Master‘s Thesis

for the attainment of the academic degree
Master of Arts
of Steinbeis-University
School of International Business and Entrepreneurship

Data Mining Risk Management Applied to Accountability

Vládimir B. de Araújo, Post-graduate

SCHOOL OF INTERNATIONAL
BUSINESS AND ENTREPRENEURSHIP
STEINBEIS UNIVERSITY

M.A. in General Management

Year 2019/2021

Data Mining Risk Management Applied to Accountability

Author:

Vládimir B. de Araújo, Post-graduate

1249-1600109

Period of project work:

01.04.2109 - 10.12.2021

Primary examiner:

Prof. Dr. Jör Fuß, SHB

Secondary examiner:

Dr. Ardin Djalali, SHB

Closing Date: 11.10.2021

DECLARATION

Affidavit

I, Vládimir Borges de Araújo

born on the: 17th April 1976 in: Aracaju-SE/Brazil

hereby declare,

1. that I have independently written this assignment. I have neither made use of any other sources, tools, and resources as those given nor have I used any other unauthorized resources,

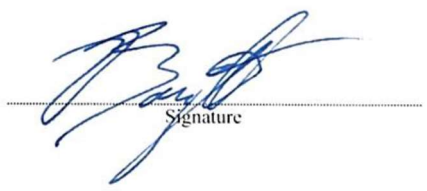
2. that I have neither at home nor abroad presented this assignment in any other form as an examination paper,

3. that I have informed my employer and obtained his understanding concerning the title, form, and content of this assignment.

4. that my paper complies with the principles of good scientific practice demanded by SIBE.

Brasília, 11/October/2021

Place, date



ACKNOWLEDGMENTS

To my parents, José Valter and Ageni, for my education. Education is the human being's most significant life-changing force. Your sacrifice to provide me first line one is the consistent seed for every success I harvest today.

To my brothers, Valter and Thaís, for the lesson in discipline and perseverance. Seeing you, made me believe I could do it.

To the Comptroller General of the Union, especially the Directorate of Social Policy and Public Security Auditing, for their generosity in showing the proper ways for efficient risk management.

To FNDE's former President, Silvio Pinheiro, for the opportunity to lead this work. To former FNDE's Head Attorney, Diogo Moraes, for first providing the concept of accountability in a "deterministic model". Thanks to his initial idea our study of a "predictive model" was legally accepted. To IT Directorate, notably to General Development Coordination, under Mr. Márcio Cunha's responsibility, at the time. And especially, to my colleagues at General Coordination of Accountability Monitoring, for their complicity and loyalty. Nobody achieves anything alone and this work couldn't exist without you.

To Matheus and André, for the brilliant work in data mining development. Your selfless commitment to transform my ideas into reality brought the project from the dream world of the impossible into the real world.

To the friends at SIBE-Brazil, especially Fábio Zimmerman, for initially recognizing project potential as master degree material.

To my little boys, Henrique and Lucas, for the inspiration to overcome limits. For you, I shall go where have never imagined possible. Everything I do is to inspire you in your own future to not limit yourself.

And last but not least, to my soulmate, Christiane, for always believing when even I hesitated. You are my shining strength inspiration, resilience, and compassion model. Love you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND DEFINITIONS

AUC	Area Under the Curve
SiGPC	Accountability Management System
PBLE	Broadband for Schools Program
BI	Business Intelligence
QUANT	Center for Quantitative Methods
CGU	Comptroller General of the Union
SIGECON	Council Management System
CV	Cross-validation
DM	Data Mining
PDDE	Direct Money in School Program
Uex	Executing Units
TCU	Federal Court of Accounts
STF	Federal Supreme Court
DIFIN	Financial Directorate
CGAPC	General Coordination of Accountability Monitoring
GPD	Gross Domestic Product
SIGEF	Integrated Financial Management Systeme
SIMEC	Integrated System of Monitoring Execution and Control
KPI	Key Performance Indicator
MEC	Ministry of Education
PNLD	National Book and Teaching Material Program
ProInfo	National Educational Technology Program
FNDE	National Fund for Education Development
PNAE	National School Feeding Program
PNATE	National Program of Support to School Transportation
SIBE	School of International Business and Entrepreneurship
SecexEDU	Secretary of External Control of Education, Culture, and Sports
DIRTI	Technology and Innovation Directorate
WBS	Work Breakdown Structure

1. INTRODUCTION AND CONTEXT

Bound to the Ministry of Education, the National Fund for Education Development (FNDE), created by Law No. 5.537, dated November 21, 1968 [1], as amended by Decree-Law No. 872 dated September 15, 1969 [2], is the major public educational policies execution enabler, with capillarity in all federative entities of the nation.

Besides innovating the government supply model, the diverse projects and programs under execution make the FNDE a reference institution in Brazilian Education: School Feeding, School Books, Money Directly to School, School Library, School Transportation, School Path, Restructuring and Acquisition of Equipment for the Public-School Network of Early Childhood Education.

To achieve improvement and guarantee a quality education, especially basic education in the public school system, the FNDE has become the main partner of the 26 states, the 5.565 municipalities, and the Federal District. In this context, the funds are divided into constitutional, automatic, and voluntary (agreements).

In short, the constitutional transfer is mandatory, in which the FNDE only calculates the portion to be transferred from its annual budgets. For the automatic transfers, there is a law and its own regulation for each public policy, which disobliges the FNDE and the beneficiary entity to make agreements: there is only adhesion. Here the beneficiary, the federative unit (a municipality or a state member), receives the resources as long as it is accountable.

Finally, there are the agreements, or voluntary transfers, which according to the Fiscal Responsibility Law, in article 25[3], are those that require a prior agreement between the parties before the federal agency sponsorship. Regardless of how the public policy is financed, the funds' transfer to a grantee creates an accountability duty to the grantor.

---

[1] **L5537**, available at: <http://www.planalto.gov.br/ccivil_03/LEIS/L5537.htm>. Accessed in: 18 set. 2019.

[2] **Del0872**, available at: <http://www.planalto.gov.br/ccivil_03/Decreto-Lei/1965-1988/Del0872.htm>. Accessed in: 18 set. 2019.

[3] **Lcp101**, available at: <http://www.planalto.gov.br/ccivil_03/leis/lcp/lcp101.htm>. Accessed in: 6 ago. 2021.

FNDE's financial scale is formidable. We must emphasize that this public agency finances "scholarships and student aid"; funding and development early childhood education; renovation, school vehicle fleet standardization and expansion for the municipal, Federal District and state networks of basic public education; continuing education in distance modality; strategy of technical and financial assistance for the federated entities as an instrument of diagnosis and educational policy planning; Broadband in Schools Program (PBLE); Direct Money in School Program (PDDE); National School Feeding Program (PNAE); National School Transportation Program (PNATE); National Program of Books and Teaching Materials (PNLD); National Program for the Restructuring and Acquisition of Equipment for the Public School Network for Child Education; and educational programs and actions of voluntary adherence that help maintain and develop education at specific levels or modalities[4].

FNDE has always been organized around the challenges that educational policies present at each moment, having as its highest instance a collegiate body, the FNDE's Deliberative Council. The institution has a well-defined and expressively small general structure concerning the scope of its activities.

The finalistic departments are responsible for managing the programs that, as a rule, are executed by the federative entities, with technical assistance from the Agency, which ensures supplementary resources, regulation, and monitoring. Finally, the Internal Audit and the Federal Prosecutor's Office at FNDE act to ensure the activities adherence to the principles of Public Administration and the institutional objectives.

This organization, whose structure acts dynamically, allows all spheres of the federation to rely on the results of the efforts made in it. FNDE's official structure is presented by the organization chart approved the Governmental Agency's new regimental structure[5].

---

[4] **Programas**, Fundo Nacional de Desenvolvimento da Educação, available at: <https://www.gov.br/fnde/pt-br/acesso-a-informacao/acoes-e-programas/programas>. Accessed in: 12 set. 2021.

[5] **Relatório de Gestão - Estrutura**, available at: <https://sites.google.com/view/relatriodegestao/apresenta%C3%A7%C3%A3o/estrutura>. Accessed in: 12 set. 2021.

Due to these dimensions and financial importance, as its special nature, the agency has a deliberative council. The Council is in charge of deliberating on the educational projects and programs financing, financial assistance to states, Federal District, municipalities, and private educational establishments; financing of scholarships, maintenance, and internship; the FNDE's budget; formulating policies for raising and channeling resources; judging resources in administrative proceedings of fiscal or extra-fiscal debts; approving FNDE's President accounts.

Therefore, the regulation, including the methodological authorization developed by this work, will depend on the deliberation and normative issued by this Council. Its composition includes:

- Minister of Education;
- FNDE's President and Chief Attorney;
- MEC's Finalistic Secretariats;
- President of the Anísio Teixeira National Institute for Educational Studies and Research (INEP).

Of the public policies described above, three will be highlighted in this project, for the reasons described in due course: PDDE, PNAE, and PNATE.

Created in 1995, the Direct Money in School Program (PDDE) aims to provide financial assistance to schools, on a supplementary basis, to contribute to physical and pedagogical infrastructure maintenance and improvement, with a consequent increase in school performance. It also aims to strengthen social participation and school self-management.

The program covers several actions that have specific purposes and target audiences, although the transfer and resources management follow the same operational patterns as PDDE.

The Aggregate Actions are grouped into different types of accounts under two fundamental axes: quality and structure. The Quality Axis aggregates the actions: Emergency, Connected Education, New High School. The Structure Axis includes: Water and Sanitary Sewage and Accessible School.

The PDDE is aimed at basic education public schools of the state, municipal, and Federal District networks, also private special education schools maintained by non-profit entities.

The National School Feeding Program (PNAE) provides school meals and food and nutrition education for students in all stages of public basic education. The federal government transfers to states, municipalities, and federal schools

supplementary financial amounts in 10 monthly payments (from February to November) to cover 200 school days, according to the number of enrolled students in each education network.

PNAE is monitored and inspected directly by society, through the School Meals Councils (CAE), and also by the FNDE, the Federal Court of Accounts (TCU), the Office of the Comptroller General (CGU), and the Public Ministry. The transfer is made directly to the states and municipalities, based on the yearly School Census prior to the year in question.

The program covers students of all basic education (pre-school, primary, secondary, and youth and adult education) enrolled in public, philanthropic, and community schools (under agreements with the government). It is worth pointing out that the PNAE budget benefits millions of Brazilian students.

And finally, the National Program of Support to School Transportation (PNATE) consists of the financial resources automatic transferred to cover expenses with maintenance, insurance, licensing, taxes and fees, tires, tubes, brake, suspension, gearbox, engine, electrical, and bodywork services, seat repairs, fuel and lubricants for the vehicle or, as the case may be, the boat used to transport public basic education students living in rural areas. It is also used to pay for services contracted with third parties for school transport.

The resources are destined for public basic education students living in rural areas who use school transportation. The amounts transferred directly to the states, the Federal District, and the municipalities are made in ten annual installments, from February to November.

This scope description reveals an annual budget exceeding many countries' Gross Domestic Product (GDP), having impressive logistics and expertise to carry out its responsibilities. Part of it is to be accountable.

As a public agent representing the state authority, the FNDE, when financing public policies by sending funds, constitutionally creates an accountability obligation to the recipient: a municipality or a state government. This means that each financial transfer, at a considered period closing, most often a fiscal year, must also render its opinion on these accounts, being directly responsible for the administrative collection in case of public resource misuse, including as the legal subject of eventual judicial proceedings for the recovery and sanctions applicable to the matter. According to the sole paragraph of Article 70, in our Federal Constitution:

> Art. 70 (…)
> Sole paragraph. Any individual or legal entity, public or private, who uses, collects, keeps, manages, or administrates public money, assets, and values or for which the Union is responsible, or who, on behalf of the Union, undertakes pecuniary obligations, will render accounts.

As part of the Brazilian public administration, only effective public servants, that are, those approved in a national civil service exam, would be able to judge the public policies beneficiaries' accounts. Here is then the challenge foundation: social policies growing demands *versus* the operational capacity to judge these programs' effectiveness.

Of course, if the beneficiary agent does not perceive the State as being vigilant as to the resources employed, misappropriating state-origin resources could become a viable risk due to the lack of possible sanctions perception. Risk management, if it exists in the FNDE, is inadequate when it accepts that most of the rendered accounts are not evaluated.

Still, with the prospects outlined above, it should be recalled that FNDE is a federal public agency, characterized as an Autarchical Fund. In this context, it is benefited by an annual budget approved by the Legislative power, executed year by year. As it is not convenient to discuss, in this case, the generation of profits or dividends distribution, given its legal nature, this work will demonstrate the value creation for the entity employing generated savings or quantifiable administrative rationality.

As will be detailed, the agency currently receives up to 35.000 accounts rendered per year, by states and municipalities benefited by its financial programs in support of basic education, of which about 10% are judged and concluded annually according to the productive capacity of the department assigned to this charge: General Coordination of Accountability Monitoring (CGAPC), of the FNDE's Financial Directorate (DIFIN)[6]. We do not understand whether the solution to the imbalance between operational capacity and judgment of accounts rendered is feasible by hiring more manpower. Not the least, it should be clarified that the agency has held only two civil service exams to hire new public servants

[6] ARAUJO, Vladimir B.; ANANIAS, Mateus C.; MACIEL, Andre L.A., Techinical Note 04.2018-DIFIN/FNDE, 2018.

in the last 20 years[7], besides the fact that in recent times we have experienced a moment of fiscal austerity, after almost two decades of social governments with growing demands and public spending. Here, then, we approach the economic concept of scarcity to determine the methodological value creation at the end of the project.

Scarcity[8] "is the condition in which our wants (for goods) are greater than the limited resources (land, labor, capital, and entrepreneurship) available to satisfy those wants". But what would be the main scarcity from a billionaire public entity, whose existence is primacy in the financial provision of Brazilian public education policy? The simplest answer: operationality. It means that, as much as it has an abundant public budget, as a federal public service provider, FNDE also suffers from a balance of its operational capacity: excess demands, qualified labor scarcity for constitutionally mandatory service. As we scale this scarcity to the segment that verifies the results achieved by public policy, accountability, operational capacity will show themselves to be collapsed. Hence this project proposal as a path to operational balance: accountability risk management.

It should be highlighted that the Strategic Planning at FNDE has evolved over the last few years. The 2018-2022 cycle focuses on knowledge transfer and innovation in public services[9]. It is now that innovative initiatives, such as risk management, become an active and structured process for this proposed cycle.

For any strategically planned *Mission* and *Vision* to be consolidated, internal processes are essential. In this regard, the public agency must consider operational excellence, especially the improvement of its accountability and integrated monitoring development. Therefore, it is essential to make visible the public value that FNDE has provided for Brazilian society, showing that the financial resources for the public programs were spent appropriately and generated the expected results.

---

[7] **Gestão de Pessoas - Portal do FNDE**, available at: <https://www.fnde.gov.br/index.php/gestao-de-pessoas>. Accessed in: 9 ago. 2021.

[8] ARNOLD, Roger A., **Economics**, 9e [ed.]. Australia ; Mason, OH: South-Western Cengage Learning, 2010.

[9] FNDE, Assessoria de Comunicação Social do, **Mapa Estratégico - Portal do FNDE**, available at: <http://www.fnde.gov.br/index.php/acoes/fnde-estrategico/plano-estrategico/mapa-estrategico>. Accessed in: 23 out. 2019.

Last but not least, for this paper it should be noticed that we structured our work development by following the understandings learned at the School of International Business and Entrepreneurship (SIBE) regarding management, in particular, establishing objectives, current situation, and project framework, all directly related to the strategy proposed for the problem-solution presented.

This systematic approach base itself on the strategic triangle of business development developed by SIBE[10]showed as follows.

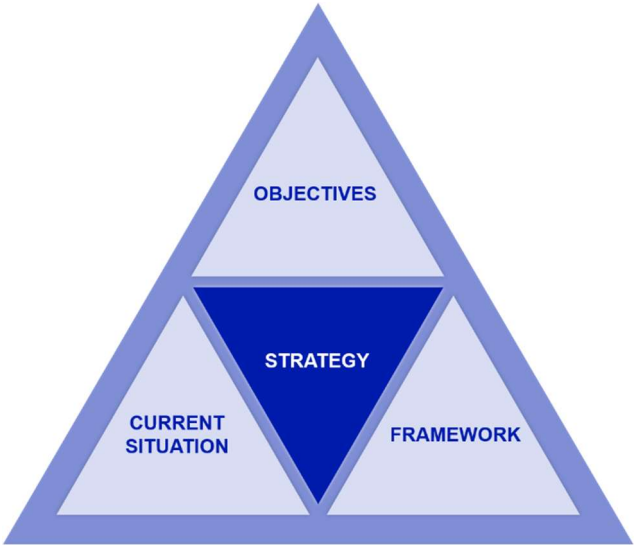


FIGURE 1. THE STRATEGIC TRIANGLE OF BUSINESS DEVELOPMENT. SOURCE: FAIX, W. G. ET AL. (1994). DER WEG ZUM SCHLANKEN UNTERNEHMEN. MÜNCHEN: MI-WIRTSCHAFTSBUCH, PP.178

As will be observed, General Management Tools[11] introduces to our core business the basis for project development and eventual report preparation, together with the additional seminars and studies developed over the last two years under the SIBE guidance.

In our work, we will attempt to demonstrate one of the main concepts learned in this book and seminars: "objectives play a fundamental role in the business development process". Applied to the public service reality, it will be perceived that, given the similarities, the theory can be fully applied. We will see that we need to revisit institutional objectives and pursue their processes formation to create value for the entity.

---

[10] NAGEL, Kurt *et al* (Orgs.), **General management tools**, 3. Aufl. Stuttgart: Steinbeis-Edition, 2013.

[11] *Ibid*.

We will see a true entrepreneurship force directly linked to institutional objectives, through an innovative project, but fully compatible with the strategic planning and structuring objectives.

If "entrepreneurial goals play a central role in the business development process"[12], then FNDE does not achieve business objectives as the essential force behind entrepreneurship, and thus, behind the Agency or Ministery of Education itself.

Note, if the entity has the fundamental duty to verify the public policies fulfillment that it finances and does not do so, for whatever reason, the institution does not fulfill its fundamental objectives. At least, it has no compliance evidence or effectiveness from these objectives covenanted with society. Our project, besides highlighting this failure, will propose a viable and scalable solution, placing the institution back on track with its strategic planning.

Following the guidelines in Figure 1, the first element is to understand the current situation in detail. Nagel[13] describes richly the need for clarity in business analysis and how meaningful it should be in representing the company. This is why we looked at various situational studies, government reports, and procedural cost studies to understand not only the gaps in FNDE but the possible solutions.

In this case, we chose to describe departmental costs, established value chains, possible success or failure factors, and the current strategic map. In this context, we seek to set up an environment focused on the institution. As we refer to a typically state-owned public service, it makes no sense to discuss competitors, so we see this business as viable to be modeled, institutional benchmarking.

Although our analysis process cannot be dissociated from economics or politics, especially given the institution nature, we do not limit ourselves to this exogenous framework. We consider laws and regulations with the potential to limit or prevent the project release, after all, as a state entity, we follow the principle of legality: only procedures foreseen by law, rule, formal regulation, etc.

In this context and continuing the guidelines solidified in the reference book, we relate the institution and its framework in a potential risks mix, where is presented the necessary remedies to each risk hypothesis confirmed.

