ANALYTICS: DASHBOARDS FOR ANALYSIS OF ACADEMIC AND ADMINISTRATIVE DATA WITHIN THE FEDERAL UNIVERSITY OF TOCANTINS

ANALYTICS: PAINÉIS DE ANÁLISE DE DADOS ACADÊMICOS E ADMINISTRATIVOS NO ÂMBITO DA UNIVERSIDADE FEDERAL DO TOCANTINS

Glauco Batista de Sousa
ORCID 0000-0000-0000-0000

Universidade Federal do Tocantins
Palmas, Brasil
glauco.sousa@uft.edu.br

Ary Henrique Morais de Oliveira
ORCID 0000-0001-5484-870X

Universidade Federal do Tocantins
Palmas, Brasil
glauco.sousa@uft.edu.

Abstract. Institutions undergo constant changes in the globalized and dynamic world, generating new positions and perspectives. Reinventions are essential to keep up with the dynamism of the market and the new social demands of customers. Therefore, many institutions are adopting digital transformation as a tool to improve processes and services. Data analysis has been shown to be effective and necessary to achieve institutional governance and management objectives. The general objective is to propose the implementation of an academic and administrative data analysis solution to promote digital transformation, monitoring of indicators and social control at the Federal University of Tocantins. The methodological approach adopted was qualitative, since data collection took place in the natural environment, enabling a deep understanding of reality. The research was applied, seeking to produce knowledge to solve specific problems. The objective of the study was explanatory, as it aims to identify factors that cause a specificity, deepening the understanding of reality. The project's outlook was positive, as UFT has highly qualified technical and human resources in the areas of education, information technology and public management, in addition to a specialized structure of its own environments for hosting systems (Data Center), which subsidized the part with the highest direct financial cost of the project. As a result, a tool was proposed for monitoring academic and administrative data, facilitating decision-making and helping to prevent problems, such as minimizing failure, retention and dropout rates at the University, in addition to enabling the creation of new public policies. As a product of the project, the proposal for the Data Management and Governance Policy, the UFT Open Data Plan and a tool with graphic panels for analyzing and monitoring academic and administrative data, with the availability of open data, were developed.

Keywords: data analysis; digital transformation; open data.

Resumo. As instituições passam por constantes mudanças no mundo globalizado e dinâmico atual, gerando novos posicionamentos e perspectivas. As reinvenções são essenciais para acompanhar o dinamismo do mercado e as novas demandas sociais dos clientes. Por isso, muitas instituições estão adotando a transformação digital como ferramenta para aprimorar processos e serviços. A análise de dados tem se mostrado eficaz e necessária para atingir objetivos de governança e gestão institucional. O objetivo geral é propor a implantação de uma solução de análise de dados acadêmicos e administrativos para promover a transformação digital, o acompanhamento de indicadores e o controle social na Universidade Federal do Tocantins. A abordagem metodológica adotada foi qualitativa, pois a coleta de dados ocorreu no ambiente natural, possibilitando uma compreensão profunda da realidade. A pesquisa foi aplicada, buscando produzir conhecimentos para a solução de problemas específicos. O objetivo do estudo foi explicativo, pois visou identificar fatores que causam um fenômeno, aprofundando o entendimento da realidade. A viabilidade do projeto foi positiva, pois a UFT conta com recursos técnicos e humanos altamente capacitados nas áreas de educação, tecnologia da informação e gestão pública, além de uma estrutura especializada de ambientes próprios para hospedagem de sistemas (*Data Center*), o que subsidiou a parte de maior custo financeiro direto do projeto. Como resultados, foi proposta uma ferramenta para o acompanhamento de dados acadêmicos e administrativos, facilitando a tomada de decisões e auxiliando na prevenção de problemas, como a minimização dos

índices de reprovação, retenção e evasão na Universidade, além de possibilitar a criação de novas políticas públicas. Como produto do projeto, foi desenvolvida a proposta da Política de Gestão e Governança de Dados, o Plano de dados abertos da UFT e uma ferramenta com painéis gráficos para análise e acompanhamento dos dados acadêmicos e administrativos, com a disponibilização dos dados abertos.

Palavras-chave: análise de dados; transformação digital; dados abertos.

1. INTRODUCTION

Institutions face constant changes in a globalized and dynamic world, which requires the continuous creation of new positions and perspectives. Strategic reinvention is essential to keep pace with the fast-moving market and meet the emerging social demands of clients. In this context, many organizations are adopting digital transformation as a key tool to enhance their processes and services, seeking an integrated and sustainable vision that shapes a future more aligned with business needs (Rogers, 2024).

Organizational Processes suggest that entities need to establish effective organizational structures and processes, ensuring compliance with legal responsibilities by both individuals and the entity, evaluating their performance, ensuring proper accountability of public funds, maintaining clear communication with stakeholders, and clarifying the roles and responsibilities of management, especially the roles and responsibilities of non-executive members of the governing and executive management bodies.

Public administration understands digital transformation as a way to offer high-quality public services while reducing costs, both financial and time-related for citizens, thus improving the relationship between citizens and the public services provided. Digital transformation in the public sector significantly reduces the annual expenses of those involved in the process and enhances the service provided to citizens, making the relationship between government and citizens more direct and transparent (Brazil, 2023a).

