learning algorithms to better compile essential data into visualizations that are easier to understand and interpret. Specifically, data visualization uses visual data to communicate information in a manner that is universal, fast, and effective. Visualized data gives stakeholders, business owners, and decision-makers a better prediction of sales volumes and future growth. Businesses with investment in data visualization can recognize patterns more quickly because they can interpret data in graphical or pictorial forms.

We cannot refer to Data visualization without describing the most common tool, the Dashboard. Despite its definition is in flux, according to a broad definition by Wexler (Wexler et al.), a dashboard is "a visual display of data used to monitor conditions and/or facilitate understanding". This concept has changed over time, in fact, has evolved from single-view reporting screens to include interactive interfaces with multiple views and purposes. (Sarikaya et al., 2019)

Visualization dashboards are ubiquitous. Nowadays, dashboards are created in a daily routine to monitor the company's processes and make them easier to monitor. For many people in these organizations, dashboards may be their first (or only) encounter with data - (Sarikaya et al., 2019). They are built and employed by nearly every industry, non-profit, and service organization to support data-driven decision making. The dashboard is, in a simple way, an easy-to-understand interface composed of Cards. Cards are used to track and display data in a summarized form. It can be a graph of your total sales, the market shares, an image, a filter or any other accounted element of your business that needs to be emphasized. In Figure 2 it is possible to see a set of examples of Cards in Domo.

Domo, Inc. is a cloud software company. It specializes in business intelligence tools and data visualization. The company is a cloud-based platform designed to provide direct, simplified, real-time access to business data for decision makers across the company with minimal IT involvement. It is a software-as-a-service (SaaS) venture.

Fig. 2 – Examples of Cards



2.3 Business Intelligence

Business Intelligence (BI) combines business analytics, data mining, data visualization, data tools, infrastructure and best practices to help organizations to make more data-driven decisions. The reason why it is so important is because it helps companies with issues such as: identify ways to increase profit, analyse customer behaviour, compare data with competitors, track performance, optimize operations, predict success, spot market trends, discover issues or problems and even anticipating them, among others.

Thus, BI is the set of strategies, processes, applications, data, products and technologies which are used to support the collection, analysis, presentation, and architecture of business information - (Kumar et al., 2017). BI technologies create the perfect solution to analyse the state of the company in relation to subjects such as the business trends, business growth, the amount of profit, employee performance,

customer satisfaction, areas of improvements in business and much more. It provides historic, existing, and predictive sights of business operations.

In today's competitive market situation, it is extremely important for small business and large corporations to have permanent access to analytical reports regarding their business activities - (Orlovskyi & Kopp, 2021). Intelligent businesses need Business Intelligence. They need it for recognizing, analysing, modelling, structuring and optimizing business processes. They need it, moreover, for making sense of massive amounts of unstructured data in order to support and improve highly sensible—if not highly critical - business decisions. The term "intelligent businesses" does not merely refer to commercial companies but also to (hopefully) intelligent governments, intelligently managed educational institutions, efficient hospitals, and so on. Every complex business activity can profit from BI. (Grossman et al., 2015)

For Vedder (Vedder et al., 1999), the product is information that will allow organizations to predict behaviours that can help organizations survive and thrive in the global economy. This is the role of Business Intelligence, translating transparency to the company so that it is possible to better understand its functioning and status both internally and externally. In fact, BI is most effective when it associates data from internal company sources to the business such as financial and operations data (internal data) with derivative data from the market in which a company operates (external data) - (Kumar et al., 2017).

2.4 Metadata & Data Governance

The word metadata only came into the English language in 1968, but the idea of metadata goes back to the first library. The word is a deliberate play on Aristotle's Metaphysics. Though Aristotle never called those works by that name, they have

historically been collected together under that title, to indicate that they came after, or dealt with topics beyond the Physics. Similarly, the word "metadata" indicates something that is beyond the data: a statement or statements about the data. Linguistically this is a loose translation of the Greek prefix "meta-", but it is consistent with what has become the everyday of the use of the word "meta", to indicate something at a higher level of abstraction-(Pomerantz, 2015). According to Pomerantz (Pomerantz, 2015), the importance of metadata lies in the fact that without information about the objects contained in a space, any sufficiently complex space is indistinguishable from chaos. Even when an object is contained within a space, if you want to find it within a timely fashion, you need metadata about it.

Metadata, or Metainformation, is information about data. An item of a metadata can tell what that data is about, usually information intelligible by a computer. Metadata makes it easier to understand the relationships and usefulness of information in the data. However, this definition is not consensual, according to Gilliland (Gilliland et al., 2016), the definition of Metadata, is literally, "data about data" and a widely used, yet frequently underspecified term that is understood in many ways by the diverse professional communities that design, create, describe, preserve, and use information systems and resources.

In fact, a search on the term "metadata" in the International Organization for Standardization's Online Browsing Platform (ISO OBP) reveals that there are 96 separate ISO standards that provide definitions of the term. However, the situation is not as problematic as it might seem: that low levels of standardization of definition across domains should not be a cause for concern. (Furner, 2020)

Metadata has been around for as long as humans have been organizing information and today we create and interact with it in increasingly digital ways (Gilliland et al., 2016). Over the last decade, digital repositories have grown at an explosive rate due to new information technologies, particularly those supporting World Wide Web (Web) applications. This growth has led to a tremendous increase in the need for data

management, an intense interest in metadata in a wide range of communities (e.g., education, government, scientific, business, etc.), and extensive development of metadata schemes. (Greenberg, 2012)

Data Governance is always related to Data quality, in fact, organizations are becoming increasingly serious about the notion of "data as an asset" as they face increasing pressure for reporting a "single version of the truth-(Khatri & Brown, 2010). Data governance (DG) is the process of managing the availability, usability, integrity and security of the data in enterprise systems, based on internal data standards and policies that also control data usage.

According to Fleissner (Fleissner et al., 2014), one example of good use of Data Governance is in the health care system. The data governance discipline is responsible for maintaining documentation, by data element, of which systems capture the data. It makes decisions on how to rationalize inconsistencies in data that is allegedly the same. It governs how the data can be used to ensure appropriate access, security and patient privacy. And if necessary, data is not captured in the way that is usable (or not captured at all), it identifies the need for potential changes in workflow and system implementation and engages the right stakeholders to affect the required modifications.

3. Corporate Metadata Dashboard

3.1 The company – EDP Comercial

The EDP Group is a global energy company, a leader in value creation, innovation and sustainability. EDP is a vertically integrated multinational utility. Over its more than 40 years of history, it has established a relevant presence in the world energy

scene, with a presence in 28 markets on 3 continents. The Group has been part of Euronext Lisbon since 1997 and EDP Brazil and EDP Renewables are also listed on the stock exchange.

The EDP Group has currently 12,180 employees, operating in 28 markets and satisfying the needs of 11,471,120 clients. The company's vision is to be a global energy company, a leader in the energy transition to create superior value. One of the main focuses of the company is its sustainability. Assuming the social and environmental responsibilities that result from its activities, contributes to the development of the regions where it is present, reducing, in a sustainable way, specific emissions of greenhouse gases from the energy it produces and actively promoting energy efficiency.

In general, the company presents itself as a reference in terms of energy in Portugal, as a company that has been growing and expanding its expansion into new markets, mainly to countries such as Spain, France, Italy, Poland, Brazil, among others with a smaller share.

In addition to the economic aspect, the company also creates an impact on a social level, and in December 2004 the EDP Foundation was created, one of the most important institutions in the cultural field in the country, in the double aspect of its own production and patronage.

3.2 Limitations of the current situation

A major problem that organizations currently face is the internal management of large disorganized data quantities. Companies are opting to invest in data structure projects that increase transparency and organisation of information. This investment facilitates the access of any member of the company to the information. Business intelligence departments have a huge impact on this problem because it is an area that creates

organized information structures that allows for greater organization and veracity of data. Sales, marketing, finance and operations departments can use business intelligence. Tasks include quantitative analysis, measuring performance against business goals, gleaning customer insights and sharing data to identify new opportunities.