---

[12] *Ibid*.

[13] *Ibid*.

In any case, well-defined, intelligible, and clear objectives for all the actors involved have proven to be more important than analyses, project developments, or even environmental concepts. It must be established that these elements are dynamic, can evolve, and can change throughout the project, but not the agreed-upon goals. Changing the objectives means a new project or objective inadequacy established for the institution. In any case, a mistaken reading provokes rework or even an entire project renunciation. Clear objectives mean a clear, well-defined strategy. The other elements and all their dynamism, are means to achieve these objectives.

Our strategy, to propose an adequate risk management methodology that balances the need for judgments on rendered accounts and operational capacity, must necessarily demonstrate, in this paper, the public agency's current situation, and how it has been led over the years to this state; the intended objectives; and the correct framework for the success of the solution we propose.

2. PROBLEM DESCRIPTION AND ANALYSIS

Over fifty years of existence and through the last decade, the organization struggles to balance its billionaire public policy funding capabilities with the capacity to evaluate its results through accountability judgments. As the operational human capacity is continuously decreasing, either by retirements or public servants migrations to more attractive state careers, with no human resources perspective to be invested in the next years as a result of the current economic austerity policy. In this period, the main strategy to try to make public resource managers accountable and the capacity to complete these processes feasibly consisted in developing the Accountability Management System – SiGPC[14].

The new electronic tool should receive, through data entry organized and declared, all accountability reports delivered to the FNDE. Resolution No. 2/2018 from FNDE's Deliberative Council[15] would establish, therefore, the mandatory

[14] ALVES, Tiago Venâncio, **SIGPC - Acesso Público - Portal do FNDE**, available at: <http://www.fnde.gov.br/fnde_sistemas/sigpc-acesso-publico>. Accessed in: 18 set. 2019.

[15] EDUCAÇÃO, Assessoria de Comunicação Social do FNDE com informações do Ministério da, **Resolução/CD/FNDE nº 2, de 18 de janeiro de 2012 - Portal do FNDE**, available at: <http://www.fnde.gov.br/acesso-a-informacao/institucional/legislacao/item/3512-

use, from 2012 onwards, of the Accountability Management System (SiGPC), developed by the FNDE for the accountability process management. This means that, as from that fiscal year, the data would be structured and available electronically, with the normative provision by the public agency to no longer receive such obligations on paper. Nonetheless, this system's development should turn, shortly, the entire accountability process automated, eliminating the new unanalyzed accounts inventories formation and mitigating the accumulation of new ones rendered without conclusive judgment.

Although it is not the purpose of this work, the challenge to be overcome should establish that the SIGPC implementation did not succeed, hence the need for an innovative alternative. It is a fact that the public agency does not have the professional vocation to develop such a complex system, since it is not its state career profile to be a public agent systems developer. It does not have the *expertise*.

Therefore, for the project's success, specialized labor should be essential, as well as to review complex processes before trying to modulate them to the system. At the moment of the tool development, internal workflows were not clear or formally charted, although the applicable legislation required the system to embrace dozens of different public policies, each one with its own normative dictating execution and, as consequence, inspection rules.

As a country of extensive social demands, especially in education, the new public policies creation or even reviews of existing ones brings to evidence a lack of normative stability, especially when considering that the process of creating and developing the SIGPC exceeded one budget year, or even consecutive government mandates. Therefore, the regulatory norms dynamics and even an agile software factory model add failure factors to the desired outcome.

The lack of compliance, or even the lack of provision for these essential requirements, has prolonged the accountability system development for more than a decade. Still, after at least ten years of planning and developing, the unfeasibility of hiring a company that specializes in systems, investments, and even multidisciplinary efforts, led the public agency to not complete the tool.

---

resolu%C3%A7%C3%A3o-cd-fnde-n%C2%BA-2-de-18-de-janeiro-de-2012>. Accessed in: 9 ago. 2021.

In any case, even though the project has not been able to automate the accountability process, as a premise, SIGPC has been successful in becoming an account rendering electronic repository. It means that there are large volumes of structured data available to be studied and diagnosed, therefore, to be analyzed even by machine learning algorithms. It is from this point that we establish the initial premise: structured data existence.

This data study will confirm the correct choice of our objective, presented in chapter 4. These objectives will be the driving force of our best efforts. Through this data, we were able to build useful information that could highlight the current situation, an essential part of any professional innovation management, as suggested by the strategic triangle of business development[16].

This first analysis should open by highlighting the FNDE's historical capacity to transfer credits to public programs beneficiaries under its responsibility. Reflecting the communication from the minister of the Federal Court of Accounts (TCU), Mr. Raimundo Carreiro, we have first public evidence of the public agency budgetary evolution[17].

On this occasion, the minister reports the chronicle accountability obligations build-up not analyzed by FNDE. See figure 2.

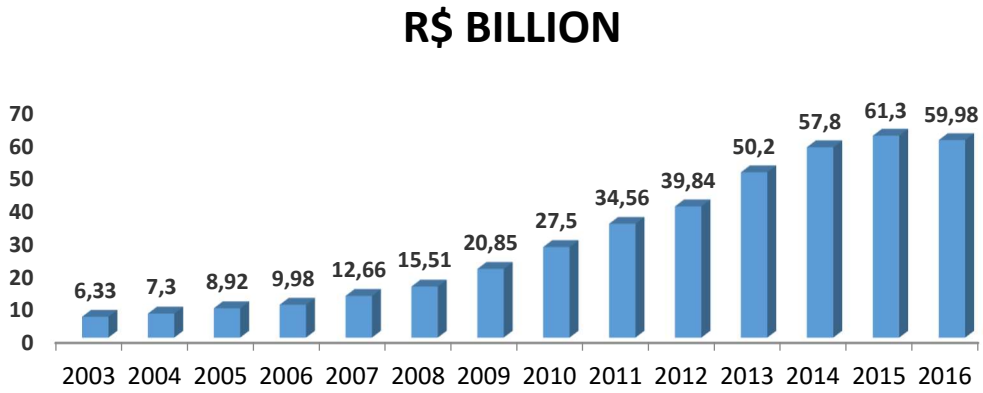


FIGURE 2. DEMONSTRATION OF FNDE BUDGET AVAILABILITY EVOLUTION, IN BILLIONS OF REAIS. SOURCE: TCU/2016.

---

[16] NAGEL *et al* (Orgs.), **General management tools**.

[17] **Pesquisa textual | Tribunal de Contas da União**, available at: <https://pesquisa.apps.tcu.gov.br/#/documento/ata-sessao/*/DATASESSAOORDENACAO%253A%255B20160921%2520to%252020160921%255D%2520COLEGIADO%253A%2528Plen%25C3%25A1rio%2529/DTRELEVANCIA%2520desc/1/%2520>. Accessed in: 9 ago. 2021.

As stated in the TCU Ordinary Session of September 21st, 2016, Minister Carreiro makes use of this graph to reveal that, in the last ten years, the volume of the resources decentralized via FNDE leaped from R$ 9.98 billion, in 2006, to R$ 59.98 billion, in 2016, which represents approximately 500% nominal growth. It also reports that such a significant increase in the volume of the decentralized resource would require a corresponding increase in the FNDE control structure, which has not occurred: 224 thousand accountability reports in stock at the Government Agency, pending analysis, 140 thousand of which on "paper" (non-electronic format); 35 thousand reports received annually on average; and 435 civil servants throughout FNDE, only 53 allocated to the accountability area, generating an equivalence of approximately 4.200 reports pending analysis per civil servant dedicated.

To properly illustrate what this situation might represent, let's assume that the operational capacity would be maintained over the years and that absolutely no new financial accounts would be received so that the current inventory would remain static and without the new increments possibility.

It means that the 224.000 accounts submitted, decreasing by human analysis by 3.500 per year, would take 64 years to complete. This data demonstrates the current model's non-sustainability, the need **for radical change, and professional planning** for it to be effective and definitive.

For the current situation, then, and for the first time using data governance[18], we used the main "accountability information concentrator", SIGPC, and other agency systems to obtain relevant information before planning the accountability solution.

Bergson L. Rego, says that Data Management and Data Governance are new and still little explored subjects[19]. Although organizations have gradually realized that the best practices adoption in data management and governance provides cost savings, increased information security, assertive information decision making, and even competitive potential, the FNDE is certainly not at the forefront. However, our efforts have the potential to change this perspective. But

[18] RÊGO, Bergson Lopes, **GESTAO E GOVERNANCA DE DADOS : promovendo dados como ativo de valor nas empresas.**, Brazil: BRASPORT; Edition: 1, 2013.

[19] *Ibid*.

before we introduce the first segments from the project data management, it is worthwhile to explore the topic a little further.

According to Rêgo, "data is essential to organizations". The author describes data should be treated as key assets for operational and strategic organizations maintenance. Something was never done in our public agency, but with a relevant precedent if we consider the successful history of Brazil Federal Revenue Service, from which we seek a baseline model for choosing according to the risk of "error".

In this sense, every organization seeking professional management compatible with the 21st century, in addition to understanding its data, must proactively seek to adopt a set of practical improvements that enable stakeholders to have the necessary security in managing their data and transforming it into useful information.

For Data Management, we need to turn data into assertive information. The first process after collection is the association and information assembly that will provide guidance. There is no wrong information if it is constructed correctly. But there is information that is useful to the process and disposable (to this process) but that can be useful to others inside the government agency, even for its financial management characteristics of public policies.

Therefore, we don't dismiss any of our acquired information, but we start directing it to the ones that leverage our process, as a strategy for retaining "advocacy".

As already state, according to the current Brazilian constitutional model, when a credit authorization (budget) becomes a financial transfer (monetary), an accountability duty arises. Consequently, the financial volume yielded to the beneficiaries is expected to be clearly compatible with the volume of the accounts for conclusive analysis, as demonstrated in figure 3, especially if compared with figure 2. See as follow:

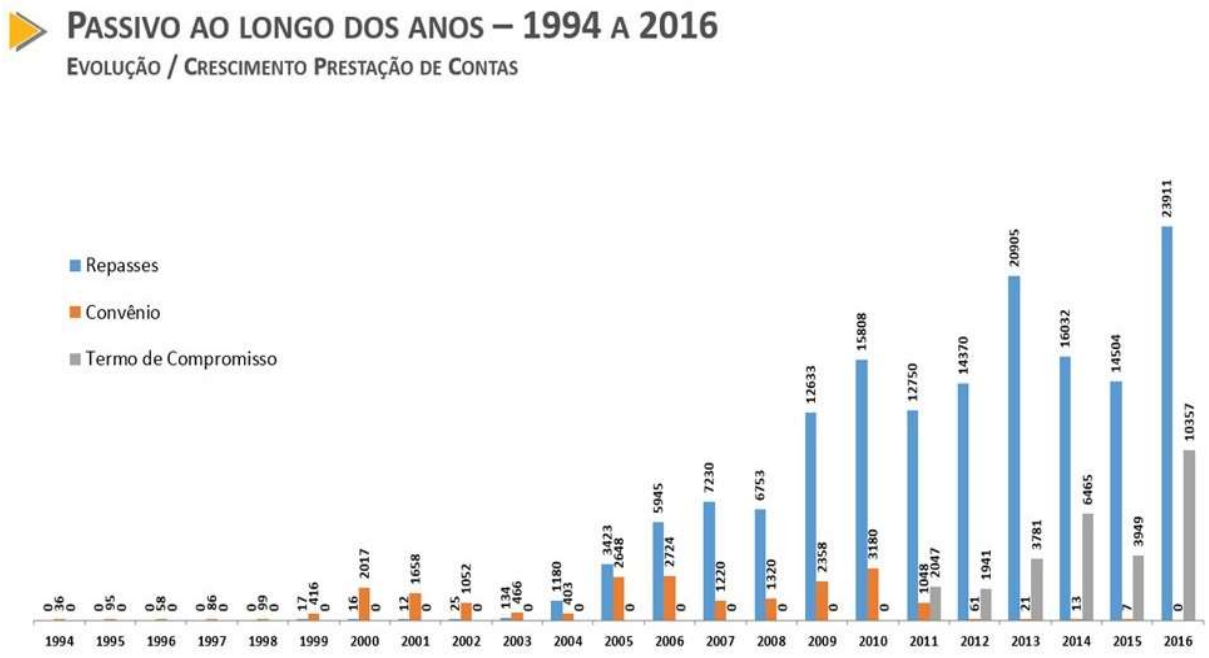


FIGURE 3. CHRONOLOGY OF NEW ACCOUNTABILITY INPUTS MATCHING BUDGET INCREASE, AS RELATED BY TCU MINISTER, RAIMUNDO CARREIRO, TO THE PLENARY COURT AT THE END OF 2016. SOURCE: DIFIN/2016.

Figure 3 shows three typical financial models, with chronology from 1994 to 2016, the communication is given year in plenary at the Court of Accounts. For the original figure, in Portuguese: “repasses”, means automatic transfers to public policy programs; and “convênios” & “termos de copromisso”, represent each transfer through public agreements.

Our paper considers an intelligence preliminary work gathering in 2018, so it will select accounts for risk management by 2016, since there wouldn't be enough input data from 2017 onwards, considering the analysis capability of traditional methods. In any case, the available period studied contains enough data, including in research attributes, so that a predictive algorithm elected could form a decision on the following rendering of accounts. This description will be further discussed in chapter 5, Approach and Method.

Finally, we were able to identify descriptively the most relevant problem, considering all resource transfer models adopted by the public entity: it lies in what we shall refer to as the "Programs" modality. These have, as main characteristics, the agreement (adjustment or similar) celebration exemption, being enough the adhesion for the automatic resources transfer, and to be accountable.

In addition, this modality is fully available in input data organized in SIGPC. It should be noted that for the order of 35.000 new *inputs* per year, close to 24.000 are only for the “Programs” modality for conclusive processes to be analyzed, our *output*. Certainly, it will not be supplied in operational equilibrium through more staff addition. Hence, this first scenario demands a viable alternative that reverses the situation presented efficiently. The need to innovate, ergo, is imperial: **add value to the institution through a new methodology**.

The major issue is that for the current situation, the FNDE, as well as many government agencies, adopt what we will call a "**deterministic model**" for judging accountability.

We have already seen that the duty to render funds accounts from the National Treasury is constitutionally protected (article 70, sole paragraph). Even so, regardless of the constitutional protection, the duty to render accounts was not only already foreseen in Decree-Law No. 200, on February 25th, 1967[20], but its article 93 already established the good and regular use proof onus relying upon who effectively used them, with its evidence following the Administration's rules:

> Art. 93 Whoever uses public money must justify its good and regular use in conformity with the laws, regulations, and norms emanating from the competent administrative authorities.

It is relevant to underline that our Constitutional guardian, the Federal Supreme Court (STF), has a clear understanding as to the obligations and responsibilities, stating that the duty to render accounts belongs to the individual responsible for public assets and values, and not to the entity, regardless of whether the individual is a public agent or not[21].

> 3. The duty to render accounts, in this case, is not of the entity, but of the individual responsible for public assets and values, whether a public agent or not. 4. although the entity is a private law entity, it is subject to the supervision of the State, since it receives resources of state origin,

---

[20] **DEL 200**, available at: <http://www.planalto.gov.br/ccivil_03/decreto-lei/del0200.htm>. Accessed in: 10 ago. 2021.

[21] **- Mandado de segurança. Tribunal de Contas da União. Supremo Tribunal Federal STF - MANDADO DE SEGURANÇA : MS 21644 DF**, Jusbrasil, available at: <https://stf.jusbrasil.com.br/jurisprudencia/14706970/mandado-de-seguranca-ms-21644-df>. Accessed in: 10 ago. 2021.

> and its leaders must render accounts for the amounts received; whoever manages public money or administers goods or interests of the community must render accounts to the competent organ for supervision. (MS 21.644/DF, Pleno, Min. Néri da Silveira, DJ de 08.11.1996).

Therefore, each federal financial transfer destined to meet society's demands will necessarily be subject to the good and regular use demonstration of the resources received, minimally, as to the aspects contained in article 70 from the Federal Constitution, in its head:

> Art.70 The accounting, financial, budgetary, operational, and patrimonial inspection of the Union and the direct and indirect administration entities, regarding legality, legitimacy, economy, application of subventions and revenue waiver, will be exercised by the National Congress, through external control, and by the internal control system of each Power.

In summary, the current accountability analysis model, in which each financial transfer destined to the execution of a public policy generates a result consideration, assumes a deterministic model of conclusive analysis, where the financial transfer to the FNDE program execution is deterministic to an accountability analysis (see Annex I, Accountability Flow in SIGPC).

Despite the clear obligation and a deterministic flow, there are no obstacles to other evaluations adoption within constitutional methods to this duty besides the deterministic and parity ones. It is worth pointing out that it is a constitutional precept, among others, the **efficiency principle**, present in article 37 of the 1988 Federal Constitution.

It is on this constitutional principle that our proposal is based, where we respect the parity between the granting of resources and the respective obligation to render accounts, in which traditional conclusive analyses are **prioritized** according to the risk and those considered low risk are conclusively filed, relying on that in a predictive model, the result for these accounts would not have a judgment different from a human analysis. In other words: all the accounts continue with the guaranteed investigation, but only those considered as high malpractice risk will be traditionally analyzed.

Last but not least, we must state clearly that the FNDE, as a government agency representing the federal State, cannot evade analyzing the rendering of accounts and, eventually, taking legal action to recover the amounts possibly misappropriated or unproven. The fact is that, according to the Brazilian legal system, although criminal and administrative punishments are prescriptible, public debt is not. This is what the Brazilian constitution dictates:

> Art. 37 (...)
>
> § 5 - The law will establish the prescription terms for illicit acts committed by any agent, public servant or not, that cause losses to the public treasury, except for the respective actions for restitution.

It is important to highlight that this process is not dissociated from the understanding of the main parties responsible for public management control in Brazil, being followed and, sometimes, discussed in technical meetings with the central office for internal control of the Federal Executive Branch, CGU, and the Federal Court of Accounts - TCU.

In the case of the Federal Court of Accounts, which has already determined that FNDE should adopt solutions to its high stock of accounts, it recommended, through judgment AC-7790-42/15-1, that the agency and the Ministry of Planning "adopt an action plan with measures to reduce the high stock of accounts that have not been conclusively analyzed by the FNDE" [22].

Also, as a demanding matter regarding the " Accountability Management of funds decentralized by the FNDE", the Secretary of External Control of Education, Culture, and Sports - SecexEDU/TCU submitted the process TC 027.076/2016-4 to the Court, holding clear adequacy of its proposed content to the FNDE project [23].

---

[22] **AC 7790/2015 - PRIMEIRA CAMERA| Tribunal de Contas da União**, available at: <https://pesquisa.apps.tcu.gov.br/#/documento/acordao-completo/7790/ANOACORDAO%253A%25222015%2522/DTRELEVANCIA%2520desc%252C%2520NUMACORDAOINT%2520desc/0/%2520>. Accessed in: 30 ago. 2021.

[23] **TC 027.076/2016-4 | Tribunal de Contas da União**, available at: <https://pesquisa.apps.tcu.gov.br/#/documento/processo/027.076%252F2016-4%2520/COPIAANO%253A%25222016%2522/DTAUTUACAOORDENACAO%2520desc%252C%2520NUMEROCOMZEROS%2520desc/1/%2520>. Accessed in: 30 ago. 2021.