Rogers (2017) points out that digital technologies have changed the world by the way data is considered. Data was expensive, from acquisition to storage, in traditional models that required large Information Technology (IT) systems to map the data flow. Today, data is generated at an accelerated rate, both by businesses and institutions, as well as by society in general, in various locations in real-time. Therefore, digital technologies transform the way organizations innovate.

Digital Transformation involves the adoption of digital technologies to transform how organizations manage their processes, organizational culture, and ability to adapt to market changes. In this scenario, data becomes one of the most valuable assets of an institution. However, when not properly managed, data may not meet the strategic objectives required to support important decision-making. Therefore, activities such as data collection, storage, processing, and analysis are essential to ensure efficient management and governance, promoting more informed decisions that align with institutional goals.

The presentation, availability, and reuse of public data can be accessed by all interested parties through open data. These data are considered open when anyone can use, reuse, and redistribute them freely, as long as proper attribution is made and the data is shared under the same license. The goal of open data is to make information accessible in reusable formats, promoting greater citizen participation in public control processes, which contributes to inclusion, transparency, and accountability.

This paper discusses the data trajectory at the Federal University of Tocantins (UFT), from the creation of data governance to the implementation of mechanisms for presenting and analyzing academic and administrative data. The objective is to monitor key indicators and promote transparency through open data, aiming for social control. Additionally, the research

focuses on structuring data presentation dashboards for Institutional Governance, with the aim of identifying opportunities for the development of new public policies and supporting decision-making by the university's senior administration.

2. THEORETICAL FRAMEWORK

The massive use of Information and Communication Technologies (ICT) has been causing transformations in all areas. In recent decades, large volumes of data, primarily digital, have been stored. However, it is necessary for this data to be transformed into useful information, so it can be appropriately applied in the construction of knowledge and the development of new public policies in Public Administration.

2.1. Digital Transformation

Digital transformation is the process of changing the behavior of organizations through the use of digital technologies to create new ways of managing business processes, organizational culture, and adapting to market changes. Digital transformation has five strategic domains that mark the transition from the analog to the digital era: Customers, competition, data, innovation, and value. These domains are interconnected to foster business growth (Rogers, 2017).

In recent years, public administration has invested resources in digital transformation, believing it to be an appropriate path to improve public services provided to society, aiming for greater ease and versatility in people's lives. The massive use of digital technologies accelerates the transformation of the services offered, serving as a philosophy of transformation in the way the government interacts with society.

The potential of digital transformation is vast, with the design and automation of high-quality services that institutions are capable of offering at lower costs, whether in terms of time or finances. This reduces the costs for the customer and makes the relationship between stakeholders more transparent and direct (Brazil, 2023a).

Public administration seeks to improve public services through digital transformation, which is happening contemporaneously. Figure 1 presents the historical evolution of digital government in its journey toward greater maturity in delivering high-quality and integrated digital public services to society.

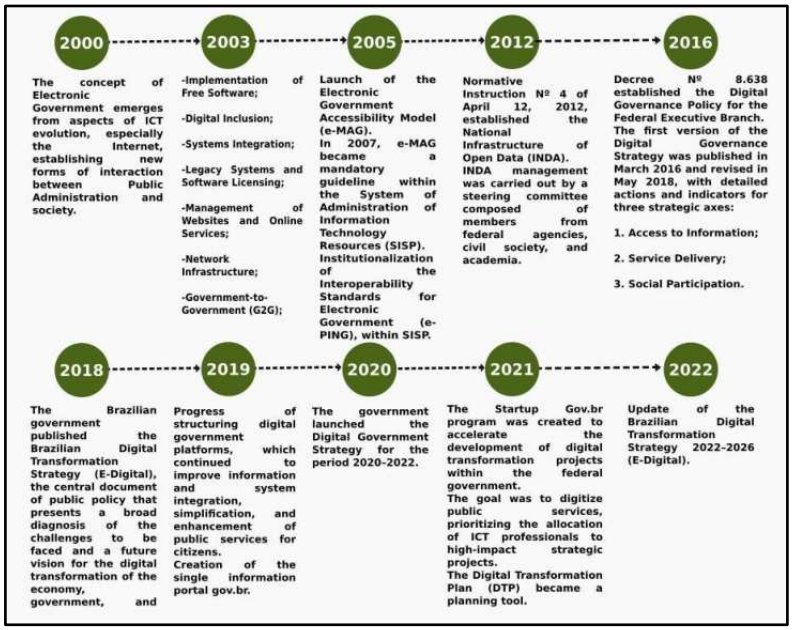


Figure 1. Evolution of Digital Government
Source: Authors (2024) adapted from Brazil (2023b)

The evolution of digital government, within this scope, began in the year 2000, when the concept of Electronic Government or e-gov was adopted, adding electronic aspects to governmental strategies. This marked the initial step toward finding new ways to engage with society. Therefore, Figure 1 summarizes the historical evolution from 2000 to 2022, focusing on regulatory milestones that contributed to the increased maturity in the area of digital transformation.

2.2. Data Fundamentals

Data, like the organizations that structure their management processes across strategic, tactical, and operational areas, are also divided into two main segments: Data Governance and Data Management. Data Governance focuses on strategies related to the use and control of data within the organization, establishing guidelines and policies. On the other hand, Data Management focuses on data processing, from source collection to the presentation of information through data intelligence dashboards.