The objective of the investment is to add value to the "Departments of Business and Decision Making". For this, the departments that deal with Data must make these structures as organized and transparent as possible so that there is a clear and simple flow of information within the companies. One of the clear goals of the companies is that the flow of information is so transparent that any employee of a department that intends to verify a certain value in a table or dashboard knows where that information is. Only in this sense companies can improve internal management with more transparent and easily accessible information that improves the organization of the company as a whole and, as a result, we can get added value from this investment and a good return on investment (ROI).

In the case of EDP Comercial - Energias de Portugal Comercial, the organization of data from extracting, cleaning, analysing, automating and creating reports for the rest of the company is handled by a group composed of almost 60 people named "Systems and Data Department". Within this department there is the division of the group into smaller departments, specialized in each function. In the case of reporting and creating Dashboards, the department that has this function is BI & Analytics. This department is divided into two subsectors: Business Intelligence and Analytics. Thus, the Analytics sub-department deals with projects more focused on automating repetitive processes within the company, forecasting models and several projects with Machine Learning that allows the company to innovate and incorporate cutting-edge technologies in its processes. As for the Business Intelligence sub-department, it reports the entire company in relation to the different subdivisions. In fact, it is in this area that Dashboards are created and updated on different subjects (Sales,

Commissioning, Debt, Electric Mobility, among others) which are then available to the Business areas, which can consult them whenever they want objective, real and verified information for consultation, verification, or decision making.

Due to the size of the company, and the growing and continuous amount of information that a company of this size faces, many obstacles arise in this organization. In fact, a high amount of information dashboards are created and updated at a fast pace and there is a need to align these contents so that they are useful, verified and used correctly by other departments. In this way, the need arose to create a Metadata project: a Repository/Dashboard with all the Metadata and Metainformation from the developments of the BI & Analytics department to facilitate the internal management. The project had as objectives the improvement of Internal Management, thus, through the development of the Repository/Dashboard, it was possible to facilitate the location, comprehension and preservation of all the digital records. Another advantage arising from the development of this project was the elimination of duplication of information: first, due to the lack of organization of the content already developed (Dashboards), work already carried out was requested and, second, it was natural to duplicate content by different areas and different departments. What in addition to being an expense in terms of human resources, to make matters worse, different business areas sometimes reached different values and conclusions on the same topic, leading to interdepartmental disorganization. In quantitative terms, as the company pays the software - DOMO (software used for the development of all information repositories in the company through data) per calculation line, the smaller this duplication of information, the lower the cost imposed to the company. Regarding the improvement of this process, which would avoid data duplication, the company would be less subject to two problems: the allocation of human resources (Business Intelligence analysts or internal dashboard creators in every non-BI department) to develop Dashboards that had already been

developed and the use of more lines of calculation in the software with associated charges.

The main part of the Project would be to gather, in this way, all the information created so far and incorporate it into an index document to facilitate the consultation of what is intended. Simplicity was requested and the simpler the more effective because the main users would be employees of both B2B and B2C business departments that do not have know-how of data structures and their visualization. The objective would thus be to create something precise and robust enough to convey credibility to the company's information. Thus, it would be expected that employees could consult this repository of credible information with standardized concepts (KPI's), reducing the time needed to search for information by each employee and reaching a higher analytical maturity.

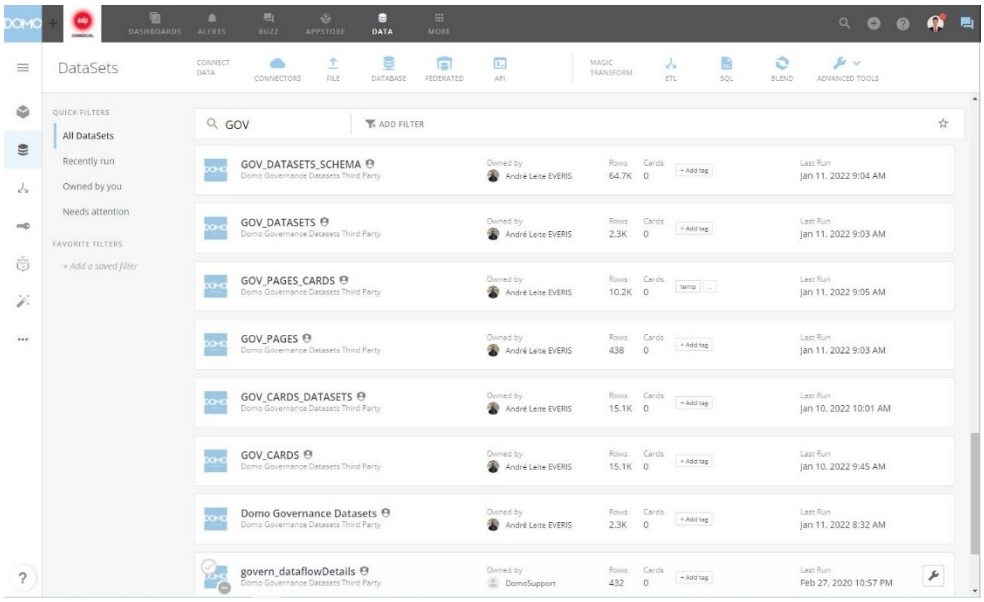
3.3 Project Overview

3.3.1 Data collection and organization

First, it was necessary to contextualize the problem, and get a clear perception of what was being asked for and for what purpose it was being asked. Thus, it was clear that there was no Dashboard with information on all the developments the Department had developed, that is, there was no document showing which Dashboards were already developed, who developed them and what content they had. This was a problem, because the other departments did not have information about what had already been developed and the Business Intelligence team would thus have to have a broad notion of what has already been done and where each information was, in case there was any request. The first step to be taken, was to organize a set of loose Governance tables, which were several tables/datasets that contained information about the dashboards: for example, a dataset with all dashboards created, a dataset with all the cards, a dataset with descriptions of the cards, a dataset with the

dashboard's developers. Briefly, crucial information spread over several Datasets. These datasets were a good start and needed to be analysed, verified and after that, structured to create a Megadataset with the adapted information that was requested. In this first period, it was important to realize that it was necessary to collect verified information. For this reason, there was a period of analysis of the Governance Datasets regarding their content and data up-to-dateness. This time to verify each Governance dataset was crucial and of high importance because it made no sense to create something structured without the project's roots containing real or verified information. In this sense, these tables were analysed in detail so that the information they contained was, in fact, aligned with the project's objectives, that is, the creation of a reliable tool that was easy to access and that contained updated, verified and correct data.

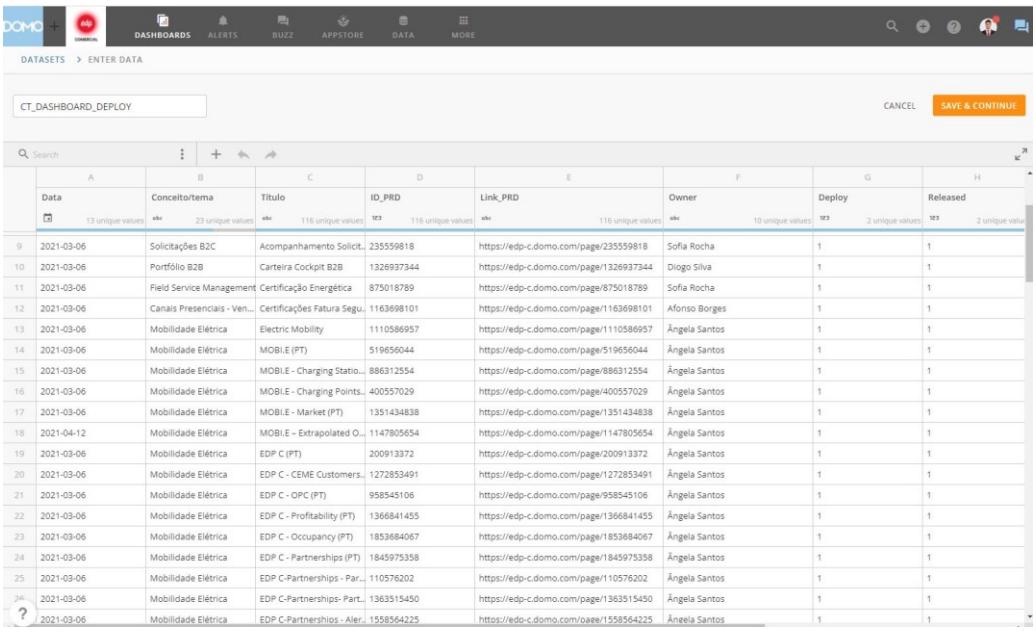
Fig. 3 – Domo's Governance Datasets



In a second phase, after having the different Governance Datasets checked as to the quality, veracity and timeliness of their information, it was necessary to start organizing the information found in different Business Segments. Thus, for example, to facilitate future research and improve the organization, the Dashboards that contained information on "Electric Mobility", which are all those related to electric