The process, among other actions, proposes that "prioritizing a set of accounts for more detailed instruction is not to fail to analyze the universe of accounts, as long as this universe, in aggregate form, is subjected to filters and tests". It means a "fine mesh" criteria that allow FNDE to classify the rendering of accounts as to the level of criticality, based on indications of non-compliance determined from the typologies defined in the risk matrix.

TCU continues in its judgment by stating that "the better this classification will be, the better the risk criteria are defined and, above all, the more qualified are the data and information that enter the system". The text further states that for better model calibration, the selected risk typologies should be regularly checked for their ability to classify accounts as correctly as possible regarding their criticality level.

It highlights that the proposed model is not an adjustment to FNDE analysis capacity, but an innovation that classifies the accounts, even making it possible to mass file those accounts classified as low misappropriation risk.

One of the premises established is "the possibility of mass filing the accounts that passed through the fine mesh and were classified as non-compliance low risk".

This premise resumes the discussion dealt with in the Courts recommendation, about the possibility of the FNDE Deliberative Council to dispose of, based on the general rules established by the creation laws, on the principles of public administration and concerning the accountability obligation, about the selection methodology and accountability analysis, by **sampling**, based on risk and data crossing. In short, according to the TCU: "as well as, consequently, on the accounts archiving with low-risk irregularities".

This last premise, crystallizes the understanding already elucidated by the legal position nº 0001/2016/DEPCONSU/PGF/AGU[24], as to the accountability archiving processes within the scope of the PNAE, PNATE, and PDDE. According to the Federal Attorney General's Office advisory body, these three programs, unlike covenants and similar instruments, are independent of the execution of any instrument, constituting automatic transfers regulated through a

[24]**SEI / FNDE**, available at: <https://www.fnde.gov.br/sip/login.php?sigla_orgao_sistema=FNDE&sigla_sistema=SEI>. Accessed in: 30 ago. 2021.

Deliberative Council resolution from the Government Agency and, for the archiving purposes, it concludes that, view of the above, "it is possible for the FNDE Deliberative Council to establish rules for the accountability archiving processes related to automatic transfers under the PNAE, PDDE, and PNATE".

Therefore, it is up to the Federal Attorney's Office in FNDE to assess any proposal for a normative act that may be presented, with the suggestions concerning this subject.

Finally, the CGU has formally recommended on two occasions to this Agency to adopt risk management based on data mining directly applied to our program's accountability.

The first, through Technical Note no. 731/2017/DIV1/CGEDUB/DS/SFC[25], states that, as for the FNDE processes analysis, for this to happen in a **timely** manner, it is suggested "the adoption of a methodology based on risk management for the processes selection that will be analyzed, appropriate to the operating capacity of the Government Agency".

CGU continues by indicating the need to elaborate indicators, data that could be used such as the involvement of large transfers, of executing entities that have been subject to inspection by the control agencies, of entities that have failed or not submitted previous accounts, or that have not been followed up for a long period, besides the use of the data mining technique to identify patterns and trends, among other factors.

To them, "the processes selected by the established criteria will be analyzed and, the others, finalized, without this preventing, however, that processes already concluded without merit analysis by the FNDE be reopened, if facts arise that point to possible irregularities".

And then, through Technical Note No. 1357/2018/CGEDUB/DS/SFC, ratifying the proposal as a **viable and appropriate solution**[26], CGU describes that "the current model adopted by FNDE is expensive and not very successful, and there is a direct dependence on the operational capacity of the unit to achieve good results". Furthermore, they continue, "despite the manager's efforts to solve the issue, the advances have been timid".

---

[25] NEVES, Clara V.B.; DE JESUS, Victor L.F.; CONCEIÇÃO, Shirley G., Technical Note Nº. 731/2017 / DIV1 / CGEDUB / DS / SFC, 2017.

[26] PEREIRA, Fernando S. *et al*, Technical Note n° 1357/2018/CGEDUB/DS/SFC, 2018.

On the other hand, CGU believes that it is possible to use data extraction, crossing, statistics, and mining techniques to help the process of accountability treatment and selection. In this sense, that is "possible to elaborate a classification index of the accounts rendered, as well as to identify a pattern of prices practiced for several items". Furthermore, they found patterns that can help the manager's decision-making.

Following the Technical Note idea, it is clarified that "the use of this methodology promotes the accountability optimization of the analysis process, making it more timely". Following TCU judgment of FNDE accounts, it happens because there is a better accounts treatment, propitiating a prioritized selection. In addition, "a risk-based methodology increases the chances of having a more assertive analysis, resulting in a better use of the limited manpower".

As will be proved, the method proposed in this paper is in line with what is proposed by the highest federal public management control instances, including the expectations from those Offices and agency senior management, especially providing an answer to society regarding the good, regular, and efficient management of public resources for public policies in Education.

Finally, as much as this process is innovative, the federal government, through Interministerial Normative Instruction nº. 5, November 6, 2018, shows itself to be favorable to efficient initiatives concerning risk evaluation methodology:

> Art. 1 This Normative Instruction establishes rules, guidelines, and parameters, based on the risk evaluation methodology, for adherence by concessor agencies and entities, following § 7º of art. 62 of the Interministerial Administrative Rule n. 424, of December 30, 2016.

It is FNDE's responsibility, in a similar way to this interministerial normative instruction, to establish its guidelines and parameters, keeping the due analogy, employing a resolution by its Deliberative Council.

Clearly, it is perceived that some mapped challenges, given the circumstance, are in **perfect timing** to be transmuted as an **innovation opportunity**. Even in the perception of self-regulatory viability, as confirmed, for the prioritization process as we propose. These facts can be better grouped and understood by tracing the risk structure and respective responses, where we will consider, in due course, the influences of these internal and external factors in our process.

Next, we will structure the problem encountered to define the design parameters.

3.STRUCTURING THE PROBLEM

As described, the SIGPC creation as a proposed automation solution proved unsuccessful. In any case, even though the project has not been able to automate the accountability process, SiGPC has been successful in becoming an electronic repository of account rendering. A stated, there are large structured data volumes available to be studied and diagnosed, therefore, to be analyzed, also by machine learning algorithms.

According to Kurt Nagel[27], "adequate objectives and a suitable strategy for achieving these can only be defined after the company's current situation and the framework are known". The project planning for this project, to avoid the same shortcoming that occurred with SIGPC, that frustrated its main goal, must be in-depth. To this end, we use unconventional approaches: in order not to fail similarly, the project propositions must be different from the conventionally tried ones.

In this principle, we continue to structure our problem with an initial basic proposal but never attempted: to know objectively all the possible accountability circumstances, now, with the SIGPC use as an electronic repository. Figure 3 has already shown us that our challenge focuses on the "Programs" modality. A deep data analysis, then, reveals the distribution proposed in figure 4.

[27] NAGEL *et al* (Orgs.), **General management tools**.

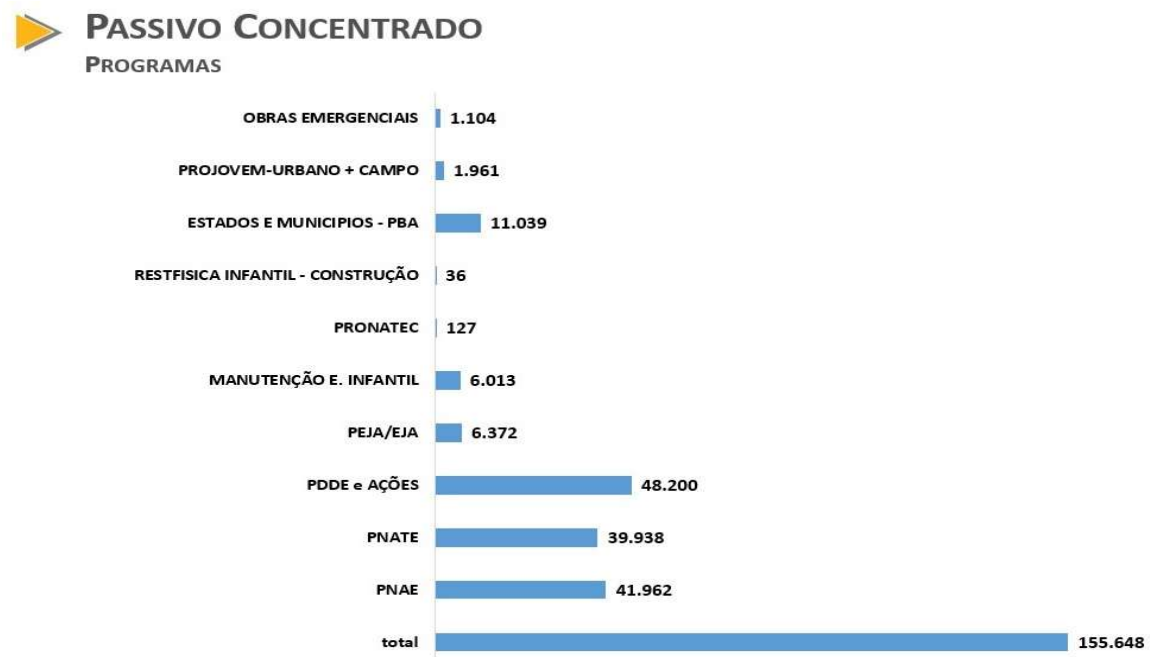


FIGURE 4. DISTRIBUTION OF THE ACCOUNTABILITY OBLIGATIONS BY "PROGRAM". SOURCE DIFIN2016.

As seen, the most relevant problem, considering all models of resource transfers adopted by the entity, lies in automatic transfers, so-called by FNDE as "Institutional Programs". From the distribution achieved, it can be observed that three program categories represent, at the time the data were obtained, as much as 83.59% of the accounts without conclusive analysis and already declared to SIGPC as input data.

Certainly, then, this should be our focus for the eventual risk management model. The modalities correspond to the Direct Money in School Program (PDDE), 48.200 records; National School Feeding Program (PNAE), 41.962 records and; National Program of Support to School Transportation (PNATE), 39.932 records.

With this database work and the **problem fully structured**, the project has the FNDE management historical survey, from the accountability point of view of its transferred resources to finance federal educational programs. Following these establishments, a statistical survey and studies using Big Data tools, and data-crossing in internal databases, shall build a risk management methodology using adequate data mining. Then classifying the accountability presented to the FNDE to indicate, with appropriate accuracy, which ones should be traditionally analyzed, due to the public fund's high-risk diversion, and which ones should be

filed as approved due to the low-risk rate. As we know that this task does not fit FNDE natural professional profile, we opted for hiring specialized consultants, a cost-effective and objective solution **to avoid past mistakes** in highly complex project conduction.

As can be seen, the problem structuring indicates that, necessarily, a solution must give up the deterministic model, at least as a priority, and adopt a risk management model. In this project, we propose a "**predictive model**" in substitution, but it is emphasized, constitutionally adequate: prediction is the main asset, but the parity analysis of each accountability under FNDE responsibility persists. All accounts are scrutinized, but only those in the selected classification range are to be submitted to human evaluation.

The project, therefore, brings an administrative rationalization solution, avoiding an increase of non-analyzed processes to become new in stock, at the same time that it provides a solution for an already existing inventory of over 200.000 accounts already pending conclusive analysis, and growing at around 30.000 new each year. More than that, it creates an abstraction to the process of checking for inconsistencies in public accounts, a so-called " **to be singled out for auditing**" or a “fine mesh” by risk management, generating a control expectation in the federal resources education manager (governors and mayors), able to make them more vigilant to its federal financed resources and creating a new accounts model to the government agency, to be called "**FNDE Fine Mesh**".

In a professionally conducted innovation project, structuring the problem also means foreseeing risks and their respective remediation actions. Kurt Nagel[28] describes that the business development process must define and assess potential risks. For this project, we can clearly list them as follows:

[28] *Ibid*.

Risk	Cause	Consequence	Response Action
Modeling Difficulty from existing data	Data entry freely. Few manual reviews compared to the rendering of accounts in stock.	Inconsistent data, inadequate data model.	Data cleansing to avoid inaccurate records from a record set, table, or database. Oversampling Techniques, increasing the representativeness of manual analyzes.
Budgetary unavailability for consultants payment	No financial allocation / no budget reserve	work Interruption due to non-payment	Block/reserve pre-budget allocation
Do not access external databases: GovData	Unavailability of data in GovData tool	Form statistical pattern considering only internal data	Search databases for the project in a dedicated way, creating web services
Refusal of finalist areas responsible for FNDE programs in participating in the project	Project unfamiliarity; scheduling incompatibility; lack of interest in managerial information; resistance to change	Temporal lapse for the understanding of technical analysis rules; delay in first deliveries; lack of sponsorship from the finalist areas	Project consequences evidence and provision of managerial information to Stakeholders with potential for political sponsorship
Inadequate or undersized IT architecture	Tools used for data mining with processing and storage architecture not available in the FNDE	Slow processing, inability to storage or improper storage	Installing R software on a dedicated server or use it on the GovData itself. Use other software; local architectures upgrade; DIRTI support

TABLE 1. PROJECT RISKS.

At this moment it must be created an understanding for drivers and inhibitors, as an enabler for scenarios using story dimensions and paths divided into driven axes that could be lighted by a SWOT analysis[29].

[29] HUBER-, Angelika *et al*, Rethinking the value chain, **KPMG | Rethinking the value chain**, p. 82, 2018.

In these terms, we have grouped factors considering internal and external aspects not yet brought up, so that we can define two variables on our axis that adequately reflect a trend portfolio.

Before going into the SWOT analysis, it is important to understand the so-called "strategic evaluation". This term refers to a process carried out to investigate both the sector in which the organization is inserted (Education), and to study the institution itself (FNDE). In other words, this analysis aims to assist in the creation of a necessarily institutional action plan.

Through this process, we were able to make the right decision, to guarantee the path to success for this endeavor, integrated with the institutional strategic plan.

SWOT aims to help evaluate and understand the current situation in which we find ourselves. This study facilitated the preparation of more adequate planning to the organizational objectives, besides contributing to the decision-making process.

Thus, the analysis consisted in filling out the following table, contemplating the internal and external scenario of the enterprise, framing "Strengths", "Weaknesses" as internal factors, and "Opportunities" and "Threats" as external factors.

As we go through the variables, the "strength" portrays the differentiations achieved or inserted in the project and which generate advantages. The "weaknesses", on the other hand, are the weak points identified in the internal context of the enterprise, providing it with disadvantages within its operations.

Externally concerning its industry, the "opportunities" contain the possibilities for the business to prosper by exploring new products or market segments. Even extending this concept to the public sector. Similarly, the "threats" are the obstacles to be faced.

In other words, through the SWOT analysis, it was possible to understand the positive points, the negative aspects, what can be improved, and what can hinder the project's success.

The result can be seen in figure 5 as follows, with each aspect registered in its respective quadrant.

FIGURE 5. SWOT MATRIX, CONSIDERING INFLUENCES FROM INTERNAL AND EXTERNAL ENVIRONMENTS.

In synthesis, due to what we have summarized in figure 5, it is possible to develop project strategies that are fully aligned with the organizational structure and FNDE culture, as will be seen further on. Furthermore, it becomes evident how our project can behave in the market in which it is inserted. Something fundamental to define a project's scalability.

Once FNDE's current situations related to the project are defined, as an endogenous framework, as well as basic constraints (resources, opportunities, norms, laws, technology, etc.) as an external framework, we can finally define the project's objectives in the next item.

4. AIMS AND TARGET SETTING

Our objective is to propose an appropriate risk management methodology that enables conclusive accountability analysis of rendered accounts, considering programs listed for the pilot project, balanced with public agency operational resources, to avoid generating new accounts pending conclusive analysis.

The appropriate indicator should measure project success: growth rate for conclusive analysis inventory. The target rate goal for generating new inventories for PNAE, PNATE, and PDDE, should be ideally 0%. It means that the algorithm predictably concludes by approving some accountability obligations and rates others to traditional analysis, respecting the operational capacity established for the process.

4.1 SETTING S.M.A.R.T. OBJECTIVES

Although we have clearly defined our objectives and targeting, it is convenient and didactic to set S.M.A.R.T. objectives and confront them in this paper results regarding the achievements.

"S.M.A.R.T. Goals" is a methodology name that establishes criteria for defining goals, whether they are personal or professional, individual or collective. Its proposition is attributed to George Doran, then an executive at Washington Water Power, who in 1991 presented it in the article "There's a S.M.A.R.T. Way to Write Management's Goals and Objectives"[30].

According to the article, Doran believed that managers get confused when describing the results they want to achieve and that this happens because of too much information they seek for it, impairing their focus on the task.

According to Charted Management Institute (CMI), objectives set out what a business is trying to achieve. They should be based on organizational strategy and be aligned with corporate vision, mission, and values[31]. For this reason, our paper brings a review of the FNDE strategic map, constantly confronting it with the very reason for the project's existence.

---

[30] S.M.A.R.T-Way-Management-Review.pdf.

[31] Setting SMART Objectives, p. 5, .

Following CMI directives, there are several different versions of the acronym with different terms associated with some of the letters. Objectives are 'S.M.A.R.T.' if they are specific, measurable, achievable (sometimes agreed), realistic (or relevant), and time-bound (or timely)[32]:

- **Specific** – to provide adequate risk management methodology for FNDE accountability
- **Measurable** – the increase of new accounts rendered without conclusive analysis of FNDE's current inventory should be *null* when fully implemented.
- **Achievable** (or agreed) - the methodology is achievable and will bring balance to the *INPUT* and *OUTPUT* of accounts rendered by using an adequate accuracy predictable method, with machine learning that runs the process increasingly reliable and accurate. All FNDE board of Directors must sign an agreement with the methodology, because it involves their departments, as public policy managers.
- **Realistic** (or relevant) – this work focuses on the outcome of providing a realistic solution to the FNDE major public management challenge. Something that lasts for 50 years with no apparent solution. Only risk management can change the current paradigm.
- **Time-bound** (or timely) – Not only is the result attainable, but the timing is perfect for the moment of austerity in public accounts. More than this, it is possible to draw a balanced plan at the end of this work, so that in five years FNDE will extinguish its stock of non-analyzed accounts and cease to generate new ones.

It will be seen that this work has developed, in a fundamental way, to meet each of these requirements successfully.

[32] *Ibid*.

5. APPROACH AND METHOD

5.1 VALUES STATEMENT

The primary approach to the project should demonstrate, as a business rule, the compatibility between project and institutional strategic planning. To this end, an introductory analysis of the FNDE strategic map is required [33].