According to Barbieri (2020), the biggest challenge regarding data is achieving an organizational perspective on these assets, ensuring that they are produced, analyzed, consumed, and managed correctly. In other words, the differentiating factor is the data managers, not the hardware or software. Therefore, it is necessary to organize the competencies and roles of each actor in the data governance and management stage, especially in public administration, with well-defined areas and responsibilities.

The difference between Data Governance and Data Management lies in the fact that the former provides direction, monitoring, supervision, and evaluation of data management practices. In other words, it establishes the guidelines and policies to be followed by Data Management, which, in turn, is inherent to and integrated with organizational processes, being responsible for planning, execution, control, and monitoring. In summary, Data Management pertains to the manipulation of the resources and powers made available to agencies and entities to achieve their objectives. Generally speaking, Data Management executes the policies and guidelines defined by Data Governance (Brazil, 2023d).

For the fundamentals of Data Governance, it is important to highlight that efforts are focused on organizing and integrating the data and metadata of organizations in a way that aligns with business needs, ensuring they are unique, integral, reliable, maintainable, known, performant, readable, and accessible to those who truly need access (Rêgo, 2013).

Rêgo (2020) adds that Data Management, in the broadest sense of the definition, is the discipline responsible for defining, planning, implementing, and executing strategies, procedures, and practices necessary for effectively managing an organization's data and information resources. This includes plans for defining, standardizing, organizing, protecting, and utilizing data.

However, "Data is not just data." In other words, data does not appear magically, perfectly aligned with the needs of those consuming it. Data is often the result of long construction processes involving various methodological decisions. Furthermore, the consumption of data is not immediate. Those who read or analyze the data must also be trained to understand it. Even experienced researchers can fall into the many pitfalls that a new and unfamiliar dataset typically presents (Shikida, Monasterio, 2021).

2.3. Business Intelligence with Data from Federal Public Administration

Business Intelligence (BI) in federal public administration involves using data to improve decision-making and optimize public service management. By utilizing data analysis tools and techniques, BI enables public entities to extract valuable insights from available data, supporting policy formulation and enhancing service efficiency.

The implementation of BI solutions in public administration includes collecting and analyzing large volumes of data from different sources, such as administrative records, surveys, and performance data. The use of dashboards and analytical reports facilitates the monitoring and evaluation of public services, enabling the identification of areas for improvement and more informed decision-making. Data integration and predictive analysis also allow public managers to anticipate problems and develop proactive strategies to address challenges (Sharda, Delen & Turban, 2019).

Business Intelligence is a powerful tool for public administration, offering an in-depth view of operations and facilitating data-driven decision-making. Adopting BI practices can lead to more efficient public service management, greater transparency, and better responsiveness to the population's needs. Investing in BI solutions and a data-driven culture is essential to improve the quality and effectiveness of public administration.

The National Digital Government Strategy (ENGD) presents data intelligence as the fifth of the ten specific goals declared in the strategy, aiming to qualify decision-making and service delivery in public organizations through the constant and ethical reuse of available data for analysis, interoperability, and personalization (Brazil, 2024e).

The table below presents the ENGD recommendations for Goal 5 - Data Intelligence:

Table 1. Recommendations for Goal 5: Data Intelligence

Recommendations for Goal 5: Data Intelligence	
Recommendation 5.1	Develop, publish, and implement a data governance program.
Recommendation 5.2	Establish and adopt mechanisms for interoperability and data sharing, both among government agencies and with other federative entities, especially those provided by the GOV.BR Platform, to enhance public policies and eliminate unnecessary data requests in the provision of public services.
Recommendation 5.3	Contribute to the development and adoption of a data-sharing model that enables citizens to use their data securely and improves their experience when accessing services.
Recommendation 5.4	Encourage the cataloging of data in order to promote the discovery and reuse of data.
Recommendation 5.5	Promote the use and adoption of data analysis and data science, ethically, in decision-making for public policies and in the personalization of services.

Source: Authors (2024) adapted from Brazil (2023e)

The recommendations presented in Table 1: Recommendations for Goal 5: Data Intelligence aim to clarify that one of the most valuable inputs in the digital age is data, which can be used to add value to services provided to citizens in various areas. There are significant and necessary advantages to basing public sector activities and services on the analysis and reuse of data: information is automatically compared and made available, and bureaucracy is minimized.

2.4. Data Analysis

The interpretation of data requires knowledge, skills, and quality data in order for analyses to be efficient. Therefore, the involvement of business areas is crucial in generating data insights through data analysis.

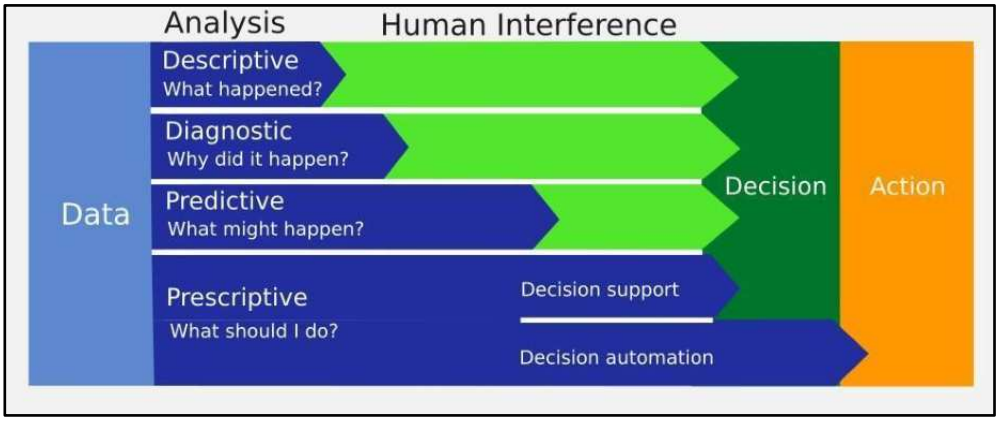


Figure 2. Data Analysis and Human Interference
Source: Ferreira Junior (2024)