chargers for electric cars, which is, information on their location, operators, maintenance cost, revenue, etc. would be separated from the Dashboards of "In-Person Channels – Sales", which present information on the Commercial Performance of Stores, Agents, Partners in detail. Thus, in this second phase there was a need to understand what each Dashboard already created contained in terms of content, what was its purpose and after having a clear perception of it then, fit it into a Segment, crated to organize the information. This step required understanding all areas of the company's business. In fact, at this stage we were verifying the content of the dashboards and allocating them to a Concept/Segment to enable a greater organization of this information within the company. The theme would be how strategically we could join the dashboards into groups to further facilitate the filtering of Concepts.

Fig. 4 – Grouping Dashboards into Concepts



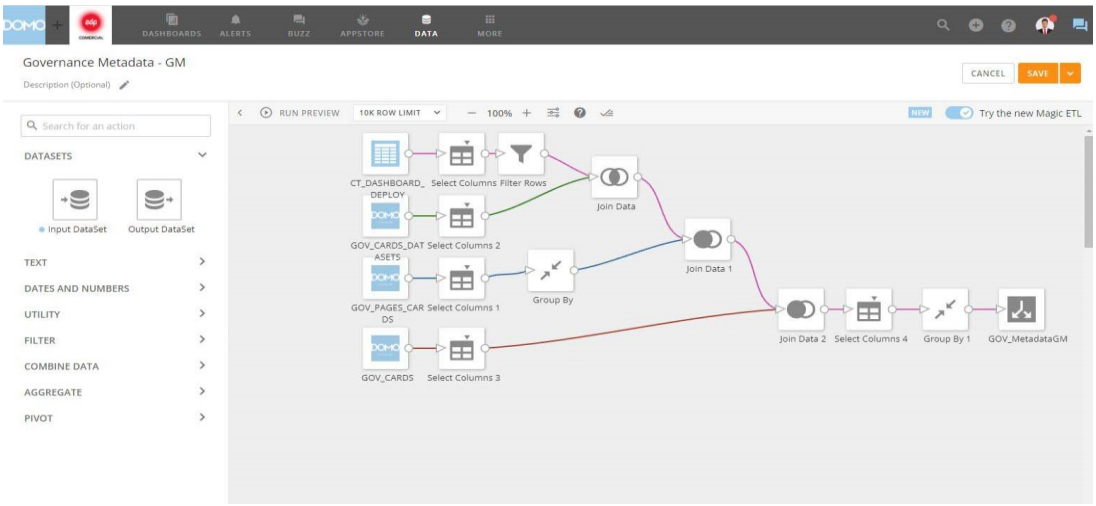
	Data	Conceito/tema	Título	ID_PRD	Link_PRD	Owner	Deploy	Released
9	2021-03-06	Solicitações B2C	Acompanhamento Solicit...	235559818	https://edp-c.domo.com/page/235559818	Sofia Rocha	1	1
10	2021-03-06	Portfólio B2B	Carteira Cockpit B2B	1326937344	https://edp-c.domo.com/page/1326937344	Diogo Silva	1	1
11	2021-03-06	Field Service Management	Certificação Energética	875018789	https://edp-c.domo.com/page/875018789	Sofia Rocha	1	1
12	2021-03-06	Canais Presenciais - Ven...	Certificações Fatura Segu...	1163698101	https://edp-c.domo.com/page/1163698101	Afonso Borges	1	1
13	2021-03-06	Mobilidade Elétrica	Electric Mobility	1110586957	https://edp-c.domo.com/page/1110586957	Ângela Santos	1	1
14	2021-03-06	Mobilidade Elétrica	MOBI.E (PT)	519656044	https://edp-c.domo.com/page/519656044	Ângela Santos	1	1
15	2021-03-06	Mobilidade Elétrica	MOBI.E - Charging Statio...	886312554	https://edp-c.domo.com/page/886312554	Ângela Santos	1	1
16	2021-03-06	Mobilidade Elétrica	MOBI.E - Charging Points...	400557029	https://edp-c.domo.com/page/400557029	Ângela Santos	1	1
17	2021-03-06	Mobilidade Elétrica	MOBI.E - Market (PT)	1351434838	https://edp-c.domo.com/page/1351434838	Ângela Santos	1	1
18	2021-04-12	Mobilidade Elétrica	MOBI.E - Extrapolated O...	1147805654	https://edp-c.domo.com/page/1147805654	Ângela Santos	1	1
19	2021-03-06	Mobilidade Elétrica	EDP C (PT)	200913372	https://edp-c.domo.com/page/200913372	Ângela Santos	1	1
20	2021-03-06	Mobilidade Elétrica	EDP C - CEME Customers...	1272853491	https://edp-c.domo.com/page/1272853491	Ângela Santos	1	1
21	2021-03-06	Mobilidade Elétrica	EDP C - OPC (PT)	958545106	https://edp-c.domo.com/page/958545106	Ângela Santos	1	1
22	2021-03-06	Mobilidade Elétrica	EDP C - Profitability (PT)	1366841455	https://edp-c.domo.com/page/1366841455	Ângela Santos	1	1
23	2021-03-06	Mobilidade Elétrica	EDP C - Occupancy (PT)	1853684067	https://edp-c.domo.com/page/1853684067	Ângela Santos	1	1
24	2021-03-06	Mobilidade Elétrica	EDP C - Partnerships (PT)	1845975358	https://edp-c.domo.com/page/1845975358	Ângela Santos	1	1
25	2021-03-06	Mobilidade Elétrica	EDP C-Partnerships - Par...	110576202	https://edp-c.domo.com/page/110576202	Ângela Santos	1	1
26	2021-03-06	Mobilidade Elétrica	EDP C-Partnerships- Part...	1363515450	https://edp-c.domo.com/page/1363515450	Ângela Santos	1	1
27	2021-03-06	Mobilidade Elétrica	EDP C-Partnerships - Aler...	1558564225	https://edp-c.domo.com/page/1558564225	Ângela Santos	1	1

After grouping the datasets by Concepts/Segments, it was necessary to start thinking about how we were going to gather all the important information in a granulated Megadataset on the card. This Megadataset thus, contained the information of all the

datasets that we had checked together, that is, it would have the information of all the Dashboards created, who developed them, which Segment/Concept they belonged to, which cards they contained - each graphic, table or relevant information, as the description of each card and what datasets were powering it. This was basically all the information we were looking for. Therefore, through an ETL we were able to group the four most important Datasets in this sense, which were enough to gather all the required information. To do this we use one of Domo 's features and created a Magic ETL.

Thus, as we can see in Figure 4 -"Creating the Magic ETL", we start by selecting the desired and relevant columns in each independent dataset. After this step, in the case of the Dataset with the intended Segment/Concept, the "CT_Dashboard_Deploy", where each line corresponded a different Dashboard, we had to filter the intended lines, that is, only consider the lines (Dashboards) with a relevant concept, leaving behind the lines (Dashboards) as " Help", "FAQ", "Instructions", among other topics, such as Dashboards related to Covid-19, which were not considered relevant for future analysis, and it made no sense to enter the department's Metadata Dashboard.