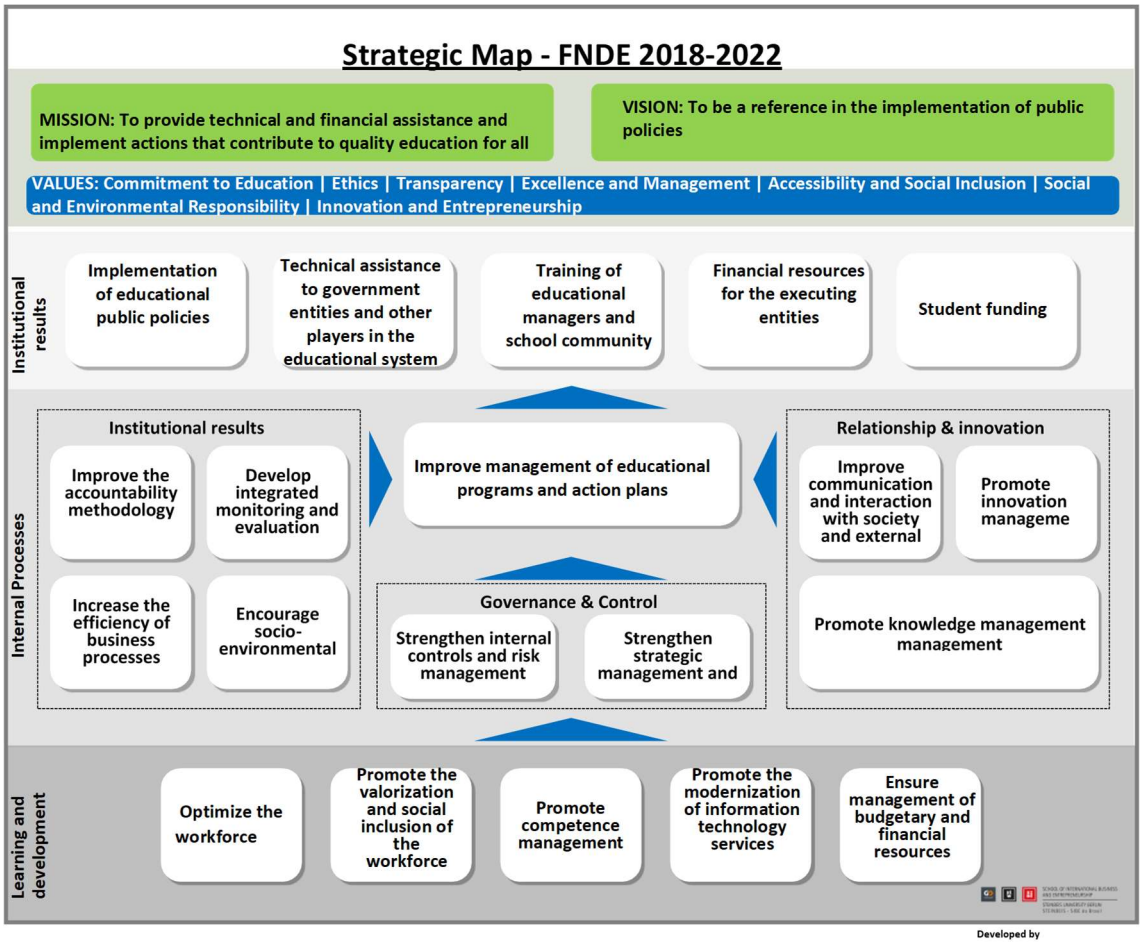


FIGURE 6. FNDE STRATEGIC MAP

It is highlighted that in "internal processes", a shared path to "Improve management of educational programs and actions" resides in achieving "Operational excellence". In this aspect, our project seems to be strengthened as a strategic core for the proposed cycle.

---

[33] EDUCAÇÃO, Assessoria de Comunicação Social do FNDE com informações do Ministério da, **Mapa Estratégico - Portal do FNDE**, available at: <http://www.fnde.gov.br/index.php/acoes/fnde-estrategico/plano-estrategico/mapa-estrategico>. Accessed in: 18 ago. 2021.

This project fits perfectly in the accountability methodology improvement, including integrated monitoring development for public policies evaluation, with the consequent increase in the business processes efficiency, as will be proven in this paper. It should be clarified that one of the data mining collateral processes, with extensive study and information development, is not only about choice efficiency and procedural economy. Data collection and organization make it feasible to build Business Intelligence (BI), with diverse managerial panels that can enable newly acquired knowledge management and even public policy monitoring, audit planning, and much more.

A mission, a vision, and a strategy are basic parameters describing the enterprise's conduct. As for the mission statement, it is a task with a high degree of importance, including the responsibility for its implementation. When formulating a mission, a statement is expressed in which the essence of the organization's operation is stated[34].

Behind this concept, FNDE organizes its actions guided by its institutional Mission: "to provide technical and financial assistance and to execute actions that contribute to quality education for all". As a Vision of the future, it focuses on "being a reference in the implementation of public policies". These elements form the strategic benchmarks that are unfolded into objectives, goals, and projects. In addition to these, a set of values was defined that translate and guide how the initiatives will be implemented in the public agency: Commitment to Education; Ethics; Transparency; Excellence and Management; Accessibility and Social Inclusion; Social and Environmental Responsibility; Innovation and Entrepreneurship[35].

We understand that the entity's strategic management has chosen to be precise in its statement. However, we believe, by technique, that missions should embrace short and objective statements: while the vision must interpret the “big picture” of what you want to achieve, the mission must bring a general statement

[34] AGNIESZKA WIŚNIEWSKA, The Mission & Vision Statements in the Consumers’ Strategies, **Acta Universitatis Nicolai Copernici. Zarządzanie**, v. 11, n. 3, p. 111–118, 2018.

[35] EDUCAÇÃO, **Mapa Estratégico - Portal do FNDE**.

of how you will achieve the vision[36]. At this stage, this paper outlines the need to propose its Mission-Vision-Values statements **directed to the project**.

In economics, the “right amount” of anything is the “optimal” or “efficient” amount, and the efficient amount is the amount for which the marginal benefits equal the marginal costs[37]. As stated in this project, we have achieved efficiency when the marginal benefits of using data mining risk management equal the marginal costs of traditional analysis of accounts rendered to the FNDE. As our proposal aims that the economic benefits of predicting the account rendering result exceed the analyzing cost indistinctly to all inventory not worked yet, then a better *Mission* statement will be “Provide accountability efficiency”.

Likewise, as we propose a paradigm-shifting in the account management of Brazilian public agencies, the Project *Vision* will be " being a reference in the accountability risk management”. For the *Values*, we summarize “Commitment to compliance; Ethics; Transparency; Excellence and Management; Innovation and Entrepreneurship”.

Considering the above description compatible with FNDE statements already validated, we would have the following proposed adaptation to the Project Strategic Map:

FIGURE 7. PROJECT MISSION, VISION, VALUE.

[36] **Vision and Mission Statements -- a Roadmap of Where You Want to Go and How to Get There | Ag Decision Maker**, available at: <https://www.extension.iastate.edu/agdm/wholefarm/html/c5-09.html>. Accessed in: 6 abr. 2020.

[37] Arnold - 2010 - Economics.pdf.

With the FNDE board of directors and the presidency commitment, also considering the environmental analysis and the project compatibility with the strategic objectives, our proposal is promising to be evidenced in its implementation within this strategic cycle.

Kaplan and Norton[38] rightly describe the multi-business corporation challenge as being "how to operate multiple units within the structure to create value beyond what individual units could achieve on their own." In that sense, we must recall that a strategic map is the instrument of the Balanced Scorecard which aims to translate the path (strategy) to achieve the vision of the future and accomplish the mission[39].

According to Robert S. Kaplan and David P. Norton, a strategy-oriented organization follows five management principles:

1. Mobilize change through executive leadership;
2. Translate the strategy into operational terms;
3. Align the organization with the strategy;
4. Motivate to turn strategy into everyone's task; and
5. Manage to turn strategy into an ongoing process.

At FNDE, these were done through the articulation of strategic objectives arranged in perspectives (Institutional Achievements; Internal Processes; Learning and growth) and connection for cause and effect relationships that could make it possible to actions identification necessary to achieve each goal.

The strategic map was organized, then, in three perspectives that distribute the objectives strategies designed for this management cycle – 2018 to 2022. The "Learning" perspective and "growth" organize the related challenges to technology, people, and institutional budget. These challenges drive the proposed objectives for the "Internal Processes", which is then when we align what must be done for the organization to have excellence in its middle and final

---

[38] MCLEAN, R., Robert S. Kaplan and David P. Norton-Alignment: Using the Balanced Scorecard to Create Corporate Synergies, **AUSTRALIAN JOURNAL OF MANAGEMENT**, n. 2, p. 367, 2006.

[39] KAPLAN, ROBERT S; NORTON, DAVID P, **Organização orientada para a estratégia : como as empresas que adotam o balanced scorecard prosperam no novo ambiente de negócios / Robert S. Kaplan , David P. Norton ; [tradução Afonso Celso da Cunha Serra].**, 17. ed. Rio de Janeiro : Elsevier : Campus, 2001.: [s.n., s.d.].

activities. Finally, the perspective of "Institutional Achievements" points to the great deliveries from FNDE to society.

The FNDE Strategic Map distribute the strategic objectives in perspectives, to be organized as follow:

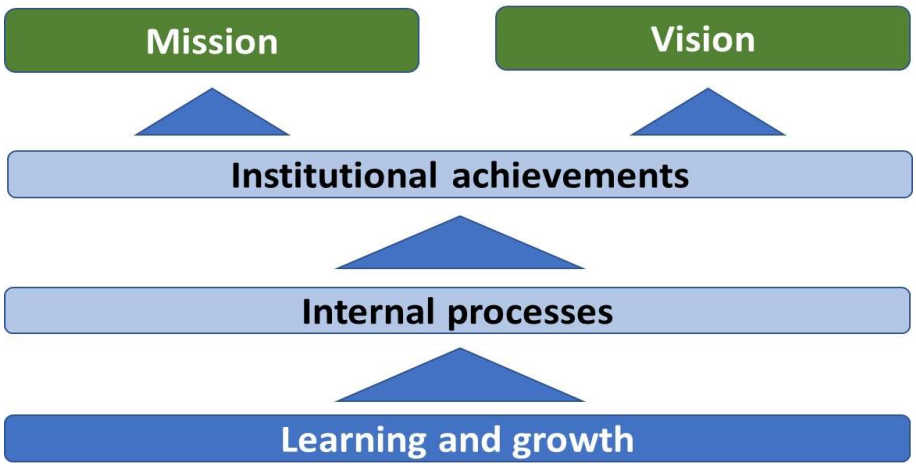


FIGURE 8. PROSPECTS FOR THE STRATEGY.

5.2 STRATEGIES FORMULATION FOR GOAL ACHIEVEMENT

The strategic objectives determine what must be done so that the organization accomplish its mission and achieve its Vision of the future: translate the challenges to be faced by FNDE in fulfilling its institutional role. To the 2018-2022 cycle, with the mission statement and strategic thrust validated, twenty strategic objectives were defined[40].

Among these strategies we have been able to include the project to which this paper perform as Master Thesis:

1. **Improve accountability**: strengthen, improve and refine the accountability system, through a computerized system and the continued training of civil servants and the sectors involved.

This strategy is our project relational center, although it is not unique. Like this one, other strategies may be related to proving the proper compatibility of our enterprise with the FNDE strategic plan. Thus, we can list them below:

---

[40] EDUCAÇÃO, Assessoria de Comunicação Social do FNDE com informações do Ministério da, **Sobre o FNDE Estratégico - Portal do FNDE**, available at: <http://www.fnde.gov.br/index.php/acoes/fnde-estrategico/sobre/sobre>. Accessed in: 6 abr. 2020.

2. **Educational public policy implementation**: identifying the enforcement capacity effectiveness of the FNDE regarding educational policies formulated by the Ministry of Education.
3. **Technical assistance to government entities and other stakeholders in the educational system**: ensuring technical assistance and financial resource transfers and materials, considering technical criteria, peculiarities, and federated entities' needs, and other educational system stakeholders.
4. **Educational managers and school community formation**: consolidating Corporate Education at FNDE and the ongoing training of managers, employees, and institutional partners, with a focus on improving educational management.
5. **Financial resources for the executors**: to provide conditions for the supply of goods and services necessary for the good functioning of the school and the education strengthening. These financial resources are destined for the execution of the final actions related to knowledge transmission, products, and services.
6. **Student Funding**: consolidate FNDE actions as an agent of the Student Fund and establish mechanisms to expand access services to higher education.
7. **Improve management of educational programs and actions**: optimize the use and financial administration, technological and people resources for the effective execution of finalists' programs, considering the target audience peculiarities.
8. **Develop integrated monitoring and evaluation**: develop methodology and systematization in an integrated manner for monitoring the programs and projects execution financed by FNDE, so that corrective measures can be adopted, and the proposed results achieved.
9. **Encourage socio-environmental management**: propose integrated actions that consider the importance of sustainability for our current and future society, using natural resources and public goods economically and rationally, avoiding waste.

10. **Increase the efficiency of business processes**: optimize and formalize work and management procedures to increase the efficiency of business processes, preserving the institution's knowledge.
11. **Improve communication and interaction with society and external entities**: improve and expand FNDE's communication channels with its various audiences.
12. **Promote innovation management**: implement actions aimed at organizing and disseminating innovation management through the development of institutional projects.
13. **Promote organizational knowledge management**: implement actions aimed at organizing and disseminating internal work knowledge and academic production generated by employees, in line with the FNDE business.
14. **Strengthening internal controls and risk management**: improving actions to strengthen and disseminate internal controls and risk management at FNDE.
15. **Strengthening strategic management and governance**: improving initiatives and projects to strengthen and disseminate good practices in strategic management and governance.
16. **Optimize workforce**: implement workforce sizing methodology to identify institutional needs in a job capacity.
17. **Promote workforce enhancement and social inclusion**: strengthen enhancement policy of the workforce and stimulate the commitment of the server with its permanent qualification and aligned with the specific position attributes.
18. **Promote skills management**: implement mechanisms capable of identifying and preserving the essential and specific skills revealed in the framework of the institution's people for the FNDE's mission and vision achievement.
19. **To promote information technology services modernization**: to provide the institution with adequate technological means, aiming at the celerity and efficiency in the management of educational programs.
20. **To ensure budgetary and financial resources management**: to manage the availability and use of the budgetary and financial limits received by FNDE for the execution of the actions under its responsibility.

It must be observed that *Data Mining Risk Management Applied to Accountability* relates directly to strategic objectives 8, 12, 14, and 15 in addition to item 1, dedicated to accountability performance.

To understand how this project should be implemented, enforcing strategies listed as 1, 8, 12, 14, and 15, we need a view to demonstrating the desired outcome. The final product should demonstrate prototype panels and reports implementation for the statistical model for the classification of ordinary government policies accounts, through data mining including planning for the production deployment of these panels and reports and the knowledge produced by the project, with a summary proposal of other data mining process uses for these policies.

The implementation steps will be further covered in chapter 6.

5.3 PROJECT JUSTIFICATION

While still considering the approach, we ought to review the operational justification. Of course, due to the project economic relevance, as well as its adequacy as a proposal for a new accountability model, this chapter cannot omit to consider the recommendations compatibility from the State entities responsible for the governance control, given the calamitous situation recognized in this project, certainly many were the recommendations from these institutions.

Thanks to the information gathered, complemented by statistical studies contracted through a specialized consultancy, using Big Data tools and cross-referencing databases, including those external to the FNDE, it was possible to build a risk management methodology, using data mining, to classify the accounts submitted to the FNDE.

This methodology relevance has been justified by the possibility of listing which accountability obligations should be prioritized before our operational capacity, determining a possible conclusive analysis based on the misappropriation of statistical risk on federal public resources.

This study, therefore, brings a proposal for administrative rationalization, avoiding a new accumulation of unanalyzed processes. More than that, an abstraction creation to the process of inconsistency verification in public accounts, a so-called "**fine mesh**" by risk management, with the potential to

generate an expectation of control to the Federal Education Resources manager with the advent of its annual publication, making it more vigilant to its management and creating a new model of accounts to the Governmental Authority.

In this context, considering as institutional business "to be a reference in the implementation of public policies", as established in the updated FNDE strategic map[41], this proposal is also justified because of the demands high number coming from judicial and public management control entities determining the conclusive analysis of accountability obligations that practically make ordinary analysis unfeasible and mitigates planning actions to address the problem.

In addition, considering providing a governance solution, and following Joint Normative Instruction MP/CGU No. 01/2016[42], which provides on internal controls, risk management, and governance within the federal executive branch, we have created a pilot project with the most representative programs of rendering of accounts obligations pending analysis, before expanding the initiatives to all financial transfers forms that generate accountability to FNDE.

Following these recommendations, the data mining project meets the need to classify the stock of pending accountability analysis to understand its dynamic behavior, especially in three pilot programs. The algorithm adopted in this project is intended to meet the need for risk management strategy, considering the accounts payable in terms of their chronology, financial aspects, default rates, and cross-referencing information that can help classify the result of accountability according to the risk to the Treasury.

This methodology, in a scientific way, will prioritize analysis by predictive behavior, in detriment to the parity and deterministic analysis of each accountability, to make the capacity of conclusive analysis compatible with the FNDE operational reality.

Finally, to submit this document's contents to an examination of legal adequacy and viability, it is convenient to inform that the FNDE does not consider in this process the mere rationality for the procedural cost of analyzing accounts.

---

[41] EDUCAÇÃO, **Mapa Estratégico - Portal do FNDE**.

[42] NACIONAL, Imprensa, **INSTRUÇÃO NORMATIVA CONJUNTA Nº 1, DE 10 DE MAIO DE 2016 - Imprensa Nacional**, available at: <https://www.in.gov.br/materia>. Accessed in: 20 ago. 2021.

This is an innovative initiative that addresses several other aspects in the risk composition, promoting explicit efficiency to the accountability management.

5.4 PRIORITIZING ACCOUNTABILITY CONCLUSIVE ANALYSES: PROPOSED PREDICTIVE METHOD

In summary, as already described, the current accountability analysis model, in which each financial transfer destined to a public policy execution generates a result counterpart, presupposes a **deterministic** model of conclusive evaluations, where the financial transfer to the FNDE program execution is deterministic to an accountability analysis.

Despite the clear obligation and a deterministic flow, we do not see any obstacles to the other measuring methods adoption of constitutional duty besides the deterministic and parity ones. We have already pointed out that it is a constitutional precept, among others, the principle of efficiency, present in article 37 of the 1988 Federal Constitution[43].

In this sense, and as already mentioned, this Institution has already carried out two situational studies of its accounts rendering, according to a **deterministic parity and traditionally manual analysis method**. We will now look at their findings and how they correlate to this project.

The first one, while studying the administrative rationality of the agency accountability[44], presented the dynamics of the accountability inventories evolution, making clear the imbalance between financial transfers with accountability obligations and the unfeasibility of its equal assessment by the current deterministic model of conclusion by analytical and manual process.

In the meantime, the current methodology is also questionable in terms of operational efficiency, because it is necessary to project which operational capacity in human and logistical resources would be necessary for the analysis of all the accounts rendered from the current stock and to supply the almost 35.000 annual accounts that are currently rendered to the FNDE yearly.

---

43 Constituicao-Compilado, available at: <http://www.planalto.gov.br/ccivil_03/constituicao/constituicaocompilado.htm>. Accessed in: 12 nov. 2019.

44 ARAUJO, Vladimir B., Technical Survey, 2017.

The study, although it presents prescriptibility hypotheses of the punitive claim and administrative rationality, is sufficient to conclude that the current model of accountability is unsustainable, given a capacity for conclusive analysis estimated in less than 4.000 accounts per year.

As for the second study, through consultancy services provided by the Ministry of Education and under the supervision and guidance of this General Coordination of Accountability, EloGroup Consultancy presented a Report on the FNDE Accountability Process Cost Analysis[45]. In this study, among other results with management proposals for the accountability stocks, the consultancy concludes on the **unit administrative cost** of the accountability conclusive analyses and the annual operational cost of FNDE organizational departments.