Figure 2 presents the evolution of data analysis and the level of human interference that each type of data analysis can provide within a dataset.

Based on the contributions of Sharda, Delen, and Turban (2020), types of analysis can be categorized as follows, as presented in Table 2:

Table 2. Types of Data Analysis

Analysis	Description
Descriptive	This type of analysis aims to understand the institution and identify trends and causes of past occurrences. The main questions this analysis seeks to answer include: "What happened?" and "What is currently happening in the institution?". Answers to these questions can be obtained through business reports, dashboards, scorecards, and data warehouses. Descriptive analysis provides a detailed and contextual view of past and current operations, helping to build a solid foundation for future decision-making.
Predictive	Predictive analysis focuses on identifying patterns and trends that can help forecast future events. The central questions are: "What is likely to happen in the future?" and "Why are these events expected?". This analysis is often performed using data mining techniques, which explore large volumes of data to identify patterns and make informed predictions about future events or behaviors.
Prescriptive	Prescriptive analysis seeks to provide recommendations on actions to take based on analytical models. The main questions addressed in this analysis are: "What should I do?" and "Why should I adopt this approach?". This analysis involves process optimization, conducting simulations, and creating models and expert systems to support decision-making. It offers guidelines on how to deal with identified situations and predicts the outcomes of proposed actions, aiming to improve performance and achieve desired goals.

Source: Sharda, Delen, e Turban (2020)

Each type of analysis plays a crucial role in understanding and managing an institution's operations, offering tools and insights that range from describing the current state to forecasting future events and recommending specific actions to improve organizational performance.

2.5. Open Data

Currently, a large volume of data can be accessed openly on the internet. This has been made possible, mostly due to the availability of public data in an open format that can be accessed and processed using information technology resources. This enables active citizen participation in the development of initiatives to monitor, oversee, and contribute to the

improvement of public resource management. The legal framework has been working towards the creation of laws to guide institutions in opening data (TCU, 2015).

In the Public Administration, the Federal Comptroller General (CGU), through the National Open Data Infrastructure (INDA), was established with a managerial and normative role, capable of setting complementary rules related to this matter. Decree 8777/2016 established the Open Data Policy of the Federal Executive Branch, and Decree 9903/2019 determined the CGU, through INDA, as responsible for managing Open Data in the Federal Executive Branch (Brazil, 2023f).

The implementation of open data in institutions is crucial to promote transparency, innovation, and collaboration. By providing accessible and usable data, these institutions can not only strengthen public trust but also stimulate interdisciplinary research and the development of innovative solutions to complex problems. Open data allows researchers, students, and the general public to analyze and use information more efficiently, facilitating scientific discoveries and the creation of evidence-based policies (Brazil, 2023g).

2.6. Statistical Data

Statistics refers to methods or sets of methods that have been adapted to collect, explore, describe, and interpret numerical datasets. The foundation of statistics lies in a dataset that is intended to be structured and interpreted. The result of observing the data depends on its characteristics, objects, entities, or attributes that constitute the reality (Silvestre, 2007).

The problem of analyzing and understanding a relevant dataset for a specific group of studies requires transforming the data into information to compare them with other results or apply them to a given theory. This can be done through statistical inference, a part of statistics that is continuously a methodology in science aimed at collecting, reducing, analyzing, and modeling data, classified by scientists to test theories and hypotheses (Bussab & Moretin, 2017).

Statistical data are essential for making informed decisions and developing effective public policies. They provide an objective foundation for understanding complex phenomena, identifying trends, predicting outcomes, and evaluating the impact of implemented actions. Therefore, proficiency in statistical techniques and a data-driven culture are crucial for enhancing the quality and relevance of decisions, significantly contributing to the advancement of knowledge and innovation.

3. METHODOLOGY – METHODS

The methodological approach adopted was qualitative, as data were collected in a natural environment, allowing for an in-depth understanding of the reality, with descriptions, interpretations, and assignment of meanings, investigating concepts, opinions, and practices. The research was applied in nature, as it sought to produce knowledge for practical application directed at solving specific problems. The objective of the study was explanatory, aiming to identify the factors that cause a particular phenomenon, deepening the understanding of reality (Mazzaro, 2016).

The methodology was divided into 7 stages: Literature review, meetings and data gathering from the business environment, research on open data and data analysis software in higher education institutions, creation of the descriptive data model, construction of the Data Governance and Management Policy and the Open Data Plan, prototype development of the data analysis dashboards software, and software registration.