Fig. 5 – Creating the Magic ETL



Then, in the first Inner Join that we do, when we add the "CT_DASHBOARD_DEPLOY" to the "GOV_CARDS_DATASETS" we do it in order

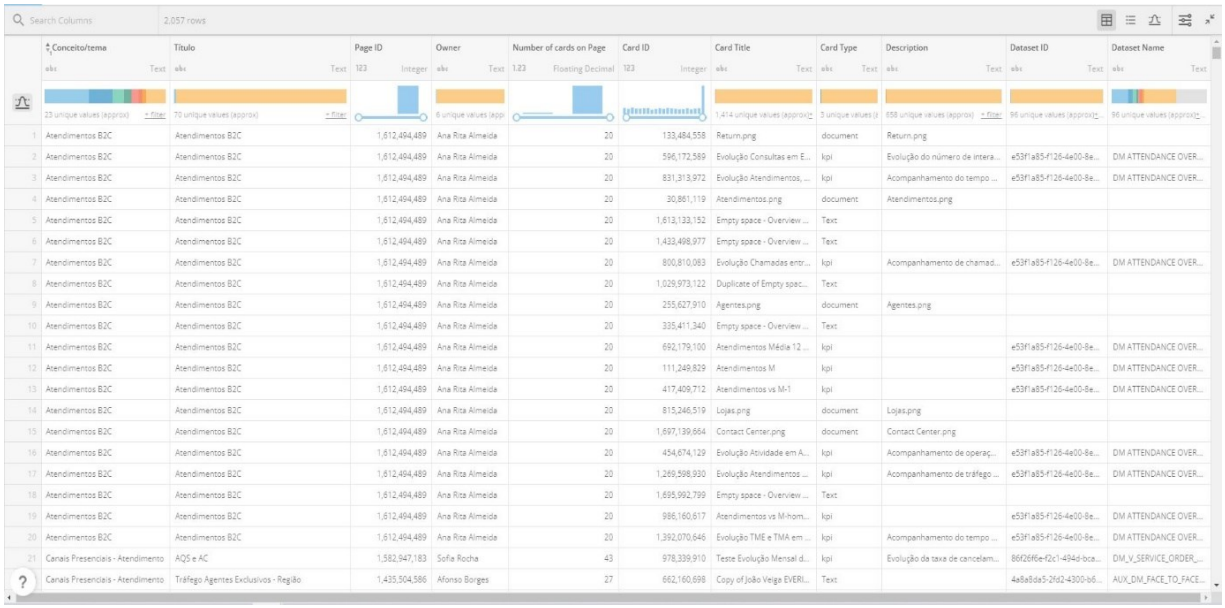
to join the: Concepts/Segments that we created and the Owners of a Dashboard that came from the "CT_DASHBOARD_DEPLOY" with the: titles of each Dashboard, the type and ID's of the cards and the titles and ID's of the Datasets that are feeding this Dashboard that were present in "GOV_CARDS_DATASETS". The linkage that allowed these two Datasets to be joined was the Page ID, which was present in both, and which is simply the ID of each Dashboard.

We followed this direction throughout the creation of the ETL, being that, in the second one, which is a left join, since we don't want what just coincides in these but also what comes from behind, when adding "GOV_PAGES_CARDS" we do it to join more information for each card: such as the number of cards in each Dashboard and their title.

Finally, we added the "GOV_CARDS" to get the descriptions of each Card, that is, an explanation of what is done in each card of each Dashboard. It is usually the description of each graph or table of a Dashboard to elucidate and explain what is being visualized through that card.

Therefore, the last steps are just to consider which columns will be the most relevant to analyse, eliminate the duplicates and group the columns by, first, Dashboard and then by card, to have a visual waterfall effect, that is, display each Dashboard and its information and then the cards of that Dashboard and their characteristics, as we can see in Fig.6 – "Magic ETL's Output - Gov_MetadataGM". The final output of this Dataset has 2057 rows, that is, as the Dataset is granulated to the Card, there are 2057 cards in the sum of all dashboards of the department.

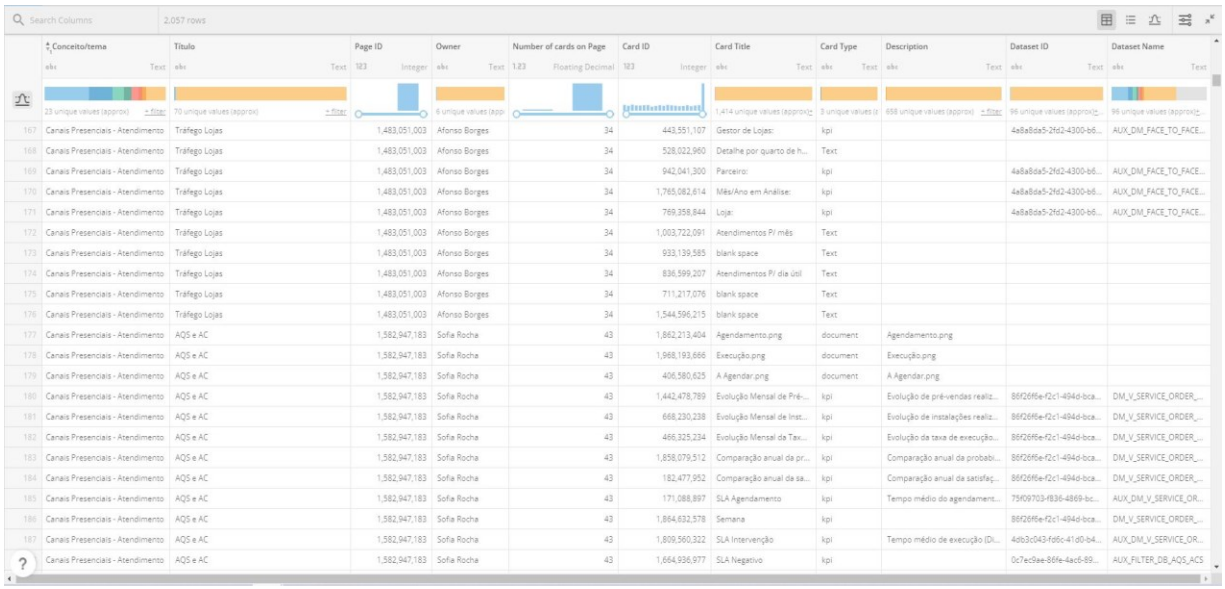
Fig. 6 – Magic ETL's Output – Gov_MetadataGM