To establish results, a methodology that states clearly the project gains and prove to add value to the institution must be defined. At this point, we will establish scenarios and set up, methodologically, project costs and effective costs.

5.5 COSTS: PROJECT vs EFFECTIVE

In this section, we establish the implementation cost and listed respective steps in Annexes IV and V, presented as project timetables. Additionally, we detailed the traditional analysis cost from the rendering of accounts as reported. For the implementation to what we should call "**scenario 2**", the costs are disposal at Annex VI, at the end of this paper, amounting to a total established in R$ 192.000,00. We list the direct costs incurred by the project, since the study on procedural costs was fully funded by the Ministry of Education, and the development and integration of BI into SIGPC was fully funded by DIRTI. In both cases, outside our cost center (DIFIN) and with no impact on what we would be authorized to spend. As seen, BI construction for the rendering of the accounts, including the "FNDE Fine Mesh" project representation did not increase project cost, given its budget forecast for the technology directorate, DIRTI.

---

[45] ELOGROUP, Consultoria, **FNDE Accountability Process Cost Analysis Report**, [s.l.]: Ministry of Education / Office of Management of Processes and Strategic Projects / National Fund for the Development of Education, 2017.

The BI would be developed for the FNDE programs accountability and execution, we simply adopted the opportunity to insert project final panels and integration.

For the actual cost to operate the rendering of accounts, what we shall call "**scenario 1**", corresponding to the current situation previously described, we shall consider the study commissioned by the Ministry of Education in 2017, used as a preliminary phase to the work that would trigger this project[46].

By then, FNDE had an inventory of 196.853 accounts to be analyzed. Considering the agency's annual analysis capacity, historically 3.500 conclusive opinions per year, that inventory number tends to increase over the years, which brings a series of losses to the control system and may even impact the public policies financed by these resources.

Considering this issue, the study aimed to determine what was then the analyzing accountability cost of the main programs passed on to educational entities and what would be the impact of reducing liabilities if the accountability obligations whose amount was lower than the analysis cost were filed, exempting a conclusive analysis of a state servant.

To carry out the process of the accounts rendering analysis, the General Coordination of Accountability Monitoring, CGAPC, had 2016, 178 employees, being 25 consultants, 52 servers, 62 outsourced, and 39 interns. In total, in that same year, using this technical staff, it was able to perform the analysis of 3.519 processes, of which 3 educational programs represented up to 80%: PDDE, PNATE, and PNAE, according to SiGPC data[47]. CGAPC is the department responsible for the analysis of the account and for interrupting the resources transfer in the event of an account negative judgment, i.e., if a loss to the treasury is identified.

To perform the work of determining the single cost of an accounting process, the consultancy hired by the Ministry used budget data from FNDE, the data from SiGPC, as well as that from other systems adopted by the public agency. The "organizational sizing method" was also adopted to measure how

---

[46] *Ibid*.

[47] ALVES, **SIGPC - Acesso Público - Portal do FNDE**.

many individuals effectively dedicate themselves to the accountability process and to map its flow and the steps that constitute it[48].

According to the study, the team decided on the "ABC method", or *Activity-Based Costing*, which calculates the cost focusing on the activity performed, not the final product obtained, and seeks to reduce distortions caused by arbitrary costs allocations, such as absorption. Thus, in ABC, the professional identifies the main activities of a unit so that the costs are apportioned according to an indicator that represents the appropriate resources consumption, which is called tracking[49].

By applying the cost drivers researched by the consultancy and relating them to expenses, the study obtained the cost of each organizational unit[50] for the year 2016. Note that the CGAPC organizational unit cost, responsible for accountability processes at the public agency, reaches R$ 20.318.239,51.

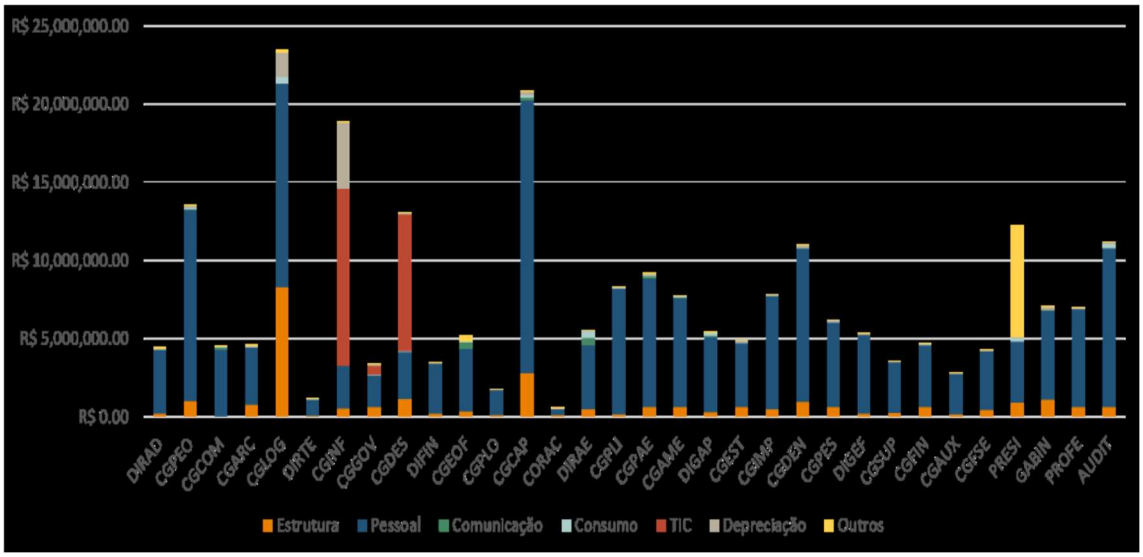


FIGURE 9. FNDE ORGANIZATIONAL UNITS COST. SOURCE: (ELOGROUP, 2017).

The next step in the method was to distribute the support units costs to those performing the finalistic processes in FNDE. As support areas, both the Presidency and the Directorates, according to the method presented, should have their costs distributed to the other units. After calculating CGAPC unit cost

---

[48] ELOGROUP, **Cost Analysis**.

[49] MARTINS, Eliseu, **Contabilidade de Custos**, 10. ed. [s.l.]: Atlas, 2010.

[50] **Quem é quem - Portal do FNDE**, available at: <https://www.fnde.gov.br/index.php/acesso-a-informacao/institucional/quem-e-quem>. Accessed in: 27 abr. 2020.

and determining the cost of the areas that help analyze the rendering of accounts, the study firm concluded that the cost of analyzing the accounts rendering for the PDDE, PNAE, PNATE programs correspond to R$ 13.384,93, R$ 13.115,67, and R$ 12.222,26, respectively [51].

However, as our federal constitution states, **even if the financial transfer does not exceed the process unit cost**, such accounts could not be filed without analysis (scenario 1). But there is no constitutional obstacle for this analysis to be performed through artificial intelligence (scenario 2).

Here is where our proposal is the most innovative. By mining data, from the rendering of accounts already concluded, we shall be able to establish a risk classification that predicts, with high accuracy, the results of an account rendering, being able to consider approved those of low risk and invoking those of high risk for traditional conclusive analysis: by a public servant. Naturally, each result is again considered in machine learning, making the artificial intelligence decision each time more precise.

The artificial intelligence approvals results override the traditional accountability analysis. This is where our **opportunity cost** comes in, where the adequate concept would be better exposed in chapter 7, Results.

To present this project in a didactic and intelligible way to those who are unfamiliar with scientific methodology and techniques, we will discuss the following concepts, work steps, algorithm adopted, development, and prototype results.

5.5.1 DATA MINING

This paper is about learning from data. According to "The Elements of Statistical Learning: Data Mining, Inference, and Prediction"[52], in a typical scenario, we have an outcome measurement, usually quantitative (such as a stock price) or categorical (such as heart attack/no heart attack), that we wish to predict based on a set of features (such as diet and clinical measurements). We

---

[51] ELOGROUP, **Cost Analysis**.

[52] HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J.H., **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**, [s.l.]: Springer, 2009.

have a training set of data, in which we observe the outcome and feature measurements for a set of objects (such as people).

The authors describe that using this data it is possible to build a prediction model, or "learner", which will enable us to predict the outcome for **new unseen objects**. A good learner is one that accurately predicts such an outcome.

The typical scenario for our case is the registered accounts and their set of quantitative characteristics (attributes): loss, federal unit, type of program, amount transferred, impugned value, etc., where we want to predict, based on the study and knowledge of the registered data (past): the result of an account rendered and not yet analyzed (prediction).

As a premise, we will also describe the method used for the desired accuracy. A demand is required by the Federal Government's central internal control (audit) office, CGU.

This work center itself on what is called the "supervised learning" problem. It is called "supervised" because of the outcome variable presence to guide the learning process[53]. In the unsupervised learning problem, according to Hastie *et al*, we observe only the features and have no measurements of the outcome. This project, otherwise, describes how and why the data are organized or clustered.

As a terminology source, we use the classical machine learning language. As we noted in our studies, the statistical literature often names inputs as predictors. In our case, the inputs (predictors) are the accounts rendered that are registered with the system that manages them (SIGPC) and pending human analysis, that is, pending the registration of a conclusive result. The outputs, on the other hand, are called responses, or classically, dependent variables. Our OUTPUT (or outcome) should be the expected accountability result. As we will always emphasize: we look for the balance between INPUT/OUPUT to find out whether the project is successful or not.

We must keep in mind, with the methodology chosen to manage risk, that the desired outputs (predict "loss" or "no loss" outcome) were meant to predict an accounting opinion production. Given some specific characteristics selected in SIGPC in which with "yesterday's" data, we try to predict "today's" results. Given the selected attributes of a registered account, we want to rank and classify it on a predetermined risk scale.

---

[53] *Ibid*.

As we can see, Hastie *et al* [54] comment that there is a distinction in the type of output that leads to a naming convention for prediction tasks: regression when we predict quantitative outcomes, and ranking when we predict qualitative outcomes. These two tasks have much in common, and in particular, both can be seen as a task in function approximation. But this work clearly "classifies" the OUTPUT "rendering of accounts delivered and not analyzed".

The inputs also vary in the type of measurement; we have both quantitative and qualitative input variables. These have also led to distinctions in the types of methods that are used for prediction: some methods are defined more naturally for quantitative inputs, some more naturally for qualitative, and some for both.

In our case, we use attributes found in the SIGPC tables that describe both quantities (loss, number of accounts, number of programs, number of bad results, etc.) and qualities (federation unit, public policy funding program, etc.).

Hastie *et al*[55] describe the development of two simple but powerful forecasting methods: the linear model fit by least squares and the k-nearest-neighbor forecasting rule. The linear model makes huge assumptions about structure and stable yields but possibly inaccurate predictions. The nearest-neighbor method makes very soft structural assumptions: its prediction is often accurate but may be unstable. It is a clustering technique.

Meanwhile, our project uses another option: k-fold cross-validation. The choice was made to avoid a process called "overfitting" when a model is overly adherent to a specific analyses base. This way, the results are more robust in a way that is independent of the input base, avoiding unrealistically good results. We will describe the method for accuracy in due course in section 6.2 Trobleshottings Solutions.

Also, according to Grollman[56], learning algorithms are methods by which data is processed to extract patterns that can later be applied to novel situations. The author states that a system "learns" when the performance of a task is improved, considering a pre-established metric. Given the hypothesis initially established, a learning algorithm should "teach" the system to better choose the

---

[54] *Ibid*.

[55] *Ibid*.

[56] GROLLMAN, Daniel; BILLARD, Aude, Learning Algorithms, *in*: SEEL, Norbert M. (Org.), **Encyclopedia of the Sciences of Learning**, Boston, MA: Springer US, 2012, p. 1766–1769.

accounts to be analyzed. It means to say that "we have no interest" in analyzing accounts that would be approved since their final destination is the archiving.

We should somehow teach the system that stores the accounts to "pick and choose" accounts that would not be approved, to bypass the careful and detailed analysis of the larger set and dedicate the scarce effort to the smaller and more harmful (to Treasure) set of accounts. "A learning algorithm is what is responsible for generating parameters for the model by processing collected data"[57].

Following this project and according to the Technical Note presenting studies performed by the CGAPC department[58], Data Mining is described as the analysis of observational datasets - usually in large volumes - to find unknown relationships and summarize the data in new ways that are understandable and useful to their owner. The relationships and summaries derived from a statistics mining exercise are referred to as models or patterns. Examples include linear equations, rules, clusters, graphs, decision tree structures, and recurrent patterns in time series.

The mining process aims to search for relationships within a set of data that are accurate, convenient, and that represent some aspect of the proposed business in a summarized and useful way. This process usually involves the steps of determining the nature and structure of the representation to be used, deciding how best to quantify and compare the different representations that fit the data, choosing the method and algorithms that will yield the best results, and finally deciding which data management principles are needed to implement the algorithms efficiently.

The data mining methodology used for the FNDE accountability risk management project consisted of CRISP-DM: Cross Industry Standard Process for Data Mining[59]. CRISP-DM is a general-purpose methodology, which is industry-independent, technology-neutral, and it is said to be the facto standard for Data Mining (DM)[60]. Didactically, it can be illustrated according to figure 10.

---

[57] *Ibid*.

[58] ARAUJO; ANANIAS; MACIEL, NT Malha Fina.

[59] CHAPMAN, Pete *et al*, Step-by-step data mining guide, p. 76, .

[60] AZEVEDO, Ana; SANTOS, Manuel, KDD, semma and CRISP-DM: A parallel overview, *in*: , [s.l.: s.n.], 2008, p. 182–185.

This methodology aims to standardize and generate a continuous workflow, which has led to good results in data mining processes.

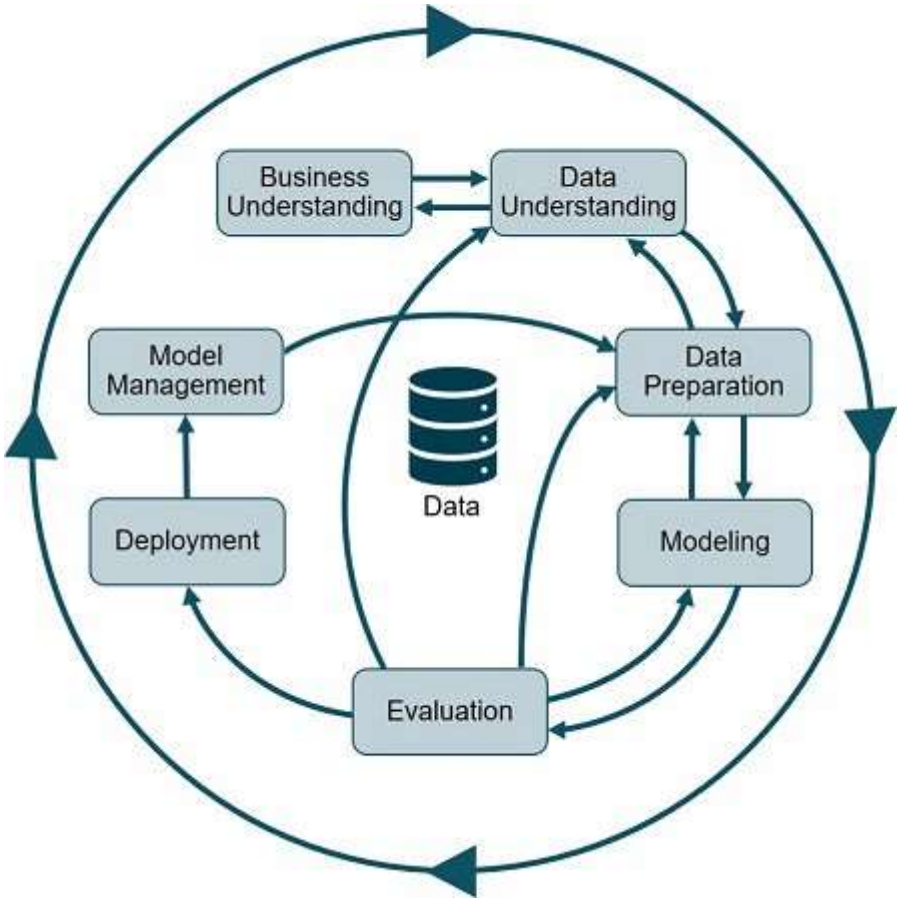


FIGURE 10. PHASES OF THE CRISP-DM REFERENCE MODEL (*CHAPMAN ET AL. - STEP-BY-STEP DATA MINING GUIDE.PDF*, N.D.).

As this work is not about a methodological data science review, but rather uses its applications to achieve the project's objectives, we will only visit basic concepts to understand the decisions, as a project manager, to achieve our pre-established objectives.

Figure 10 shows the current process model for data mining providing an overview of the data mining project life cycle. Observe that it contains the project phases and respective tasks already established, including the relationships between these tasks. The methodology, therefore, being pre-structured, enables the construction of an adequate work breakdown structure, as laid out in the appendix of this paper.

Of course, at this description level, it is not possible to identify all relationships, which were established throughout the development as the data mining tasks were assigned to the consultants, scaling these tasks up as we tried

to achieve the project objectives. In this process, we have learned that more important than the mining technique itself, which is just the tool chosen, is to **know what to actually "ask" to the algorithm** (or technicians): the whole process is centered on this aspect. Hence, it is crucial to know your institutional process intimately, otherwise, you will inevitably delay the enterprise development.

The data mining project life cycle consists of six phases[61], shown in figure 10. It will be further detailed in item 6.2 of this paper.

It must be clarified that these phases' sequence is not rigid. It could be active between different stages as required. It is the outcome of each phase that will determine how each step or task is to be accomplished. Hence the need for proximity between the project manager and technicians, as well as the manager's knowledge of the business rules.

Chapman explains that the arrows indicate the most important and frequent dependencies between phases, with the outer circle symbolizing the cyclical nature of data mining itself. Evidently, data mining does not end once a the solution is deployed.

As we will demonstrate, the lessons learned during the process and from the deployed solution will enable new gains and even new processes not visualized before, often more focused business questions. It is a natural process in which the data study generates extensive information, knowledge. In addition, there is the whole continuous learning process and post-implementation improvement.

6.TROUBLESHOOTING SOLUTIONS AND IMPLEMENTATION STEPS

6.1 IMPLEMENTATION STEPS

Once we have decided on the predictive model and knowing costs (both, project and opportunity cost), we can describe the implementation steps. The effect expected to this initiative is to analyze account performance accordingly to the human capacity available, avoiding the generation of accounts rendered and pending traditional conclusions, therefore, avoiding new inventories formation.

---

[61] CHAPMAN *et al*, Step-by-step data mining guide.

For indicator, as the project concluded, we propose a *New Inventories Generation Rate*. Indicators for PNAE, PNATE, and PDDE should be 0%. It means that the algorithm concludes by the statistical approval of some obligations to be accountable and will bring the need for other traditional analyses within operational capacity.