4. RESULTS

As results, academic and administrative databases of the Federal University of Tocantins (UFT) were collected, proposed documents for the Data Governance and Management Policy were created, the UFT Open Data Plan was drafted, and a tool for visualizing and monitoring academic data through graphic dashboards was developed, thus facilitating decision-making and the availability of open data, streamlining the addressing of potential problems as well as the creation of new public policies and social control.

4.1 Literature Review

As results, several meetings were held to guide the work with the advisor, and a bibliographic review of the data field and its surroundings was conducted, as well as training related to the proposed project theme, primarily at the National School of Public Administration (ENAP), with the results presented in the Literature Review section

4.2 Meetings and Data Gathering from the Business Environment

Meetings were held with the institution's core departments managing data in their respective Pro-Rectories, listing representatives from the Pro-Rectories of Undergraduate Studies, Student Assistance, Administration, Planning, Research, and Extension to gather the most commonly used datasets by each area. The gathered data were compiled along with the data suggested for publication by the CGU for educational institutions.

Two sets of data were generated: the priority sets for the business areas and the sets of data suitable for release, as well as the UFT's own datasets.

4.3 Data Governance and Management Policy and Open Data Plan

The proposal for the Data Governance and Management Policy is crucial to establish the structural rules for data. The policy outlines principles, guidelines, assignments, and responsibilities for the management and governance of data, information, and knowledge, ensuring efficient organization within the data ecosystem. The policy allows for informed decision-making with greater security in processes, supported by high-quality data.

The structure proposed in the Data Governance and Management Policy of UFT includes the following actors: the UFT Digital Governance Committee (CGD), the Information Management Committee (CGI), the Board of Trustees, and the technical data team.

The proposed Open Data Plan for the Federal University of Tocantins was drafted considering the Open Data Policy established by Decree 8777/2016 and other regulations, which define rules to promote the opening and analysis of government data within federal agencies and entities, such as ministries, agencies, regulatory bodies, and public foundations.

After the data collection, the following results were obtained: 313 requests for access, 32 responses in 12 databases, and 76 datasets, as shown in the dashboard in Figure 1: UFT Priority Databases.

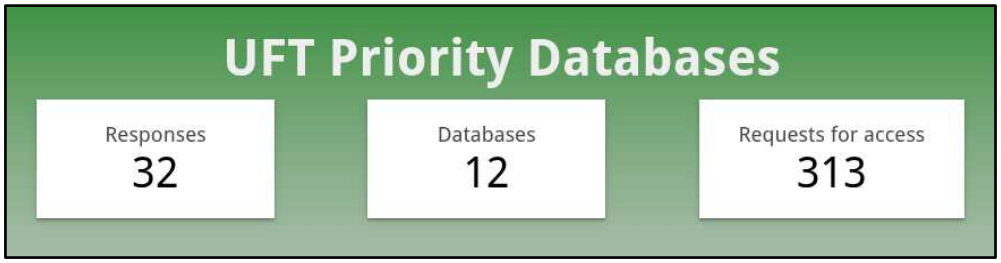


Figure 3. UFT priority databases
Source: Authors (2024)

The databases were prioritized according to the datasets they contained, as detailed in Table 1 - Breakdown of the responses from the UFT datasets. This table shows the top 10 votes from the datasets ordered in descending order. The votes were recorded in the databases with marking in the corresponding datasets. Data compilation was carried out in the databases.

Table 3. Detailed Analysis of UFT Dataset Responses

Order	Database	Data Sets	Number of Responses
1	Students	Number of Students per Course	11
2	Academic	Undergraduate and Graduate Academic Information	11
3	Academic	Undergraduate Enrollment	11
4	Academic	Electronic Academic Journals	10
5	Administrative	Selection Processes	9
6	Students	Socioeconomic Analysis of Students	8
7	Professors	Professor's name	7
8	Administrative	University restaurant menu	7
9	Administrative	SIE organizational chart	7
10	Academic	Scientific initiation project management	7

Source: Authors (2024).

The respondent could select up to 15 predefined response options for each dataset provided in the questionnaire. All options were listed alphabetically, following the categorization pattern by the dataset's name/identifier, along with a subjective field for any other suggestions regarding data openness.

After the prioritization of the data, dates for data release were established, and the document is now ready for submission to the CGU for suggestions and verification to ensure compliance with the INDA guidelines, as outlined in the PDA Preparation Manual. It will then be reviewed and approved by internal bodies, usually the Digital Governance Committee (CGD) of the institution.

4.4 Technological Prospecting

Regarding the technological prospecting of the proposed software, an initial search was conducted in the database of the Brazilian Industrial Property Institute (INPI). The first term used was “data,” which returned a total of 522 computer program records, revealing a diversity of information about data in general (BRASIL, 2023c).

The second term used was “analysis,” which returned a total of 507 computer program records, indicating a variety of topics on different types of analyses, with many systems being predominantly web-based.

Subsequent terms included “Data Analysis” and “Analysis of Data” both of which returned 38 computer program records. In this case, the topics were noted to be more specific to data analysis, with a focus on various fields of knowledge. So far, a low number of registered software programs has been observed in the area of data analysis, particularly in the field of educational data analysis.

4.5 Technological Infrastructure

The proposed infrastructure, as illustrated in Figure 2 - Prototype of the Data Infrastructure Architecture, involves creating a data repository to handle data flow without compromising the performance of the official applications that send and consume data from the university's

databases. Additionally, it aims to structure the data for presentation on dashboards, improving the processes of extracting, transforming, and loading data.