	Conceito/tema	Título	Page ID	Owner	Number of cards on Page	Card ID	Card Title	Card Type	Description	Dataset ID	Dataset Name
	abc Text	abc Text	123 Integer	abc Text	1.23 Floating Decimal	123 Integer	abc Text	abc Text	abc Text	abc Text	abc Text
	23 unique values (approx)	70 unique values (approx)		6 unique values (app			1,414 unique values (approx)	3 unique values (a	658 unique values (approx)	96 unique values (approx)	96 unique values (approx)
1	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	133,484,558	Return.png	document	Return.png		
2	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	596,172,589	Evolução Consultas em E...	kpi	Evolução do número de intera...	e53f1a85-f126-4e00-8e...	DM ATTENDANCE OVER...
3	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	831,313,972	Evolução Atendimentos, ...	kpi	Acompanhamento do tempo ...	e53f1a85-f126-4e00-8e...	DM ATTENDANCE OVER...
4	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	30,861,119	Atendimentos.png	document	Atendimentos.png		
5	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	1,613,133,152	Empty space - Overview ...	Text			
6	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	1,433,498,977	Empty space - Overview ...	Text			
7	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	800,810,083	Evolução Chamadas entr...	kpi	Acompanhamento de chamad...	e53f1a85-f126-4e00-8e...	DM ATTENDANCE OVER...
8	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	1,029,973,122	Duplicate of Empty spac...	Text			
9	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	255,627,910	Agentes.png	document	Agentes.png		
10	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	335,411,340	Empty space - Overview ...	Text			
11	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	692,179,100	Atendimentos Média 12 ...	kpi		e53f1a85-f126-4e00-8e...	DM ATTENDANCE OVER...
12	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	111,249,829	Atendimentos M	kpi		e53f1a85-f126-4e00-8e...	DM ATTENDANCE OVER...
13	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	417,409,712	Atendimentos vs M-1	kpi		e53f1a85-f126-4e00-8e...	DM ATTENDANCE OVER...
14	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	815,246,519	Lojas.png	document	Lojas.png		
15	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	1,697,139,664	Contact Center.png	document	Contact Center.png		
16	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	454,674,129	Evolução Atividade em A...	kpi	Acompanhamento de operaç...	e53f1a85-f126-4e00-8e...	DM ATTENDANCE OVER...
17	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	1,269,598,930	Evolução Atendimentos ...	kpi	Acompanhamento de tráfego ...	e53f1a85-f126-4e00-8e...	DM ATTENDANCE OVER...
18	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	1,695,992,799	Empty space - Overview ...	Text			
19	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	986,160,617	Atendimentos vs M-hom...	kpi		e53f1a85-f126-4e00-8e...	DM ATTENDANCE OVER...
20	Atendimentos B2C	Atendimentos B2C	1,612,494,489	Ana Rita Almeida	20	1,392,070,646	Evolução TME e TMA em ...	kpi	Acompanhamento do tempo ...	e53f1a85-f126-4e00-8e...	DM ATTENDANCE OVER...
21	Canais Presenciais - Atendimento	AQS e AC	1,582,947,183	Sofia Rocha	43	978,339,910	Teste Evolução Mensal d...	kpi	Evolução da taxa de cancelam...	86f26f6e-f2c1-494d-bca...	DM_V_SERVICE_ORDER_...
?	Canais Presenciais - Atendimento	Tráfego Agentes Exclusivos - Região	1,435,504,586	Afonso Borges	27	662,160,698	Copy of João Veiga EVERI...	Text		4a8a8da5-2fd2-4300-b6...	AUX_DM_FACE_TO_FACE...

In this graph, the Concept/Segment and the title have the same name because in relation to the Concept/Segment: "Atendimento B2C" all the information is encompassed in a single Dashboard, which, as we can see, has the ID: 1612494489, the Owner is Ana Rita Almeida and has 20 cards (in the Dataset we can also check the information of each card that makes up the Dashboard)

Fig. 7 – ETL's Output 2 – Gov_MetadataGM



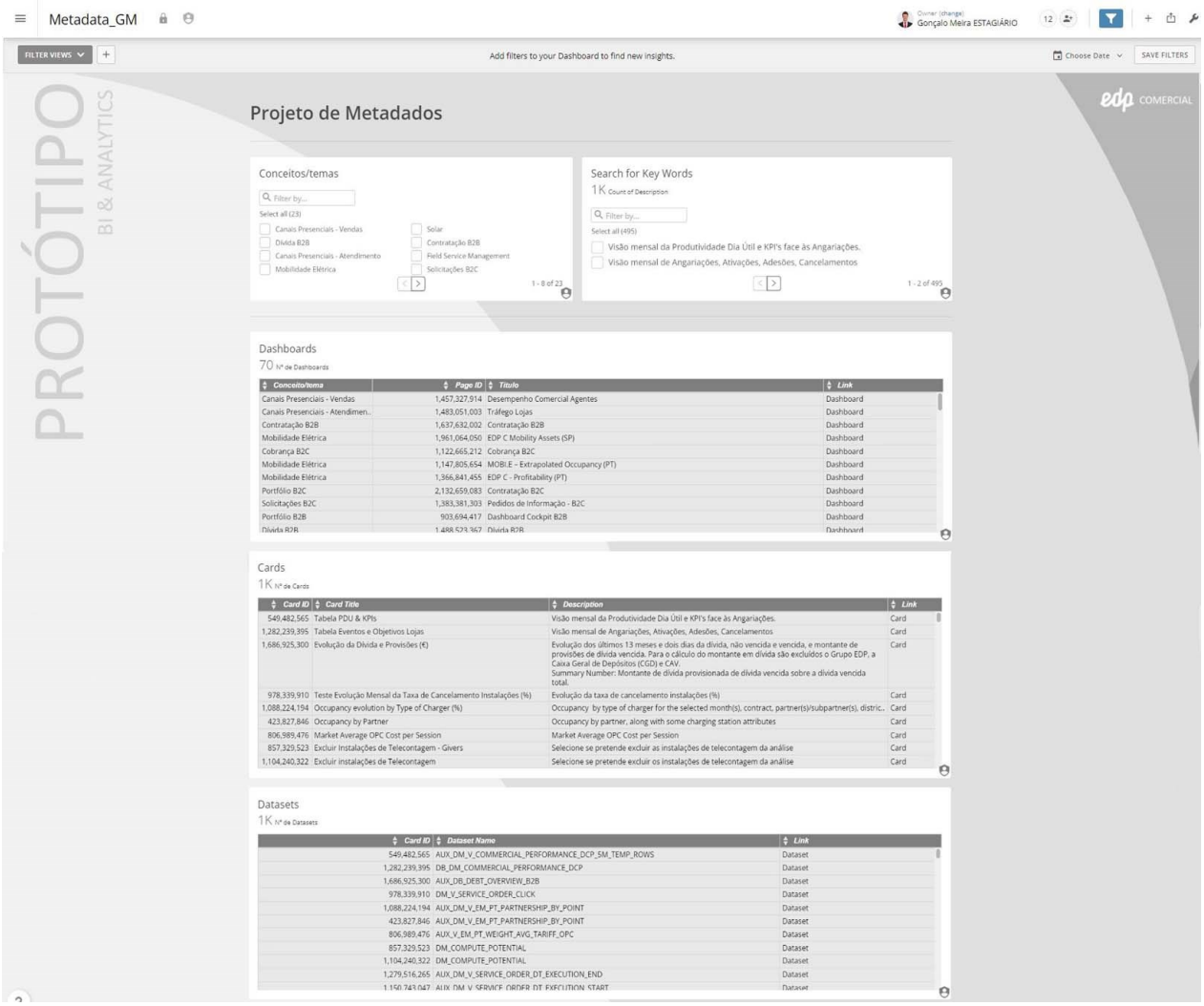
	Conceito/tema	Título	Page ID	Owner	Number of cards on Page	Card ID	Card Title	Card Type	Description	Dataset ID	Dataset Name
	abc Text	abc Text	123 Integer	abc Text	1.23 Floating Decimal	123 Integer	abc Text	abc Text	abc Text	abc Text	abc Text
	23 unique values (approx)	70 unique values (approx)		6 unique values (app			1,414 unique values (approx)	3 unique values (a	658 unique values (approx)	96 unique values (approx)	96 unique values (approx)
167	Canais Presenciais - Atendimento	Tráfego Lojas	1,483,051,003	Afonso Borges	34	443,551,107	Gestor de Lojas:	kpi		4a8a8da5-2fd2-4300-b6...	AUX_DM_FACE_TO_FACE...
168	Canais Presenciais - Atendimento	Tráfego Lojas	1,483,051,003	Afonso Borges	34	528,022,960	Detalhe por quarto de h...	Text			
169	Canais Presenciais - Atendimento	Tráfego Lojas	1,483,051,003	Afonso Borges	34	942,041,300	Parceiro:	kpi		4a8a8da5-2fd2-4300-b6...	AUX_DM_FACE_TO_FACE...
170	Canais Presenciais - Atendimento	Tráfego Lojas	1,483,051,003	Afonso Borges	34	1,765,082,614	Mês/Ano em Análise:	kpi		4a8a8da5-2fd2-4300-b6...	AUX_DM_FACE_TO_FACE...
171	Canais Presenciais - Atendimento	Tráfego Lojas	1,483,051,003	Afonso Borges	34	769,358,844	Loja:	kpi		4a8a8da5-2fd2-4300-b6...	AUX_DM_FACE_TO_FACE...
172	Canais Presenciais - Atendimento	Tráfego Lojas	1,483,051,003	Afonso Borges	34	1,003,722,091	Atendimentos P/ mês	Text			
173	Canais Presenciais - Atendimento	Tráfego Lojas	1,483,051,003	Afonso Borges	34	933,139,585	blank space	Text			
174	Canais Presenciais - Atendimento	Tráfego Lojas	1,483,051,003	Afonso Borges	34	836,599,207	Atendimentos P/ dia útil	Text			
175	Canais Presenciais - Atendimento	Tráfego Lojas	1,483,051,003	Afonso Borges	34	711,217,076	blank space	Text			
176	Canais Presenciais - Atendimento	Tráfego Lojas	1,483,051,003	Afonso Borges	34	1,544,596,215	blank space	Text			
177	Canais Presenciais - Atendimento	AQS e AC	1,582,947,183	Sofia Rocha	43	1,862,213,404	Agendamento.png	document	Agendamento.png		
178	Canais Presenciais - Atendimento	AQS e AC	1,582,947,183	Sofia Rocha	43	1,968,193,666	Execução.png	document	Execução.png		
179	Canais Presenciais - Atendimento	AQS e AC	1,582,947,183	Sofia Rocha	43	406,580,625	A Agendar.png	document	A Agendar.png		
180	Canais Presenciais - Atendimento	AQS e AC	1,582,947,183	Sofia Rocha	43	1,442,478,789	Evolução Mensal de Pré-...	kpi	Evolução de pré-vendas realiz...	86f26f6e-f2c1-494d-bca...	DM_V_SERVICE_ORDER_...
181	Canais Presenciais - Atendimento	AQS e AC	1,582,947,183	Sofia Rocha	43	668,230,238	Evolução Mensal de Inst...	kpi	Evolução de instalações realiz...	86f26f6e-f2c1-494d-bca...	DM_V_SERVICE_ORDER_...
182	Canais Presenciais - Atendimento	AQS e AC	1,582,947,183	Sofia Rocha	43	466,325,234	Evolução Mensal da Tax...	kpi	Evolução da taxa de execução...	86f26f6e-f2c1-494d-bca...	DM_V_SERVICE_ORDER_...
183	Canais Presenciais - Atendimento	AQS e AC	1,582,947,183	Sofia Rocha	43	1,858,079,512	Comparação anual da pr...	kpi	Comparação anual da probabi...	86f26f6e-f2c1-494d-bca...	DM_V_SERVICE_ORDER_...
184	Canais Presenciais - Atendimento	AQS e AC	1,582,947,183	Sofia Rocha	43	182,477,952	Comparação anual da sa...	kpi	Comparação anual da satisfaç...	86f26f6e-f2c1-494d-bca...	DM_V_SERVICE_ORDER_...
185	Canais Presenciais - Atendimento	AQS e AC	1,582,947,183	Sofia Rocha	43	171,088,897	SLA Agendamento	kpi	Tempo médio do agendament...	75f09703-f836-4869-bc...	AUX_DM_V_SERVICE_OR...
186	Canais Presenciais - Atendimento	AQS e AC	1,582,947,183	Sofia Rocha	43	1,864,632,578	Semana	kpi		86f26f6e-f2c1-494d-bca...	DM_V_SERVICE_ORDER_...
187	Canais Presenciais - Atendimento	AQS e AC	1,582,947,183	Sofia Rocha	43	1,809,560,322	SLA Intervenção	kpi	Tempo médio de execução (Di...	4db3c043-fd6c-41d0-b4...	AUX_DM_V_SERVICE_OR...
?	Canais Presenciais - Atendimento	AQS e AC	1,582,947,183	Sofia Rocha	43	1,664,936,977	SLA Negativo	kpi		0c7ec9ae-86fe-4ac6-89...	AUX_FILTER_DB_AQS_ACS