The main activities overview demonstrates these project requirements, as well as describe the resources needed to overcome them:

- Available computational infrastructure analysis, considering its adequacy to the project purposes, including the proposal of feasible improvements to reach the ideal scenario.
- Methodology to support project development under PNAE, PNATE, and PDDE programs, as well as the respective schedules, including milestones and partial deliveries.
- Data available on the SiGPC basis for statistical analysis, including the technical and financial profiles of PNAE, PNATE, and PDDE accounts, considering 2011 to 2016 intervals.
- PNAE, PNATE, and PDDE business rules description, related to the technical and financial analyzes of the rendering of accounts, evidencing the changes occurred year by year, in the period selected.
- Technical and business requirements proposal, including detailed project objectives, to subsidize data mining project, as well as prototypes of the panels and reports that will be made available, regarding PNAE, PNATE, and PDDE.
- Models and algorithms, duly validated, to subsidize the selection of accounts for PNAE, PNATE, and PDDE, with a greater probability of not achieving the program objectives and/or public resources embezzlement, taking into account the technical and financial aspects, including the years 2011 to 2016, including also shreds of evidence and summaries of the results obtained when executing these models and algorithms in the statistical environment R[62], and the respective source code.

[62] **R: The R Project for Statistical Computing**, available at: <https://www.r-project.org/>. Accessed in: 3 mar. 2020.

These steps can be summarized in a Work Breakdown Structure (WBS), demonstrating milestones to be overcome from initial planning to implementing evidence. The WBS is shown in Annex III, when we consider each deliverable activity, including academic. Likewise, a schedule was initially established as per Annex IV and V. For the 2020 schedule, the steps related to academic production, starting in August 2020, have been shifted to 2021 due to the COVID-19 pandemic. Other previous steps were maintained and completed in the project as scheduled.

6.2 TROUBLESHOOTING SOLUTIONS

CRISP-DM proposes an iterative process flow, with non-strictly defined loops between phases, and the overall iterative cyclical nature of the DM project itself. The outcome of each phase determines which stage must be performed next [63].

According to Olegas Niakšu, six phases of CRISP-DM are:

1. Business understanding,
2. Data understanding,
3. Data preparation,
4. Modeling,
5. Evaluation, and
6. Deployment

Based on the CRISP-DM methodology for understanding the data contained in the FNDE accountability systems, several meetings were held with those responsible for managing the programs and accountability of PNAE, PNATE, and PDDE, including technical and financial analysis and SIGPC system management, as an accountability electronic repository.

It was noted during this process that it would be necessary to identify the attributes useful for mining, given that the databases suggested by the various responsible parties had not yet been analyzed and the data, not necessarily, were already structured. Thus, the project team chose to use data from SIGECON

[63] Niakšu - CRISP Data Mining Methodology Extension for Medica.pdf.

(Council Management System), SIGEF (Integrated Financial Management System), and SIGPC (Accountability Management System).

During our work, in a collaborative activity with the data mining team of the Ministry of the Controller General of the Union (CGU), we created an index for identifying abnormal prices, using the PNAE price databases contained in the SIGPC. This information was later incorporated into the final model.

It should be noted that data modeling, in other words, understanding and preparing the data, is the most arduous and time-consuming step in the data mining process, been the main goal during most of the entire project. This work result identifies the main dimensions to be used in the final model. Once modeling is complete, definitions are assigned to generate the necessary results.

In FNDE, an account rendering may have five final states: approved, approved with exceptions, partially approved, partially approved with exceptions, and not approved. Initially, it was expected that the statistical model could classify an account rendering in one of these categories, however, as the number of accounts rendered effectively analyzed is low (considering the stock available), it was decided that the best model accuracy was to separate the five available final states into two groups: with "loss" to the treasury and "without loss" to the treasury. Where the classifications that lead to the conclusion "loss" to the treasury are Partially Approved, Partially Approved with exceptions, and Not Approved. Obviously, the classifications that do not lead to the conclusion of losses to the treasury are Approved and Approved with exceptions.

Once the response variable and the attributes to feed the model were defined, the eXtreme Gradient Boosting or *XGboost*[64] algorithm was chosen to predict the identification probability of whether accountability can generate some loss to the treasury. This modeling type is called supervised learning, since the machine is presented with data that has already been manually analyzed, and the algorithm used "learns" which patterns in the data define and lead to a certain classification. We also considered the need for result efficiency (accuracy), which is the method adopted one of the most powerful and used in data science.

---

[64] CHEN, Tianqi; GUESTRIN, Carlos, XGBoost: A Scalable Tree Boosting System, **Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16**, p. 785–794, 2016.

The idea behind the algorithm is quite simple. Assuming that "small learning" can generate "great learning", it is possible to build a method that learns in each iteration with the errors generated in the previous procedure.

In general, the algorithm initially generates a decision tree, in which all individuals in the training base are classified. From the second iteration on, the residues or "errors" of the previous interaction are used to generate a new tree, which tends to err less than the previous one. These steps are repeated a finite number of times until the generation of a new tree does not generate more gains. This method is called *Gradient Boosting*[65]. With the algorithm defined and the data selected, a prototype was then created.

Due to the need that all FNDE accounts be submitted to technical and financial regularity analysis, by normative rule, variables and indicators that could represent these two analysis dimensions were found and created, avoiding legal rejection anyhow. The project then formed grouped attributes as follows:

- Financial attributes: data and indicators of bank account balance, bank statements, bank orders from the entities for each concession, among others.
- Attributes related to the Social Councils: questionnaires analysis, and opinions from the councils.
- Time attributes fiscal year, the period until the entity is accountable, or the period beyond the deadline for accountability.
- Attributes related to price: prices indicator charged by the federal entity.
- Attributes related to the types of contracting instruments used by the entities.
- Attributes related to the Family Farming policy, PNAE
- Attributes related to the non-compliance entity chronology.

The complete attributes set used can be found in Annex II.

In any case, it is convenient to didactically explain the reason for the choice of search attributes for a better understanding of the desired view. In data mining, the "importance of knowing how to query" is usually emphasized. In this aspect, we need to "ask SIGPC", employing the statistical software "R", the "search questions" listed in Annex II, to obtain, by combining this process and the ranking

[65] Li - A Gentle Introduction to Gradient Boosting.pdf.

attributed to the questions, the answer to the following issue: "which account rendering should be considered as high risk of loss to the public treasury". In annex II, next to each attribute, there is a brief description.

SIGPC, therefore, is the transactional system. As a new account is entered into the system, with its various classifications (attributes), it generates a single sequential number (NUNSEQ). Each attribute is the generation of a table extracted from SIGPC, the proper combination according to supervised machine learning and developers' experience, will determine the process result.

At this point, we remind you that this work is not about data mining. This is only the technique used. This is a project management study. As a manager, we will explain why, of all the available attributes, we chose the ones presented here as the best path in our judgment (of course, there are others), so that the algorithm can find the answer to our question.

Due to manual analysis limitations and the large existing database, it was possible to obtain for the training base 1.222 observations, with classification defined in the five aforementioned analysis states. Once the project is deployed in production, a sure recommendation will be to gradually enlarge the model's training sample with new traditional analyses performed by the civil servants.

Finally, the evaluation phase consists of the final model deployment. It is a review of the construction steps model, making sure that business CRISP-DM objectives, in our case, accuracy, are properly achieved. For this project, model accuracy would be ideal as close as possible to 90% to be valid, when comparing the training sample and validation sample. The result for this phase defines whether the DM results are to be used in practical settings: to record account classification data to SiGPC.

To illustrate the Cross-Validation method, it will be worthwhile to understand the following figure 11:

FIGURE 11. K-FOLD CROSS-VALIDATION. SOURSE: SEBASTIANRASCHKA

In this sample, 10 random data groups equally sized are spat, commonly called “folds” and represented as “k”[66]. If we use 9 folds to build a model, applying it to the remaining fold initially let out, this will return the accuracy metric. Repeating it “k” times (for the example, 10), we will have 10 (or “k”) different accuracies. The result will be the average of them.

In our project, we use a separate random sample in k-folds without an intersection. For the first fold, we use a complementary part to it by training the model and making the algorithmic prediction for the fold. The process is repeated for k-folds: each one trains with different models until all the training base applications are used. Then, the accuracy percentage is observed. The average is the one to evaluate if a model is good or not. For this project, as said before, the model will be as good as it gets closer to 90%.

The accuracy of the "fine mesh" prototype is around 90%, with a 95% confidence interval equal to the CI: (0,8881; 0,9217). Table 2 shows the predicted and reference ratings used to train the model and their respective percentages. It is observed that the percentages of false positives and false negatives are low.

[66] RASCHKA, Sebastian, **How do I evaluate a model?**, Dr. Sebastian Raschka, available at: <https://sebastianraschka.com/faq/docs/evaluate-a-model.html>. Accessed in: 3 mar. 2020.

	Reference		P_Reference	
Prediction	**Loss**	**NO Loss**	**P_Loss**	**P_NO Loss**
Loss	580	58	47,46%	5,75%
NO Loss	57	527	4,66%	43,13

TABLE 2. MODEL CLASSIFICATION PARTIAL RESULTS

According to table 2, the yellow result gives us a "false positive". It means that there is a 5,75% chance, for the learning sample in the partial results, that a result will be classified with treasury loss and be false. Similarly, there is a 4,66% chance ("false negative") that a sample is classified as regular, but a confirmatory manual analysis does in fact contain treasury loss. The partial results have shown promising results, indicating model continuity as designed. However, as will be detailed in the conclusion, there is a need for continued revision (augmentation) of the model's training sample.

The study built by the two data science consultants hired for the project and presented by the Technical Note DIFIN/FNDE No. 04/2018[67] demonstrates the proposed model classification. This can be illustrated by figure 12. The algorithm proves to be very effective because the further the curve is away from the dotted line in the center, the better the classification. The Area Under the Curve (AUC), is a fit quality measure for the model, where the closer it is to 1, the more effective the classification is. For the prototype, the AUC is approximately 0,96

The vertical axis shows the true positive rate, while the horizontal axis shows the false positive rate. In signal detection theory, the Receiver Operating Characteristic Curve, or simply ROC curve, is a graphical representation that illustrates the performance of a binary classifier system and how its discrimination threshold is varied[68].

[67] ARAUJO; ANANIAS; MACIEL, NT Malha Fina.

[68] FAWCETT, Tom, An introduction to ROC analysis, **Pattern Recognition Letters**, v. 27, n. 8, p. 861–874, 2006.

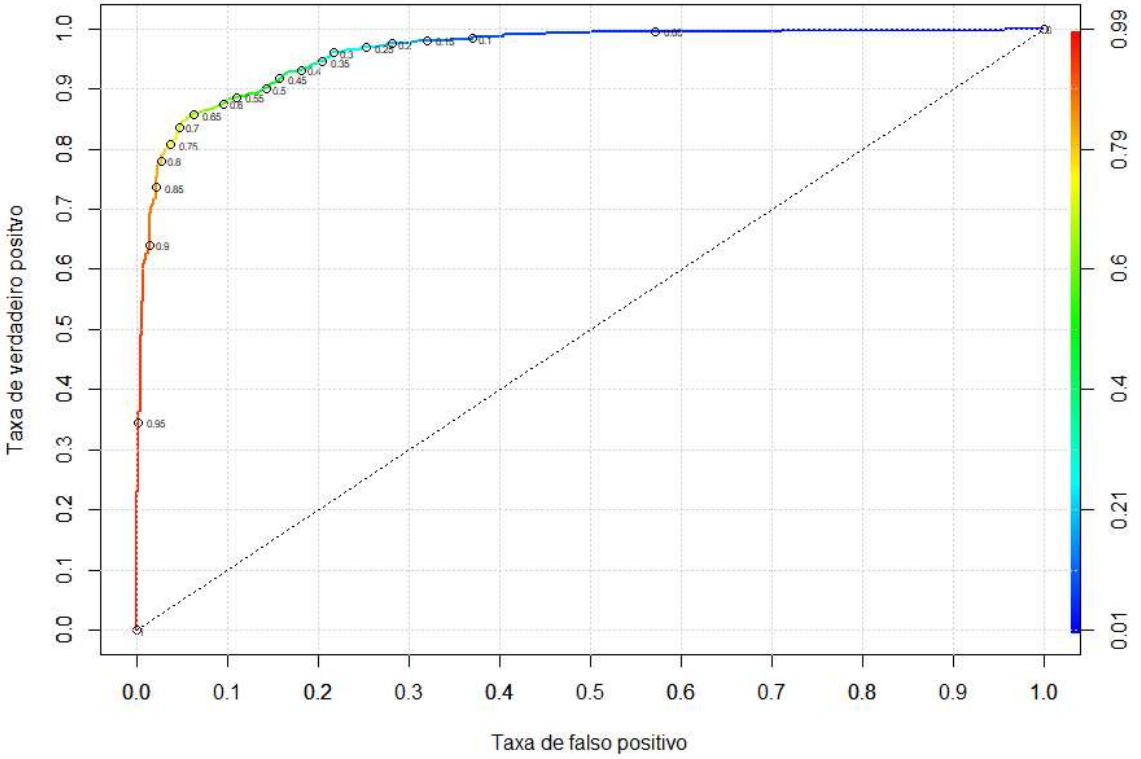


FIGURE 12. ROC CURVE OF THE PROPOSED MODEL.SOURCE: SOURCE (ARAUJO ET AL., 2018)

The ROC curve is a plot of sensitivity versus specificity (often called the false-positive rate) that offers a summary of sensitivity and specificity across a range of cut points for a continuous predictor. The area under the curve, or c-statistic, ranges from 0,5 (no discrimination) to a theoretical maximum of 1[69].

Nancy R. Cook[70] explains that a "perfect discrimination corresponds to a c-statistic of 1 and is achieved if the scores for all the cases are higher than those for all the non-cases, with no overlap". She continues by stating that "the c-statistic is equivalent to the probability that the measure or predicted risk is higher for a case than for a noncase".

[69] COOK, Nancy R., Use and Misuse of the Receiver Operating Characteristic Curve in Risk Prediction, **Circulation**, v. 115, n. 7, p. 928–935, 2007.

[70] *Ibid.*

What we wish to clarify is that c-statistic is not the probability that individuals (or simply the subject investigated) will be classified correctly, or that a subject with a high test score will turn out to be a case. In this hypothesis, there is more proximity to the predictive value or the probability of the risk that we monitor according to the test result.

The c-statistic also describes how well models can rank order cases and noncases. For our project, the cases are typically above 0,7 on the curve. For the noncases, from 0 to 0,7, we sub-classify the ranges to include hypotheses far from the extremes (where there is more precision). We define a range for "approval" and a range for "suspension" differently, preventing the case due to classification imprecision.

Given the model good qualities, it is up to the FNDE's top management to suggest in its own regulation which of the predictive ranges presented is considered adequate for what we call the "fine mesh": approval/approval with reservations, and suspension in the face of statistical impracticability of making a decision with adequate accuracy. Thus, we suggest the template below be considered as a decision proposal.

An example of the risk ranges to be considered for the algorithm is shown in table 3:

.Risk Ranges to be defined	Algorithm decision	Accuracy Percentage
Greater than or equal to 0 and less than 0,3	approval/approval with exceptions	90%
Greater than or equal to 0,3 and less than 0,7	suspended	60%
Greater than or equal to 0,7 and less than or equal to 1	"fine mesh"	93%

TABLE 3. ILLUSTRATIVE EXAMPLE WITH THE RISK RANGES FOR EACH CLASSIFICATION. SOURCE (ARAUJO ET AL., 2018).

The project, utilizing a predictive algorithm, proposes to provide an adequate solution to the unanalyzed rendering of accounts stock and can, eventually, be a valid methodology to totally supply this need, finally bringing balance to FNDE public accounts. Our proposal consists in segmenting the result into 3 parts:

I. **Automatic approval of accounts rendered**: considering the low probability of losses to the treasury presented by the algorithm, all the accounts rendered with a probability higher or equal to 0 and lower than 0,3 to the risk range of approval or approval with qualifications would be considered as definitively approved with a precision up to 95%.

II. **The suspension under uncertainty of result conclusion**: all the accounts with probability greater than or equal to 0,3 and less than 0,7 to the risk range of approval or approval with qualifications would be suspended, so that they could be analyzed in case of inspection, or reclassified in the future with the algorithm improvements.

III. **"FNDE Fine Mesh"**: all account renderings with probability greater than or equal to 0,7 would be classified with a very high risk of causing losses to the public treasury with accuracy up to 95%. The order of analysis is given by the highest probability and highest value.

It must be pointed out, however, that there are several limitations to this project's success. Some of them still persist. At this point, we will address them before presenting the work breakdown structure guideline.

During the data mining process, a few points were found that aggravate the FNDE accountability process. The PDDE proved to be the main generator of new accountability backlogs. The problem cause is the program's splitting into several "sub-programs", each one subject to a dedicated accountability obligation, creating the need for the accountable entity to meet its constitutional obligation successively to the same finalistic program several times and for the same fiscal year.

The statistical analysis processes revealed that this program, in particular, is made up of a services portfolio directly destined to the so-called Executing Units (Uex), **which could be clearly summarized as the execution of a single program** [71].

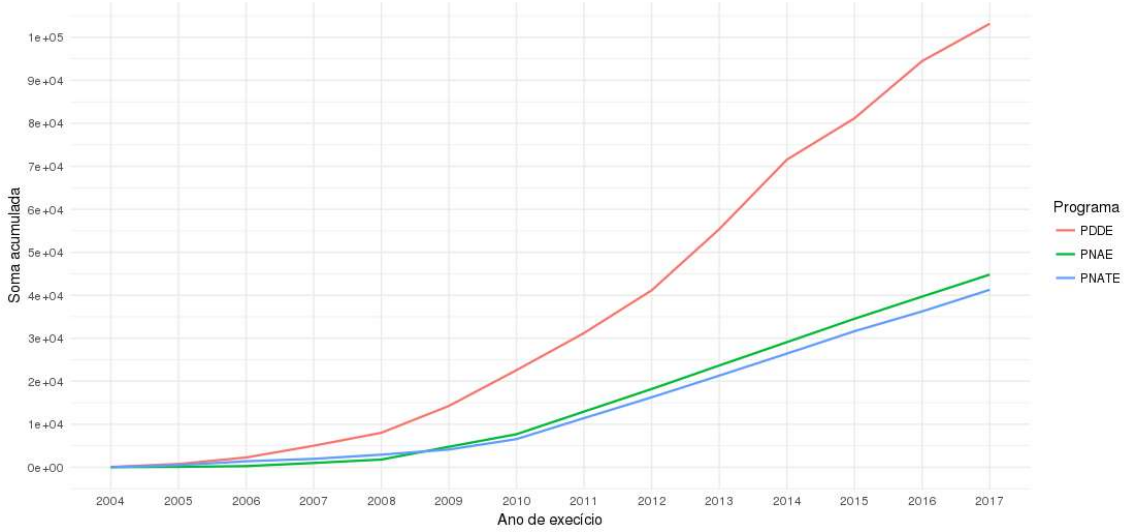


FIGURE 13. ACCUMULATED SUM OF PDDE, PNATE AND PNAE ACCOUNTS RENDERED OVER THE YEARS. SOURCE (ARAUJO ET AL., 2018)

Adopting a solution to this problem, grouping PDDE rendering of accounts into a single accountability obligation, maintaining the specificities of each action as a list of services objectively applicable to the program (providing autonomy and discretion to the UEx manager, who would manage only one specific bank account for the program), would provide about 33% reduction in the annual new stocks increasing of accounts rendered.

[71] ARAUJO; ANANIAS; MACIEL, NT Malha Fina.