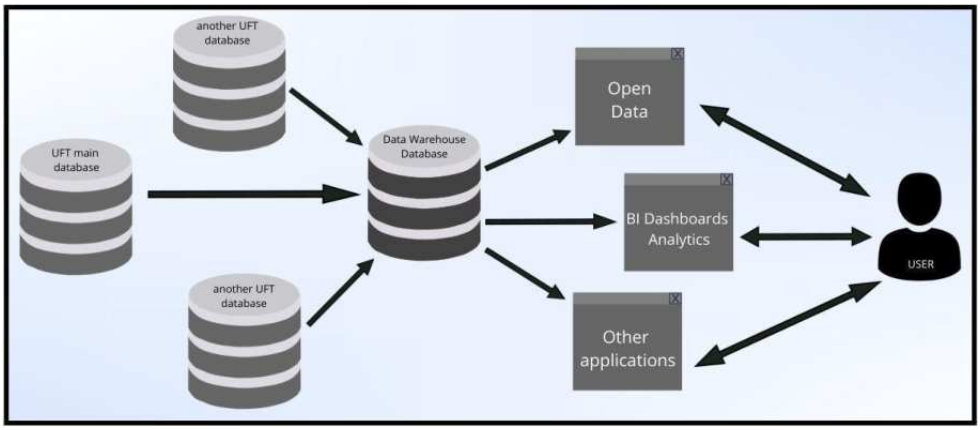


Figure 4. Prototype of the Data Infrastructure Architecture
Source: Authors (2024)

Figure 4 illustrates the Prototype of the Data Infrastructure Architecture, outlining the entire data lifecycle and interconnection, from the consumption of UFT's databases to the creation of the data warehouse and the data flow to applications that interact with users.

A database was created using the open-source object-relational PostgreSQL Database Management System (DBMS), which uses and extends the Structured Query Language (SQL) and incorporates many features to securely store and scale even the most complex data workloads.

The foundation of the project is the Data Warehouse, designed to centralize the storage and consolidation of data received from various application-related databases.

4.6 Software Development

The proposed software was developed to serve as a central tool for presenting and analyzing UFT's data, aiming to integrate the institution's most relevant information. The development began with a test version and project sketches.

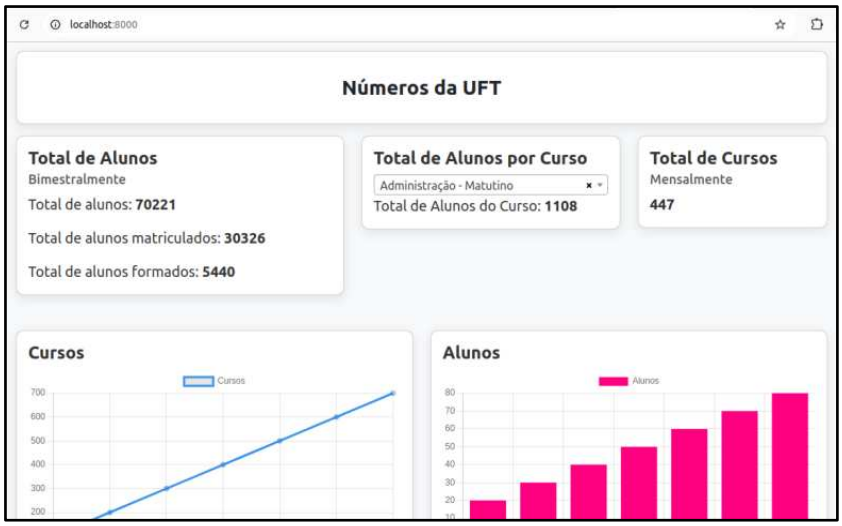


Figure 5. Initial Version of the UFT Numbers Project
Source: Authors (2024)

Figure 5 illustrates the initial prototype version of the UFT Numbers project, which featured a card-based design for organizing visual elements and information about the institution. It also included charts to visually depict data for comparison purposes. The first version prioritized a minimalist color palette, aiming for simplicity with sophistication.

As the project evolved and new components were incorporated, a redesigned data environment and layout were developed, leading to the renaming of the project as UFT More Numbers. This updated version was published in a beta testing environment. Following a presentation of the project during a UFT management meeting at the university's Digital Governance Committee, it was suggested to rename the project UFT in Numbers, which became its official title.

The project team conducts weekly meetings to address software development, including identifying impediments and proposing solutions to challenges, as well as revisiting priorities for the project's ongoing evolution. After developing a version of the Minimum Viable Product (MVP), the first published version of the project was made available in a staging environment.



Figure 6. Staging Version of the UFT in Numbers Project
Source: Authors (2024)

The project progressed with internal meetings, and presentations were initiated for representatives of the university's Pro-Rectors, as well as external stakeholders familiar with data ecosystems. For instance:

- Professor Dr. José Jorge Júnior, coordinator of the Ecograd project, which presents a data ecosystem for undergraduate studies among Federal Institutions of Higher Education (IFES) associated with Andifes.
- Federal External Control Auditor Daniel Jezini, who provided valuable advice for the project and indicated that it is on a promising path regarding institutional audit insights.
- Dr. Jeferson Alves dos Santos, Head of Internal Audit at UFABC, contributed with insights from his experiences with data and transparency, offering a legal framework perspective for the project.