In this case, we have the same Concept/Segment "Canais Presenciais - Atendimento" and 2 of 6 different Dashboards that constitute it, which in the image are the Dashboard of "Trafego Lojas" and "AQS e AC". Each one with a different ID, different Owner, different number of cards that make it up, different Datasets that feed them and each of the cards with its own information.

3.3.2 Creation of the Dashboard

Once built the Megadataset with all the desired information, it was time to think about the structure of the Dashboard that would allow this information to be illustrated in a succinct and clear way. It should be noted that the main objective was to facilitate the processing of information, so, simplicity was intended in this dashboard. Thus, after meeting several formulations, it was concluded that the easiest way to present the data was following a 'Concept/Segment-Dashboard-Card-Dataset' waterfall effect, that is, the project would be based on filters so that the information could be found more rapidly. For this, several tables were created, each one of them to facilitate the flow of information within the metadata dashboard. A table with all the dashboards available, one with all the cards and a last one with all the datasets. Thus, in the dashboard there would be a filter for each of the subtopics described above. That is, if the user was looking for a particular dashboard of the Electric Mobility Segment, he could filter by the Concept/Segment: "Electric Mobility" and find all the dashboards of that theme, all the cards of each of those dashboards and all the datasets that were feeding each of the cards. Then, if the user wanted to filter only one dashboard, the same thing happened, in the case of the dashboard filter, only the desired dashboard appeared with all the associated cards and the datasets that feed that card, thus explaining the waterfall effect. The same thing happened if the objective was to find a particular card, it was possible to filter by just that card and thus perceive the dataset that was feeding that card, the dashboard to which that card belonged and the Concept/Segment to which it was inserted.

Fig. 8 – Dashboard's Draft



In addition to the filters created on these subtopics, a description filter was also created, as you can see in Fig.8 – Dashboard's Draft. This request made sense because often, the user looks for a particular card whether it has already been created or not and does not need to go through the traditional waterfall format, that is, filter by Concept/Segment, then look in which dashboard is the card and check which of the cards has the desired information. Instead, in the word filter – the table "Search for Key Words" -, the user writes something about the intended card description that allows it to filter and retain only cards with those descriptions. Then, the user can check if that card exists and uses it. Or even if he wants that card but with another

type of information, he already knows exactly which dataset has that information and the request becomes easier.

For example, if the user wants cards with information on "taxas de cancelamento" (cancellation rates), and filters by that description in the word filter, only 6 cards will be retained: 3 on the evolution of installation cancellation rates (%) and 3 on the evolution of pre-sales cancellation rates (%). All these cards fed by the "DM_V_SERVICE_ORDER_CLICK" dataset and all in the Concept/Segment of "Canais Presenciais-Atendimento" (Personal Service-Service Channels), as seen in Fig.9 – Filtering by Descriptions I and Fig.10 - Filtering by Descriptions II. The user thus has access to a quick filter that allows him to search for information more easily through descriptions.