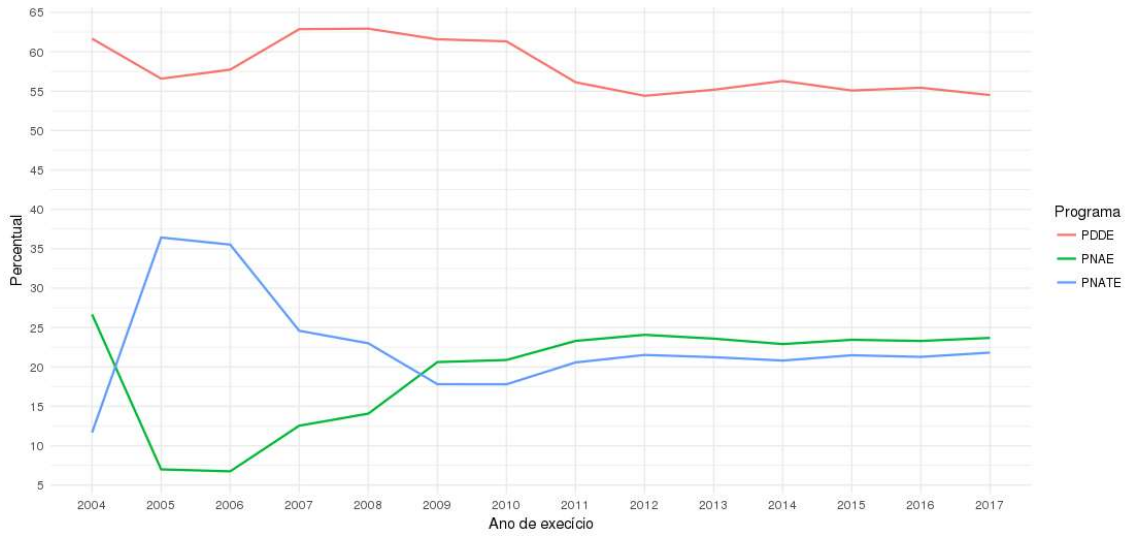


FIGURE 14. PDDE, PNAE AND PNATE ACCOUNTS RENDERED PER YEAR. SOURCE (ARAUJO ET AL., 2018).

Another important point is the communication improvement with the financial institutions that manage the entity's bank accounts from FNDE programs. It was found to be difficult to obtain data from certain entities, as these managed to check accounts in different banks from those indicated by the FNDE. The lack of standardization in the bank transactions description is also noteworthy, which hinders the analysis of the financial resources used by the beneficiaries.

Even if the SIGPC automation process has been interrupted, manual analysis demands a large number of resources and time. Thus, resuming automation would necessarily lead to a considerably more relevant traditional analysis capacity, certainly adding value to a future manual analysis by risk-based predictive choice.

At this point, it is recommended to apply the automated analysis already developed for PNAE 2011/2012 to the other FNDE programs in SIGPC[72]. We found that there is financial rules standardization and adopting as a technical rule the same financial charges criteria already adopted for the PNAE module, it is assumed that all current stock can be subjected to the same model of initial automation treatment. The expectation is to contribute to the algorithm's machine

[72] Among dozens of accountability receipt modules at SIGPC that represent each program funded by FNDE, this was the only one completed in ten years of work.

learning, possibly generating results for those accounts where mining has been inconclusive.

As noted, despite the troubleshooting encountered, one of the positive side effects of consistent data analysis is the information useful establishment for decision making and eventual data governance[73]. Hence, the project results have been evidenced utilizing panels, built as an information aggregation: Business Intelligence.

The results can be shown in an appropriate panel built by the FNDE's Technology and Innovation Directorate (DIRTI), in Microstrategy language[74]. They will be better presented in chapter 7 of this paper.

Finally, a limitation that persists given the government agency's characteristics is the lack of specialized personnel in data science. Although the entity has at its disposal more than enough resources to hire specialized scientists, these, for not belonging to the functional staff, are a constant risk of work discontinuity.

To supply this demand definitively, it is proposed to establish a Technical Cooperation Agreement between the Ministry of Education, through the National Fund for Education Development (FNDE), and the Office of the Comptroller General (CGU), aiming at the exchange of data and technologies in the scope of the "Fine Mesh", related to the rendering of accounts of FNDE programs[75].

The agreement brings mutual responsibilities between FNDE and CGU. This Federal Executive Branch Central Office of Internal Control has a department dedicated to data science, statistical analysis, and the development of predictive tools to help government management control. In the agreement, CGU assists in the development, improvement, and enterprise reliability. While FNDE provides the control body with data and information exchange. The strategy of using a related power institution that advocates for the project have

---

[73] RÊGO, **GESTAO E GOVERNANCA DE DADOS : promovendo dados como ativo de valor nas empresas.**

[74] **Discover MicroStrategy 2021**, MicroStrategy, available at: <https://www.microstrategy.com/en/get-started/2021>. Accessed in: 23 ago. 2021.

[75] UNIÃO (CGU), Brasil Controladoria-Geral da; EDUCAÇÃO (FNDE), Brasil Fundo Nacional de Desenvolvimento da, Acordo de Cooperação Técnica n. 19. de 18 de maio de 2020, **Diário Oficial da União n. 95, de 20 de maio de 2020, seção 3, p. 93**, 2020.

yielded good results, the agreement was concluded on May 18, 2020. Steps and responsibilities for each institution are listed in the following table:

#	Responsável	Atividade
1	FNDE	Data and information available related to the programs execution and rendering of accounts, and from the model initially developed by FNDE
2	FNDE	Other programs analysis with accounts rendered in SIGPC and SIMEC (materiality, quantity, data availability, etc.)
3	CGU	Model analysis, aggregation of other data available at CGU, and development of a proposal for improvement
4	FNDE	Account rendering sample Analysis for model validation.
5	FNDE/CGU	Work process definition in both institutions.
6	CGU	Classification process implementation and execution
7	FNDE	Risk appetite definition and integration of the rating into the accountability process
8	FNDE	Periodic sending of the data required for process execution
9	FNDE/CGU	FNDE's accountability and other programs evaluation to create a plan for the development of new models

TABLE 4. STEPS FORESEEN IN THE TECHNICAL COOPERATION AGREEMENT 19/2020 - FNDE/CGU

7. RESULTS

7.1 OPPORTUNITY COST

As an economic concept, we have established that individuals and institutions must make choices because scarcity exists. In other words, seemingly unlimited wants are confronted with limited resources: some *wants* must be unsatisfied and we must choose which *wants* will be satisfied and which we will not. The most valuable opportunity, or missed alternative, in choosing is known as opportunity cost. Each time a choice is made, an opportunity cost is incurred.

In this sense, consider that of a total non-analyzed inventory, PDDE, PNAE, and PNATE represent, as last updated, 119.393 renderings of accounts delivered to FNDE and not conclusively analyzed. Our predictive model, in the current results, recommends the predictive approval due to the low risk to the Brazilian Treasury of 99.793 accounts (83,58%), according to MicroStrategy BI from SiGPC developed by Technology and Innovation Directorate, DIRTI, to support this project.

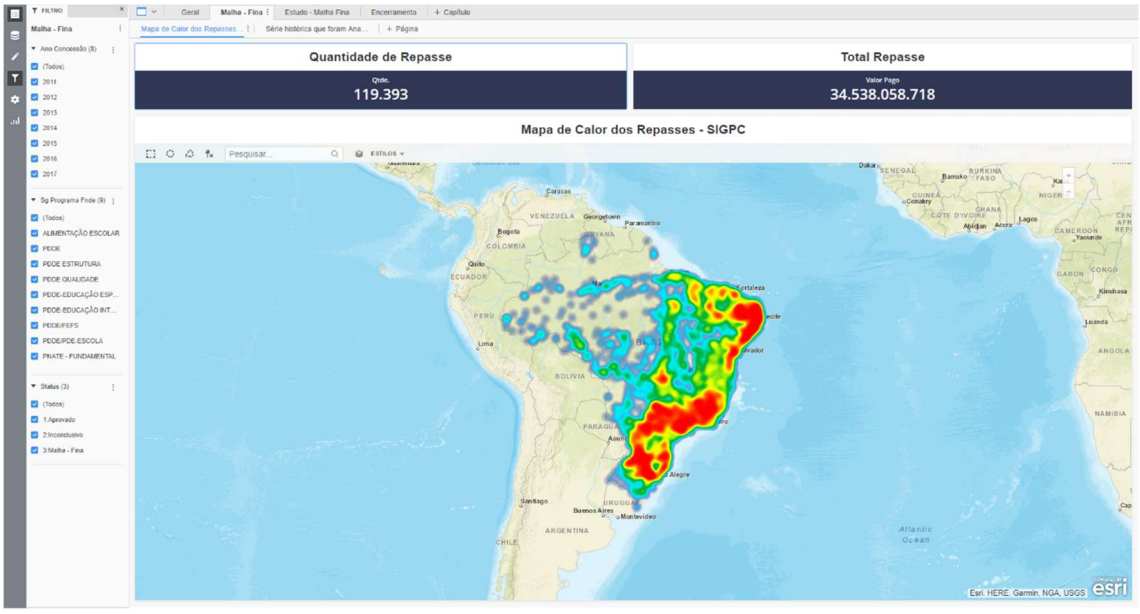


FIGURE 15. UNANALYZED RENDERING OF ACCOUNTS HEAT-MAP IN SIGPC. SOURCE: SIGPC BI/FNDE

Figure 15 shows the BI graphical representation of accounts rendered. In it, we have 119.393 accounts delivered and not analyzed for the three public policies listed in this pilot project, the object of this master's degree, to the value of R$ 34.5 billion. An additional construction advantage from this dashboard is the various ways of representing the information avocated. Here, we have a heat-map, in which the highest concentrations of accounts rendered, and consequently of financial transfers, are found in the " red-hot zones". These, for example, should be considered in terms of criticality and materiality in eventual public policy audit or monitoring planning.

For a sample range of almost 120 thousand accounts in SIGPC submitted to the statistical predictive process of risk analysis to the Treasury, representing more than R$ 34.5 billion in concessions to the three programs, the algorithm was able to classify the accounts according to the risk of non-performance by financial loss, with up to 93% accuracy[76]. Our work revealed that the predictive method is more accurate in pointing out the accounts with higher risk, precisely those that should be traditionally analyzed:

- 99,79 thousand concessions would be statistically compliant, representing R$ 25.90 billion in transferred values;
- 6.290 concessions would be statistically non-compliant, representing R$1.83 billion in transferred values;
- 13,31 thousand accounts were rendered classified as "inconclusive", representing up to R$ 6.80 billion.

[76] ARAUJO; ANANIAS; MACIEL, NT Malha Fina.

In the figure, the pinpoints predictively classified at high risk as so-called "Fine Mesh" are georeferenced in red. These should be subjected to traditional analysis by a public servant and the results fed back into the learning sample for model accuracy.

FIGURE 16. PLOTTING THE "FINE MESH" RESULTS. SOURCE: SIGPC BI/FNDE

The pinpoints labeled in yellow were classified by the model with inadequate precision and, therefore, cannot be filed although their analysis order is later than the ones labeled in red.

Those rated in green should be approved due to the low risk verified by the predictive rating, but it will be up to FNDE to define a sample of these accounts to confirm the result by manual analysis. The final classification should again be inserted into the learning sample with the intent of improving the model's classification.

Eventually, as the model evolves its machine learning, inconclusive classifications should gradually be redefined as "fine mesh" or "approved".

Note there is consistency in the algorithmic prediction in the historical series that follows the SIGPC data insertion beginning. In 2016, there is an apparent discrepancy, but this does not compromise the results achieved. It is up to the FNDE to investigate the data quality in a future project or to invest in the model accuracy by expanding the learning sample with more traditional analyses available to the model.

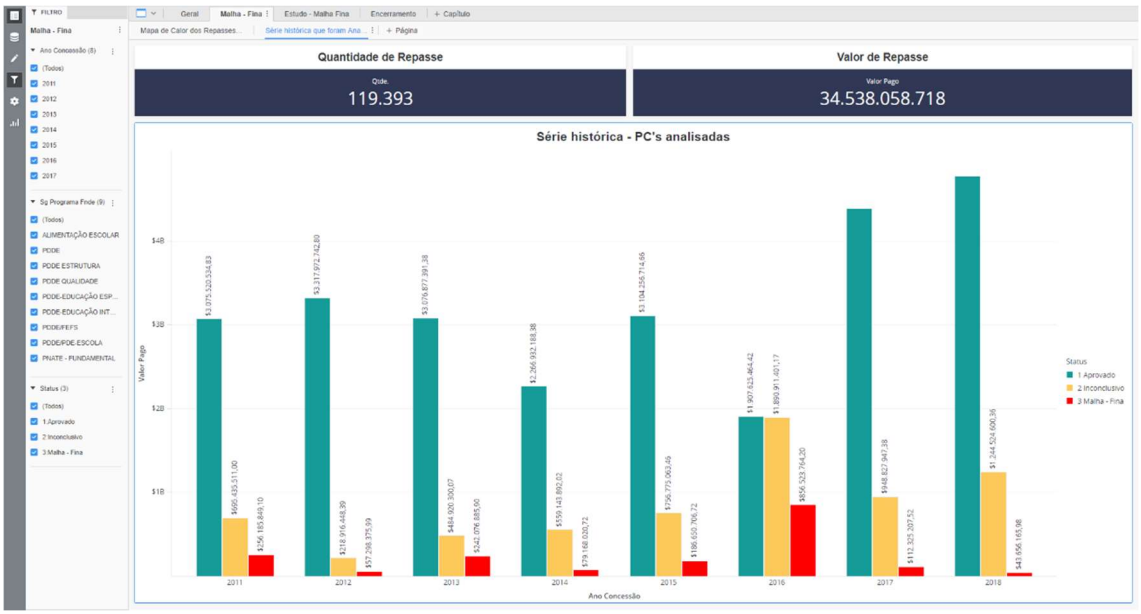


FIGURE 17. HISTORICAL SERIES SINCE THE SIGPC CONCEPTION

Now let us consider that the average cost of accountability for these programs corresponds to R$12.907[77], being the effective cost each time accountability for these three programs was analyzed, regardless of the financial amount transferred to the public policy execution. In this way, for the analysis of these accounts, effectively R$ 1.288.028.251[78] would be spent (scenario1).

The opportunity costs are incurred when choices are made[79]. If the institution does not choose the most innovative alternative, the opportunity cost would be up to R$ 1.3 billion.

Besides, the FNDE operational capacity, as already mentioned, produces only close to 3.500 conclusive opinions per year. The proposed solution puts forward at least 28 years of ordinary operation (scenario 2).

---

[77] € 2.068,43, Euro value at 09/09/2021.

[78] € 206,4 million, Euro value at 09/09/2021.

[79] Arnold - 2010 - Economics.pdf.

Considering that of the approximately 35 thousand accounts, at least 22 thousand belong to these three programs chosen for this project each year, maintaining the expectations and approval and consequent filing, we would have more than 18 thousand accounts approved and filed by artificial intelligence. This corresponds to at least R$ 232 million[80] per year in procedural rationality.

We must consider FNDE has other public policies that, although not as representative in quantity, are as materialistic. These policies concern student financing, schools construction and repair, school vehicles purchase, and more. The project outcomes highlight that it is viable to suggest other programs, even if they are found in other systems, to be subjected to risk classification, because, as our results prove, the deterministic parity model of accountability analysis is outdated and obsolete in the face of growing demands for public resources and the available means scarcity for the results accountability demonstration to society.

7.2 INVENTORY GROWTH RATE

The appropriate indicator should measure project success after implementation: Inventory Growth Rate. The target rate goal for generating new inventories for PNAE, PNATE, and PDDE, initially stated as ideally 0%. It means that the algorithm predictably concludes by approving some accountability obligations and rates others to traditional analysis, respecting the operational capacity established for the process.

To depict the real viability the project now exposes, notice annual accounts rendering OUTPUT have reduced from 35.000 to a little over 6.000 per year. From figure 18, we see the advantage of georeferenced information for high-risk rating:

[80] € 34,5 million, Euro value at 09/09/2021.

FIGURE 18. DASHBOARD FOR "FINE MESH" GROUPED BY REGION.

As a collateral result, the manager can program public policy monitoring according to the high-risk concentration, in an attempt to provide dynamics to the monitored public policy risk map. Of course, the resources beneficiary expectation is not to be part of this rank. Naturally, his behavior shall change, affecting the INPUT data quality, the caution with the information provided to SIGPC, the documents trustworthiness, etc. This is what we call the vigilant expectation of state control, which the funds administrators from FNDE certainly did not have.

Extinction of the entire stock of non-analyzed accounts has been FNDE greatest operational challenge since its foundation as a public agency. With the project results, we expect that in up to five years this liability will be extinguished and FNDE will no longer form new accumulations of unanalyzed financial obligations each financial year.

To be both viable and realistic, the process must continue machine learning and technological evolution. At this point, it is recommended that the agency create a post-project action plan. For the first moment, by redirecting the current workforce and temporary recruitment, is possible to eliminate the high-risk classifications.

To avoid the algorithm becoming too accurate in getting high-risk scores correct but unfit to improve its accuracy in "approval", at least 10% of this "task force" should be dedicated to analyzing "low risk" accounts, decreasing the exposure of machine learning to the "false negative".

A simple work plan, respecting FNDE operational capacity and with the correct decrease from the chosen Key Performance Indicator (KPI) can be followed as shown in the following table:

YEAR	OUTPUT	"FINE MESH" ANALYSIS	"APPROVED" SAMPLE	KPI
1	6.290	3.397	377	0,4
2	6.016	3.790	421	0,3
3	5.305	3.819	424	0,2
4	4.561	3.694	410	0,1
5	3.956	3.560	396	0

TABLE 5. SUGGESTED WORKPLAN FOR THE NEXT 5 YEARS.

For the first year, respecting the operational capacity, only 60% of OUTPUT is handled. For the following years, maintaining the classification of INPUT 90%, and addressing the previous year residual, gradually we will have the desired KPI decrease, until the optimal point in the fifth year post-project.

In this period, it is expected that inconclusive classification would be gradually eliminated as learning evolves. For this range, the accounts provided should simply be resubmitted to the algorithm at each new learning cycle, until they are classified in the desired target zones. At the optimal point, where every 35.000 INPUT we have 100% of the classification/analysis set, with OUTPUT filing or analysis (high-risk rate), we expect minimal or no classification outside the desired ratings.

Following this decreasing pattern and respecting the operational capacity, after five years it is possible to achieve the desired KPI for the 3 programs considered in the initial project.

7.3 OUTCOMES FOR S.M.A.R.T. OBJECTIVES

Finally, we can unveil the S.M.A.R.T. objectives in the outcomes achieved, projected to the near future, or even not achieved.