After these presentations and subsequent refinements, the project was published in a production environment, retaining the visual characteristics of the project's MVP while incorporating conceptual changes informed by internal and external feedback.



Figure 7. Production Version of the UFT in Numbers Project
Source: Authors (2024)

Figure 7 shows the initial screen of the UFT Numbers project in the production environment, focusing on mature data—validated by business areas. The software also includes other dashboards with filtering options, enabling new data insights and perspectives.

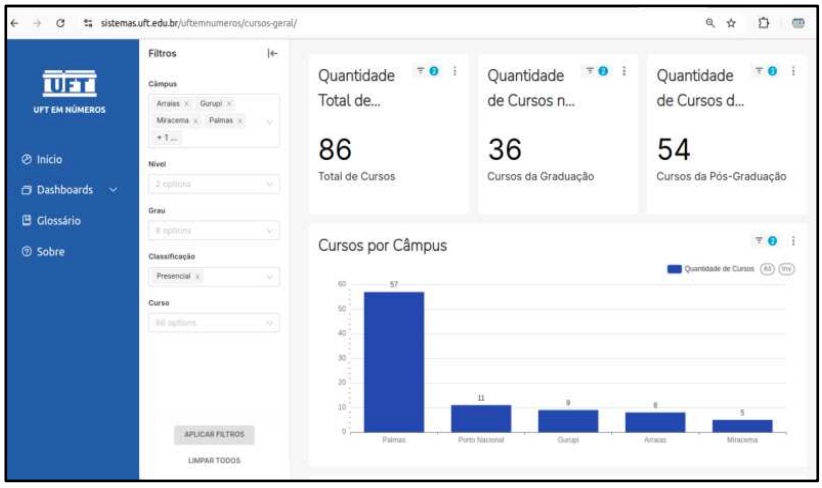


Figure 8. Production Version of the UFT in Numbers Project with filter
Source: Authors (2024)

Figure 8 shows the courses screen, including basic information and the ability to apply filters using the side panel.

4.7 Data Ecosystem

Based on all prior discussions, a simplified model of the processes for building a data ecosystem in higher education institutions is presented to guide their integration into a comprehensive data ecosystem.

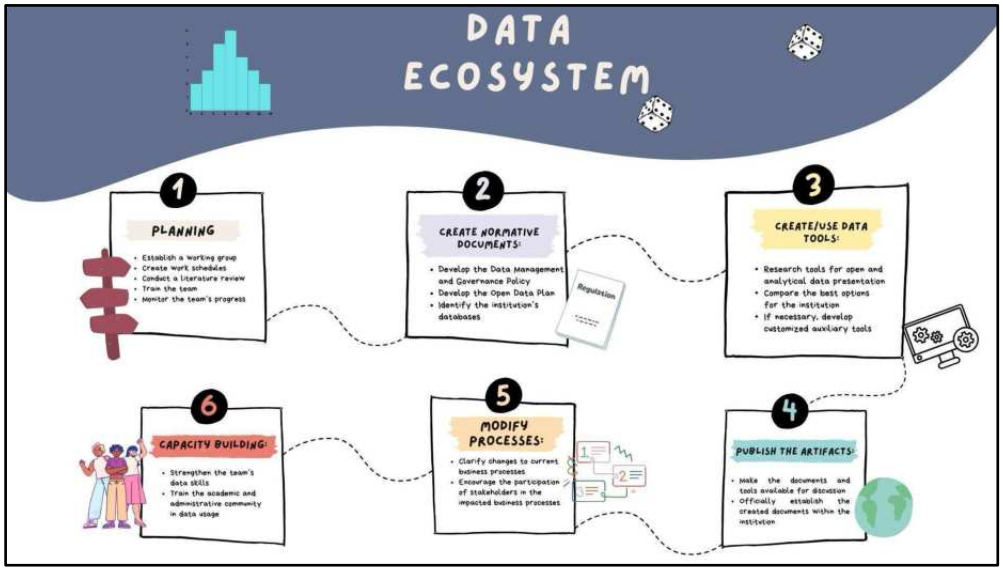


Figure 9. Data Ecosystem
Source: Authors (2024)

Figure 9 outlines six global steps for public higher education institutions to adopt and utilize a data ecosystem:

Step 1:

- Form a working group: Recruit individuals from diverse areas to compose the data management team.
- Create work schedules: Develop well-defined timelines for each team to collect and analyze the necessary information.
- Conduct a literature review: Understand relevant legislation, explore academic research, institutional websites, and gather techniques and practices from other organizations.
- Train the team: Compile and share the collected information with all teams to disseminate knowledge and facilitate project understanding.
- Monitor team progress: Address barriers encountered by each area, ensuring smooth progression in data expertise.

Step 2:

- Develop a Data Management and Governance Policy: At the strategic level, create an institutional policy to support data-driven decision-making with a structured and standardized framework.
- Prepare an Open Data Plan: At the tactical level, design the plan in accordance with the National Open Data Policy (established by Decree 8777, May 11, 2016) and notify the CGU upon publication.
- Inventory institutional databases: As a foundational step, catalog all institutional databases for use in data-related activities.