Fig. 9 – Filtering by Descriptions I

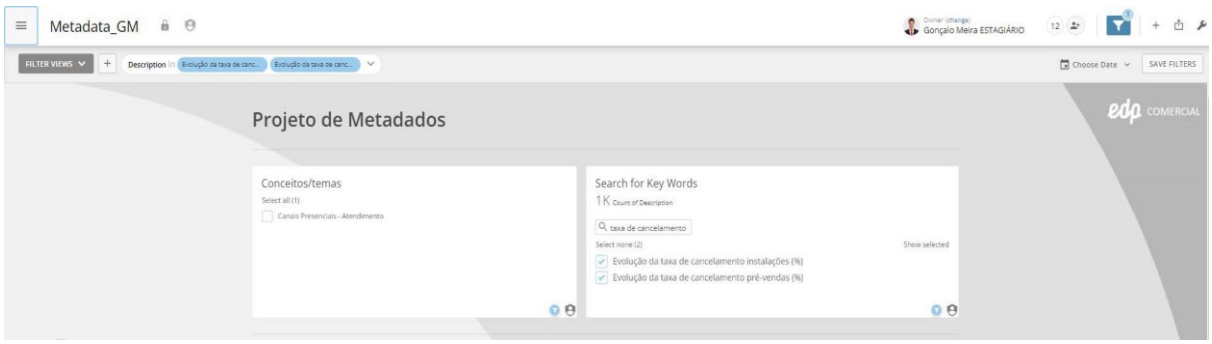
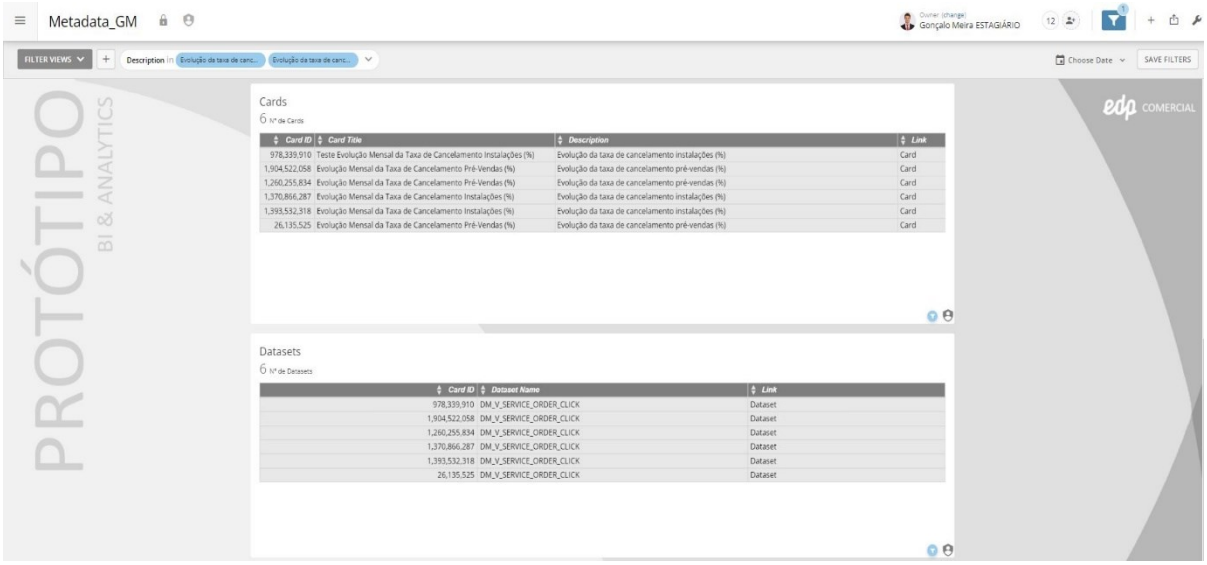
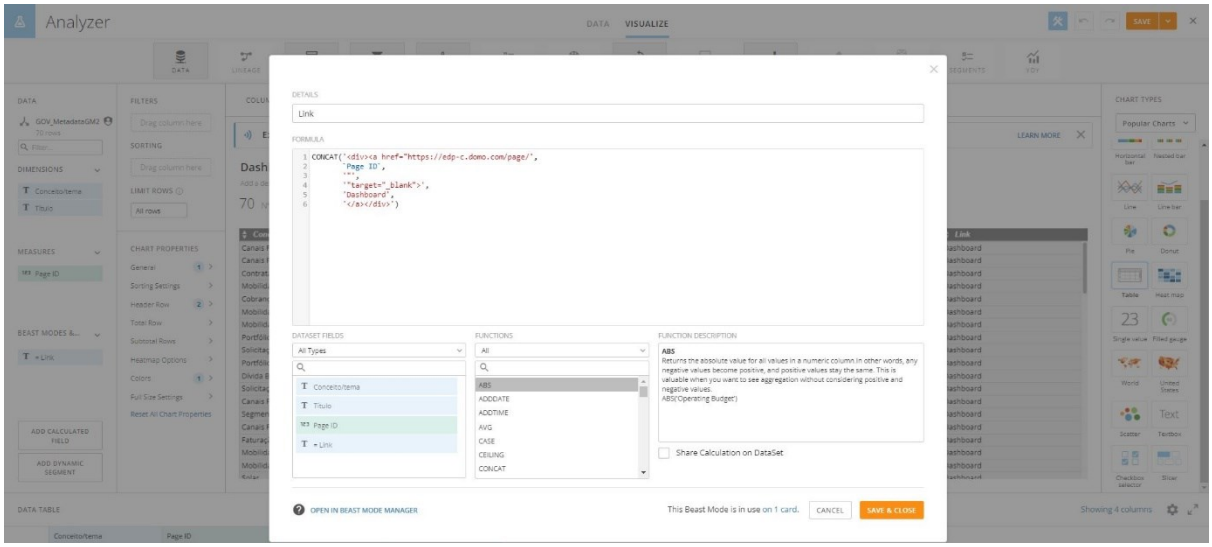


Fig. 10 – Filtering by Descriptions II



Finally, to speed up the process and facilitate access to data, it would be important to use links between the new Metadata Dashboard created and each of the different dashboards, cards or datasets. Thus, as the link of each dashboard is given by a different code ("https://edp-c.domo.com/page/" + Dashboard ID), it was possible to create a code that would allow this quick connection to be made through one of the features of the DOMO software in Analyzer, the Beast Mode, and thus create quick links that, with just one click, would open us in a different tab the dashboard we clicked on. The code that was developed for this new functionality can be confirmed in Figure 11 – Hyperlink Beast Mode Code.

Fig. 11 – Hyperlink Beast Mode Code



The same reasoning was developed for each card where the hyperlink code consists of:

- "https://edp-c.domo.com/page/" + Dashboard ID + "/kpis/details/" + Card ID.

Just like it was developed for each Dataset, where the code was:

- "https://edp-c.domo.com/datasources/" + Dataset ID + "/details/data/table".

So, with the help of DOMO's Beast Mode we created a code that allowed the introduction of the Hyperlink in our Metadata Dashboard and thus, we adapted all our tables. In fact, in all of them, a column called "Link" was created, which, if clicked, would allow the desired dashboard, card or dataset to be opened in a new tab.

The process of hyperlinks is very useful because it facilitates the search of the contents in the dashboard. This new functionality allows a fast, clear and objective search. The advantage is that it is a simplified tool so that any company user can navigate the metadata panel and keep opening tabs of any dataset, dashboard or card to check what information is on each one until they find what they are looking for.

To summarize, the process designed for this dashboard was to reduce the search for information to the shortest possible time and make it easier. The dashboard thus has, in a practical way, a very simplified visual content composed of only three tables: Dashboards, Cards and Datasets. It is in these tables that all the information related to the projects developed by the department is compiled, and this information is reviewed and confirmed. In practical terms, a user looking for specific information also has two extra tables available to assist in the search, the table of contents - with 23 general contents that the department works on and reports on - which serves the purpose of a general filter fast to make the search journey easier and less extensive; and a second instrument, which is the search for words to filter on topics more quickly if the user already has a very specific idea of what he is looking for. The process was made simpler with the use of hyperlinks to facilitate access to content, opening a new browser tab with the display of the clicked content. In this sense, the visualization of data and information is easier and more accessible to any user of the Metadata Dashboard.

3.4 Results

As for the Metadata Dashboard, after being developed and tested it is necessary to check what the 'expectation vs results' was, and if the major objectives were achieved. Thus, in relation to the results, it is important to subdivide this topic into two sub-themes: the results in relation to visualization, that is, how the concrete objectives were achieved in view of the visualization of the data of this project. And a second sub-theme: economic, that is, understanding how this project also had a positive economic character.