Step 3:

- Research tools for open and analytical data presentation: Explore tools for data presentation and identify options aligned with the institution's strategic and tactical goals.
- Compare the best tools for the institution: Assess available tools for compliance with the Data Management and Governance Policy and the Open Data Plan.
- Develop customized auxiliary tools if necessary: If existing tools are insufficient, establish a technical working group to create or enhance data analysis tools.

Step 4:

- Share documents and tools for discussion: Distribute initial drafts institution-wide for feedback and suggestions, and obtain approval from all relevant areas.
- Formalize the created documents within the institution: Submit strategic documents, such as the Data Management and Governance Policy, to the institution's highest council and tactical documents, such as the Open Data Plan, to the Digital Governance Committee.

Step 5:

- Clarify changes to current business processes: Clearly communicate the reasons for any process changes resulting from the Policy or Plan to ensure stakeholder buy-in.
- Encourage stakeholder participation: Involve affected business areas to foster collaboration and minimize resistance.

Step 6:

- Strengthen the data team: Continuously train and update the technical data team to support and guide other teams.
- Educate the academic and administrative community on data use: Provide comprehensive information about the data and methodologies to ensure broader understanding and adoption.

5. CONCLUSION

The project has demonstrated positive outcomes, showcasing significant relevance to the institution with clear and achievable goals, considering the available resources for its development.

It is anticipated that, in a short time, the solution will deliver substantial benefits to the university, providing high-value insights for institutional knowledge from a new perspective that supports decision-making and fosters innovative development.

Moreover, implementing the solution promotes integration across different university areas, encouraging interdisciplinary collaboration. This synergy will not only enhance decision-making quality but also foster idea-sharing and creative exchanges within the academic community, creating a more dynamic and innovative environment.

As the university adapts and innovates, it will be better positioned to address contemporary challenges effectively, ensuring its relevance and competitiveness in academic, administrative, and social contexts. Thus, the project will benefit not only its direct stakeholders but also strengthen the institution's mission and long-term objectives.

REFERENCES

Barbieri, C. (2020). *Governança de Dados: Práticas, conceitos e novos caminhos*. Rio de Janeiro: Alta Books.

Brasil, M. G. I. S. P. (2023a). *Transformação Digital o que é?.* Recuperado de: https://www.gov.br/governodigital/pt-br/transformacao-digital/o-que-e

Brasil, G. D. (2023b). *Do Eletrônico ao Digital.* Recuperado de: https://www.gov.br/governodigital/pt-br/estrategia-de-governanca-digital/do-eletronico-ao-digital

Brasil, I. N. P. I. (2023c). *Consulta à Base de Dados do INPI.* Recuperado de: https://busca.inpi.gov.br/pePI/jsp/programas/ProgramaSearchBasico.jsp

Brasil, G. D. (2023d). *Governança e Gestão de Dados.* Recuperado de: https://www.gov.br/governodigital/pt-br/infraestrutura-nacional-de-dados/governancadedados

Brasil, G. D. (2023e). *Inteligência de Dados.* Recuperado de: https://www.gov.br/governodigital/pt-br/estrategias-e-governanca-digital/estrategianacional/objetivo-05

Brasil, C. G. U. (2023f). *Dados Abertos.* Recuperado de: https://www.gov.br/cgu/pt-br/acesso-a-informacao/dados-abertos

Brasil, C. N. A. (2023g). *Resolução nº 3/2017.* Recuperado de: https://www.gov.br/conarq/pt-br/legislacao-arquivistica/resolucoes/resolucao-n-o-3-de-13-de-outubro-de-2017

Bussab, W., & Moretin, P. (2017). Estatística Básica. São Paulo: Editora Saraiva.

Ferreira Junior, J. R. (2021). *Análise Descritiva, Prescritiva e Preditiva.* Recuperado de: https://www.linkedin.com/pulse/an%C3%A1lise-descritiva-prescritiva-preditiva-jose-r-f-junior/

Mazaro, R. E. (2016). *Metodologia da Pesquisa Científica*. Valinhos.

Rego, B. L. (2013). *Gestão e Governança de dados: promovendo dados como ativo de valor nas empresas*. Rio de Janeiro: Editora Brasport.

Rego, B. L. (2020). *Simplificando a governança de dados: governe os dados de forma objetiva e inovadora*. Rio de Janeiro: Editora Brasport.

Rogers, D. L. (2017). *Transformação Digital: repensando o seu negócio para a era digital*. São Paulo: Autêntica Business.

Rogers, D. L. (2024). *Transformação Digital 2: um roadmap para superar os obstáculos e implementar a transformação digital de forma contínua na sua organização.* São Paulo: Autêntica Business.

Silvestre, A.L. (2007). *Análise de Dados e Estatística Descritiva.* São Paula: Escolar Editora.

Sharda, R., Delen, D., & Turban, E. (2019). *Business Intelligence e Análise de dados para gestão do negócio.* Porto Alegre: Bookman.

Sharda, R., Delen, D., & Turban, E. (2020). Analytics, data science, & artificial intelligence: systems for decision support. São Paulo: Pearson.

Shikida, C. D., Monasterio, L, & Nery, P. F. (2021). *Guia brasileiro de análise de dados: armadilhas & soluções*. Brasília: Enap.

TCU, T. C. U. (2015). *5 Motivos para a abertura de dados na administração pública.* Brasília: Secom.