3.4.1 Visualization Results

From the beginning, simplicity and objectivity were requested in this project. Creating a functional and useful dashboard for the organization was in fact the most important and so, the simpler, more logical and effective, the better. In this sense, as previously seen, a carefully thought out and revised dashboard was created so that the information was reliable. According to the literature review, data visualization follows a set of general ideals that can make the project more effective the more "easy-to-understand" the project is. In this way, the Metadata Dashboard on the projects developed by the BI & Analytics department paid special attention to the importance of the design of the elements that compose it. In summary, after studying the best way to visualize the data for the preparation and organization of the data, a specific format and design was chosen to facilitate the visualization - waterfall format between tables to facilitate the follow-up of information search and, in addition, the possibility of searching for information through a word filter. It is possible to perceive a balance, a proportion, a sense of relevant scale - all the information is divided between tables of equal size and all the tables were granulated to the card so that it was possible to easily locate and study each card. It was displayed an accessible complexity of detail -

showing all dashboards, datasets and cards in a detailed and easily accessible way. And finally, all elements of the Metadata Dashboard were drawn in a professional manner, with the technical details of production done with care and avoid content-free decoration, including in this case the excessive use of colours, explanations or chart junk.

3.4.2 Economic Results

In relation to the economic impacts of the project on the company, it is important to understand that the goal was always to improve the department's internal management in view of the projects already developed and their organization. Thus, through the development of the metadata dashboard, it was possible to facilitate the location, comprehension and preservation of all the digital records as intended and proposed from the beginning. It is foreseen that using the developed dashboard the company will be able to substantially reduce the number of duplications of content. It is expected to put an end to what happened before: where work already carried out was requested to the department. Thus, with the improvement in the accuracy and localization of the information it will, for sure, reduce the allocation of human resources (Business Intelligence analysts or internal dashboard owners in every non-BI department) to develop dashboards that had already been developed.

Furthermore, one of the main issues that is difficult to measure is the gain in terms of time for locating the desired documents. Thus, the loss of time looking for information is a topic that is expensive for the company because during this period the employee can be focused on working in his business area without being lost looking for necessary information, usually having to ask third parties for that same information.

Finally, one of the department's main expenses for the normal exercise of its functions, as mentioned above, is the use of the DOMO software. According to the contract made

with the company, the amount payable is defined by price intervals in view of the use of calculation lines used. That is, the more lines used, the higher the cost charged to the company. With the Metadata project carried out, which facilitates access to information, avoiding duplication of information and preserving the department's records, it is to be expected that the use of calculation lines will naturally reduce and that the cost attributed to the company should be lower than it was previously.

In summary, despite being economic gains of difficult precision and measurability, this internal project is characterized as being quite useful and it is important to highlight the importance of such an instrument for the management of the BI & Analytics department and the company's information in relation to data.

4. Conclusion

The use of data in everyday business is undoubtedly an activity that is transforming the business sector and, consequently, the entire population. This is a situation that is not expected to slow down in the future, quite the opposite. The increase in the ability to collect, process and analyse data tends to be more and more specialized and useful.

In a nutshell, information has power. The power to better understand customers to sell the right product or in greater quantity, the power to better understand the company to reduce costs, the power to better understand employees and the reasons that may anticipate their departure reducing the turnover, and many others. In all areas, the more information we have, the more confident we are that we will make better decisions. However, this information cannot only be relevant in terms of quantity but, more importantly, in terms of quality and credibility. Clean, verified and organized information is what brings the real benefits to the company.

In this sense, the project developed aimed to identify what was known, to date, about the potential of data and metadata for the development and competitiveness of

companies. In the case of EDP, the company has been investing heavily in data analysis teams to help generate valuable information for the entire company, avoiding silos and creating a transparent flow of information.

It is believed that the increase in information will lead to the need for more and more this type of metadata projects to organize and improve management of internal information in companies. These projects will be important because, as in the Metadata Dashboard of all the projects developed by the BI & Analytics department, they go after the information and certify its authenticity and veracity and after that, they organize the information and adapt it through of specialized visualization techniques and Business Intelligence software so that your content is visually easy to understand. The main objective of projects like this will always be the improvement of the organization of information and its credibility.

In the case of the project developed, it allowed the improvement of the internal management of the department. The dashboard was a solution to some department problems related to the duplication of information and credibility and organization of the data. The metadata project also took advantage of the reduction of indirect costs associated with the time spent by employees searching for information and direct costs associated with the payment of the Software Domo.

5. References

Brinch, M., Gunasekaran, A., & Fosso, S. (2021). Firm-level capabilities towards big data value creation. *Journal of Business Research*, *131*(July 2020), 539–548. https://doi.org/10.1016/j.jbusres.2020.07.036

Fleissner, B. B., Jasti, K., Ales, J., & Thomas, R. (2014). *The Importance of Data Governance in Healthcare*. 1–11.

Furner, J. (2020). Definitions of "Metadata": A Brief Survey of International Standards. *Journal of the Association for Information Science and Technology*, *71*(6), E33–E42. https://doi.org/10.1002/asi.24295

Gilliland Anne J., Baca Murtha (2016). Introduction to Metadata, Getty Research Institute

Greenberg, J. (2012). Understanding metadata and metadata schemes. *Metadata: A Cataloger's Primer*, *9374*(2005), 17–36. https://doi.org/10.4324/9780203051238-5

Grossman W., Rinderle-Ma S. (2015). Fundamentals of Business Intelligence, Springer Publishing Company

Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, *154*, 72–80. https://doi.org/10.1016/j.ijpe.2014.04.018

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, *53*(1), 148–152. https://doi.org/10.1145/1629175.1629210

Kumar, SM., & Belwal, M. (2017). Performance Dashboard Cutting-Edge Business Intelligence and Data Visualization. 2017 International Conference On Smart Technology for Smart Nation

Mayer-Schonberger, V., & Cukier, K. (2013). Big data: a revolution that will transform how we live, work and think, Houghton Mifflin Harcourt Publishing Company

Monino, J. L. (2021). Data Value, Big Data Analytics, and Decision-Making. *Journal of the Knowledge Economy*, *12*(1), 256–267. https://doi.org/10.1007/s13132-016-0396-2

Nestorov, S., Jukić, B., Jukić, N., Sharma, A., & Rossi, S. (2019). Generating insights through data preparation, visualization, and analysis: Framework for combining clustering and data visualization techniques for low-cardinality sequential data. *Decision Support Systems*,

125(August), 113119. https://doi.org/10.1016/j.dss.2019.113119

Orlovskyi, D., & Kopp, A. (2021). A business intelligence dashboard design approach to improve data analytics and decision making. *CEUR Workshop Proceedings*, *2833*.

Pomerantz Jeffrey (2015). *Metadata,* The MIT Press Essential Knowledge Series

Raguseo, E., & Vitari, C. (2018). Investments in big data analytics and firm performance: an empirical investigation of direct and mediating effects. *International Journal of Production Research*, *56*(15), 5206–5221. https://doi.org/10.1080/00207543.2018.1427900

Sarikaya, A., Correll, M., Bartram, L., Tory, M., & Fisher, D. (2019). What do we talk about when we talk about dashboards? *IEEE Transactions on Visualization and Computer Graphics*, *25*(1). https://doi.org/10.1109/TVCG.2018.2864903

Schoenherr, T., & Speier-Pero, C. (2015). Data science, predictive analytics, and big data in supply chain management: Current state and future potential. *Journal of Business Logistics*, *36*(1), 120–132. https://doi.org/10.1111/jbl.12082

Toasa, R., Maximiano, M., Reis, C., & Guevara, D. (2018). Data visualization techniques for real-time information - A custom and dynamic dashboard for analyzing surveys' results. *Iberian Conference on Information Systems and Technologies, CISTI*, *2018-June*, 1–7. https://doi.org/10.23919/CISTI.2018.8398641

Tufte, E. R. (2001). Aesthetics and Technique in Data Graphical Design. *The Visual Display of Quantitative Information*, 177–191. http://www.amazon.co.uk/dp/0961392142

Vedder, R. G., Vanecek, M. T., Guynes, C. S., & Cappel, J. J. (n.d.). *Perspectives Competitive Intelligence*.

Wexler, Shaffer, and A. Cotgreave. The Big Book of Dashboards: Visualizing Your Data Using Real-World Business Scenarios. John Wiley & Sons, 2017.