- **Specific** – the risk management methodology has, in fact, been completed. It has proven to be not only plausible but also revolutionary, a solution that replaces the deterministic parity model of accountability analysis that will now give way to a predictive model based on scoring. Specifically, the accounts with a high rating, classified as "high risk", are selected by the SIGPC system, for traditional analysis by FNDE's technicians. Those classified with a "low risk" rating are filed as approved. As a rule, the approved accounts may, at any time, be called by the existing legal means (judiciary, TCU determination, CGU, or prosecutors) for manual review and result confirmation, as determined by a specific legal process.
- **Measurable** – The Key Performance Indicator (KPI) chosen has proven adequate and fair for evaluating compliance. The model does classify accounts as intended, although it still has a relevant classification as "undefined". However, as the learning sample increases, the expectation of classification as "undefined" decreases. In any case, a viable plan is presented in table 5 for eventual stock complete elimination. Zero inventory expectation, including for new entries, should occur in five years following the planning. Less, if a dedicated task force is created for this specific point.
- **Achievable** (or agreed) - The same table 5 demonstrates and predicts the point at which the desired result can be reasonably achieved. In the fifth year, for the current workforce scenario, there will finally be a balance between INPUT and OUTPUT. As expected, all FNDE directors signed an official declaration of agreement with the project utilizing a circular letter, registered in the FNDE's electronic information system (SEI).
- **Realistic** (or relevant) – The process was relevant because of its economic impact and procedural efficiency. In the current results, we proved real savings up to R$1,3 billion[81] from the very beginning, with 28

[81] € 206,4 million, Euro value at 09/09/2021.

years of efficiency when compared to the current methodology. Once the paradigm shift is achieved, the methodology becomes a one-way process.

- **Time-bound** (or timely) – All the supervisory entities that monitor the activities conducted by FNDE, especially CGU, which has become a technological partner, are longing for the project release and its authorizing act. Other agencies will be able to benchmark our model, which can become a national reference.

As can be seen, even if not immediately to all items, this work fulfills its intended purpose and exceeds all initial expectations. More than that, it demonstrates scalability and collateral utilities for data governance and monitoring, and audit planning.

7.4 SCALABILITY

Despite the results and the project economic potential for the near future, a very important collateral asset consists in scalability. FNDE does not represent an isolated reality regarding public policy accountability in Brazil. However, as it has one of the largest single annual budgets, it is certainly the country's largest accountability management in the number of cases, being one of the entities with the most public judgments by the Federal Court of Accounts, TCU, therefore, one of the most audited entities.

This proposal, if satisfactorily enforced, will inaugurate a paradigm-shifting in social policy accountability management for the country, as the Court will possibly determine other agencies, in a similar situation to the FNDE, to adopt solutions such as this.

During the seminars and activities at SIBE, while working on this paper, we studied the success factors of international business development and globalization. However, long before we explored the rollout possibilities for the project, we must be aware that the FNDE does not represent a corporation but, as presented in the introduction, the State, especially for government support to education development in Brazil.

Thus, we must think about how to adjust this public agency to the concepts studied then, to understand its challenges for “internationalization” or even local expansion of its services, surely this will be the option applicable to the case.

Although FNDE is not a multinational (or even a multi-regional) corporation with several headquarters and pulverized operations, we should not deceive the reader, it is a multi-billion public agency in annual resource management. Thus, we could situate it as a regionalized or even an international expansion of its portfolio through risk management services derived from the expertise brought by our project.

Products and services for international markets, as well as the competition with multinational competitors in local markets, demand the observation of external and internal factors. This does not change for a public agency scopes that seek to expand its successful experiences to other national agencies.

It should be observed if FNDE culture and structure are suitable to participate in this process, or if it will need an extensive reengineering rational process and convincing of internal players for a new strategic objective to be achieved: the rollout to other agencies, ministries, or even a project internationalization.

FNDE will need to have a clear strategic plan and its employees, public servants, in this case, need to understand a new vision for the future. This new structure could consider the new playbook for improving its capabilities to be more efficient. The production chain's influence of its services, timing, understanding the desire target audience, internal culture, all must be rationally considered in the new scenario planning for determining new directions. This is where the analysis tools' importance comes in, foreseeing internal and external factors according to opportunities and ranking driving forces or those that anchor failure (figure 5).

Different business types of organizations internationalization are available to research as subject to international business, nevertheless, we shall always consider FNDE as a unique case, subject to its own local, regional, or maybe international potentialities and limitations. It means that for each approach we list, we must fit a reality suitable for a public agency that seeks to export its services to other institutions, whether this reality is regional or even international to government agencies interested in the benchmark.

Tarun Kanti Bose, in her Journal of Small Business Strategy,[82] has provided a recent literature review. She introduces her paper by “The Stage Based Approach”, by defining it as a linear and sequential process that constitutes a group of unique stages. In short, it consists of approaching the rollout as a gradual process of learning through experiences acquired in "other" markets (in our case, in other public agencies or even government sectors).

For the theory described, and we believe this to be the case even in a governmental environment, one should consider the acquisition process as formal interaction with other governmental agencies (in our case services or technologies applied to the Agency's state purpose), integration, and use of both: knowledge and expertise in foreign operations (regional, international, etc.) with incremental participation in foreign markets. For our case, the "foreign" market is any activity outside the FNDE that consumes its services or its pioneering expertise in Risk Management.

Following Bose’s quote, “Network theory” of internationalization places importance on the intra and inter-organizational networks for a successful internationalization process. This progression consists of relationships with other institutions or governments, adapting to this present case. It means that an externalizing activities pattern during the rollout process will depend heavily on establishing a network.

As for “The concept of international entrepreneurship”, presented in the article, we have an approach focused on rapid "scalability". As she describes it, this concept is typically applicable to international operations developed by start-ups. In this aspect, we ought to recognize FNDE already has a half-century of public service experience and can hardly be fitted into a start-up, even if it transcends concepts. It does not, in any case, prevent us from considering it at an “export” entrepreneurial juncture for the project in question: risk management.

Having considered this theoretical review, we will delve into the important aspects, according to the journal, for the FNDE “internationalization”, adapting concepts. Bose[83] lists ten. We will address each one by making appropriate adjustments to the Accountability Risk Management project and FNDE as a government agency.

---

[82] BOSE, Tarun, Critical success factors of SME internationalization, v. 26, p. 87–109, 2016.

[83] *Ibid*.

7.4.1 NETWORK, ALLIANCE, CLUSTERS

Business linkages such as networks, joint ventures, and subsidiaries play an important role in increasing the export probability[84]. Networking capability enables the identification and market opportunities exploitation, which facilitates the knowledge-intensive products development, and firm international market performance[85].

Transposing to the public sector, the relationship will be no different. FNDE relies on developing strategic relationships with other agencies and even other governments, both regional and international, to export its new risk management approach. It is therefore essential to identify agencies that share similar needs to the accountability challenges, as these will be the candidates for exporting services, solutions, and eventual cooperation agreements.

7.4.2 CAPACITY BUILDING

Financing strategies and commensurate finance management capabilities play dominant roles in the sustainable success of international business particularly for small and medium enterprises[86]. As a state-budget execution giant, FNDE excels in funding public policies, and over the past few years has excelled in the public sector in innovative solutions.

7.4.3 POLICY

Small and medium-sized enterprises from developed economies faced issues such as differences between regions, States, general business risks, and logistics, which can be challenging in a country of continental dimensions such as Brazil. It is to be expected that stability in public policies, or at least that these

---

[84] GUMEDE, Vusi; RASMUSSEN, VKamilla, Small manufacturing enterprises and exporting in South Africa: A preliminary assessment of key export success factors, **Journal of Small Business and Enterprise Development**, v. 9, p. 162–171, 2002.

[85] WEERAWARDENA, Jay; SULLIVAN MORT, Gillian, Investigating Social Entrepreneurship: A Multidimensional Model, **Journal of World Business**, v. 41, p. 21–35, 2006.

[86] SASI, Viveca; DARLING, John; GABRIELSSON, Mika, Finance Strategies of Rapidly-Growing Finnish Smes: Born Internationals and Born Globals, **European Business Review**, v. 16, 2004.

were fairly defined ones, could act as a catalyzing factor for the outreach of initiatives.

7.4.4 INNOVATIONS

There is an empirical relationship between organizational learning and organizational performance. In practice, this means that organizations reaching higher levels of organizational learning probably achieve higher performance[87].

This is an inescapable truth, whatever the business segment, state-owned or not, innovativeness, value creation for the entity is crucial to its success in the environment. Our project fulfills this expectation, by creating value through administrative rationality that represents billions of Reais in savings to the public coffers through a predictive selection of accounts to be analyzed and waiver of human analysis according to low risk detected.

7.4.5 BENEFITS AND BARRIERS

Detecting external and internal barriers can foresee serious impediments to both the rollout and the international operations success. In this sense, the strategic map should be prepared using planning tools that can foresee instruments and realities that can be catalysts or impediments to success, such as SWOT analysis, as we present in this paper.

7.4.6 FUTURE PROSPECTS

We understand that a successful externalization requires the existence of entrepreneurial strategies, such as opportunity creation, innovative products/services closely targeted to the beneficiaries, oriented growth, communication capacity, standardization, etc. Some requirements are complex for a governmental agency, due to the dependence on annual legal and budgetary forecasting, but others, especially those dependent only on planning and discretion, are fully feasible for our project's reality.

---

[87] MICHNA, Anna, The relationship between organizational learning and SME performance in Poland, **Journal of European Industrial Training**, v. 33, p. 356–370, 2009.

7.4.7 TIMING OF INTERNATIONALIZATION

Sometimes re-internationalization and de-internationalization are essential as entry and exit should not be universal, rather should be based on situations and facts[88].

Considering the state entity reality, the assertion is flawless. A government agency respects governmental agendas according to periods, taxpayer concerns, electoral seasonality, politics. Analyzing the current situation, FNDE is going through a restructuring moment of its accountability with broad support from the main Controlling Agencies. And in fact, an innovative solution seems to have been encouraged by economic results, both politically and financially.

It means a favorable situation and fact for expanding or exporting its activities through bilateral entity-entity or entity-government agreements.

7.4.8 MODES OF INTERNATIONALIZATION

Acedo and Jones [89] studied the rate of internationalization and focused on four managers' aspects in international operations. Those are risk perception, proactivity, tolerance for ambiguity, and international orientation. Transporting this to our state situation, the aspects remain, even if adaptations are respected. As with any ambitious project, risks must be mapped and listed in terms of tolerance, acceptance, and remediation, including by establishing typography and impact gradation. The agency also needs to be proactive in managing them, including identifying any ambiguities that duplicate efforts. There is no time to be lost in respecting the process timing and the synergies that converge to its success.

---

[88] FREEMAN, Susan; DELIGONUL, Seyda; CAVUSGIL, S., Strategic re-structuring by born-globals using outward and inward-oriented activity, **International Marketing Review**, v. 30, 2013.

[89] ACEDO, Francisco; JONES, Marian, Speed of Internationalization and Entrepreneurial Cognition: Insights and a Comparison between International New Ventures, Exporters and Domestic Firms, **Journal of World Business**, v. 42, p. 236–252, 2007.

7.4.9 DESTINATIONS OF INTERNATIONALIZATION

We recognize that appropriate targeting is pivotal to the eventual success of any service exporting to other agencies or governments, local or beyond. Managing cultural distances and providing support and response to the beneficiary of the innovative tool to be transferred to the new government entities seems to be the key to success. This aspect highlights the importance of the FNDE social capital and its network of relationships with actors and stakeholders.

7.4.10 OPERATIONAL DECISIONS OF INTERNATIONALIZATION

Our research reveals that a kind of entrepreneurial culture exists, particularly when one wishes to operate abroad. It is a behavior or institutional competence that seems to allow corporate entities to gain significant operational, strategic, and competitive advantages.

We realize that internal culture is a decisive factor of leverage, at least synergy in the rollout process. Ibeh[90] identified several factors that lead to positive international corporate performance, including decision-makers' prior experience, outward contacts, and orientation, specific competencies related to planning orientation, innovation, research, and more.

There is also a deep relationship between risk, operating characteristics. Notice that this is not distinguished between the public and private sector, as long as we consider the public as an exporter of services or solutions. For the present case, we can finally evidence the rollout after the perception of the above aspects regarding project scalability. In this case, it is already a reality even before the appropriate normalization by the FNDE Deliberative Council.

As mentioned, the studies and work done to perform the accountability BI became more than an information resource, but rather a planning tool for management control actions. With the signing of a technical cooperation agreement between FNDE and CGU for sharing data and technology[91], this

---

[90] IBEH, Kevin, On the internal drivers of export performance among Nigerian firms: Empirical findings and implications, **Management Decision - MANAGE DECISION**, v. 41, p. 217–225, 2003.

[91] UNIÃO (CGU); EDUCAÇÃO (FNDE), Acordo de Cooperação Técnica n. 19. de 18 de maio de 2020.

project fulfills its role in being scalable to other agencies or government offices. Hopefully, it is just the first step in a long journey.

8.CONCLUSIONS AND NEXT STEPS

This paper consists in its objective to propose an appropriate risk management methodology to enable conclusive accountability analysis of rendered accounts, considering programs listed, to balance itself with public agency operational capacity, avoiding new accounts pending conclusive analysis to increase actual stock.

Furthermore, the project hypothesis states that the proposed risk management, in addition to the efficiency of the agency‘s accountability process, will provide a viable solution to the current stock of accounts that have already been rendered and not analyzed. As stated above, the assumed hypothesis holds true.

As we assumed, the appropriate indicator should measure project success: growth rate for conclusive analysis inventory. The target rate goal for generating new inventories for PNAE, PNATE, and PDDE, should be ideally 0%. With the project concluded and the results evidenced utilizing the program business intelligence that manages the accounts rendered, SIGPC, we presented an action plan to demonstrate the expected decrease in the inventory of accounts until the operational balance is reached.

As assumed, the algorithm predictably concludes by approving some accountability obligations and rating others to traditional analysis, respecting the operational capacity established for the process. More than that, the adoption of the methodology brings to the FNDE technological forefront the use of evidence-based data in its governance process.

The business intelligence tool, among other results, provides a general and geo-referenced view from the rendering of accounts inserted in SIGPC. In this aspect, the results exceed the project's intended goals. The tool shows financial values, stock chronological evolution, stock formation dynamics, breaking point between current analysis capacity and the formation of new inventories, accounts rendered distribution per loss identification, the relationship between transferred value and the number of annual accounts rendered, and

tangible representation of the FNDE "Fine mesh" applied to the three programs listed in the pilot project.

Being built on MicroStrategy[92], The tool enables countless information crossings, simply by selecting metrics and attributes according to the user's objectives. It is at this point, especially, that we exceeded our goals since the methodology used starts to demonstrate scalability to be exported to other agencies or even signed a partnership with government authorities of public management control, through agreements for sharing data and technology. Such scalability has already been evidenced by an agreement between the CGU Minister, Education Minister, and FNDE Presidency[93].

It should be noted that data analyzed shows PNAE managing to boost the number of its accountability obligations, in remote past, by subdividing the program into other specialties that would generate identical individual accountability obligations to be analyzed. However, the accountability obligations chronological curve from these programs shows a return to linearity, evidencing that the regrouping of its "derived" programs into a single school feeding policy could be a solution to be emulated.

From the SIGPC data lake for PDDE, it is evident that the agency has the accountability largest stock in this single policy. In this case, clearly associated with the public policy subdivision into several "son" actions, all of which creating individual accountability obligations for the FNDE.

It is concluded, therefore, that the highest manager should consider the successful PNAE experience aggregating "sub-programs" from PDDE into a single "father" program, with a broader and more objective services menu (aggregating all PDDE actions objectively into a single accountability obligation), reducing by up to 1/3 the total increase in accountability liabilities under FNDE responsibility. This measure depends on a regulation drawn up and published by the FNDE Deliberative Council.

---

[92] **Discover MicroStrategy 2021**.

[93] UNIÃO (CGU); EDUCAÇÃO (FNDE), Acordo de Cooperação Técnica n. 19. de 18 de maio de 2020.

Finally, we conclude this process preserves an administrative rationality relation, and criteria reasonableness, respecting the due process of law applied to the accountability Constitutional obligation. Therefore, it is efficient and, by making it both conventional and public, it will promote administrative morality by creating a vigilant attitude toward the party responsible for the management of federal public resources.

The model, according to its actual results, has a promising impact, therefore, as a paradigm-shifting in Brazilian public administration, it would be responsible to carry out new validations as next steps or even tests of other methodologies or algorithms in order to evolve it or to establish appropriate risk appetites for the government agency.

In short, it is up to FNDE, after regulating the resolution that authorizes risk management by prediction, to establish what its "risk appetite" is. In other words, from **what value the agency will not submit itself to predictive approval**, necessarily performing traditional judgment on the beneficiary's accounts of the public policy financed with FNDE resources, regardless of the predictive result proposed by the "Fine Mesh".

Other statistical criteria can be combined with materiality, such as high-risk classification frequency, the sum of programs classified as high-risk in defined time intervals, etc.

The idea is that whenever the risk exceeds the established appetite, FNDE must necessarily judge the account rendered, independent of the risk range. Likewise, the independent sector responsible for auditing the public agency may propose an audit track, listing attributes that, whenever consolidated, should determine the accounts judgment, regardless of the risk range.

A stage that has not been reached, but is independent of this research, consists of "Resolution" from FNDEDdeliberative Council. As the project's next step, it is expected to elaborate this normative, to legally allow the administrative process. We consider this political-administrative stage to be overdue.

Because of the SARS-COVID19 pandemic, priorities and policies in Brazil underwent major changes.

FNDE, as a government agency and with great political participation, especially in a fiscal year that forced the education digitizing[94] without adequate planning, was not immune to such political oscillations.

Several authorities, among them ministers and Agency presidents, all deliberative Council members, were switched in this period, severely impacting the legal and formal launching of this already developed measure. Even management control entities' priorities, were redirected to the fraud prediction associated with the pandemic care.

Nevertheless, with evolutions in vaccines and advances in immunization, the necessary formalities for the formal launching enterprise are imminent, since the cooperation agreement with CGU already benefits from the results from this project. Moreover, thanks to the results achieved, FNDE has created a Nucleus of Quantitative Methods, directly linked to the presidency and ensuring projects continuity[95].

As a further next step, it is recommended that the entity anticipate the control organs' impositions and endeavor to expand this initiative to all FNDE accounts and applicable systems, by adopting the so-called "FNDE Fine Mesh", already on 2022 fiscal year, for all accounts under its responsibility.

---

[94] Digitization refers to the business process of going digital, which requires deeper changes, from the business model to the value stream. To this end, the organization uses new processes, systems, tools, and means of collaboration.

[95] **Novo método moderniza análise de prestação de contas no FNDE - Portal do FNDE**, available at: <http://www.fnde.gov.br/index.php/acesso-a-informacao/institucional/area-de-imprensa/noticias/item/12815-novo-m%C3%A9todo-moderniza-an%C3%A1lise-de-presta%C3%A7%C3%A3o-de-contas-no-fnde?highlight=WyJlc2NvbGEiXQ==>. Accessed in: 3 set. 2021.

9. REFERENCES

ACEDO, Francisco; JONES, Marian. Speed of Internationalization and Entrepreneurial Cognition: Insights and a Comparison between International New Ventures, Exporters and Domestic Firms. Journal of World Business, v. 42, p. 236–252, 2007.

AGNIESZKA WIŚNIEWSKA. The Mission & Vision Statements in the Consumers' Strategies. Acta Universitatis Nicolai Copernici. Zarządzanie, v. 11, n. 3, p. 111–118, 2018. Available at: <http://search.ebscohost.com/login.aspx?direct=true&db=edsbaz&AN=edsbaz.171542092&site=eds-live>.

ALVES, Tiago Venâncio. SIGPC - Acesso Público - Portal do FNDE. Available at: <http://www.fnde.gov.br/fnde_sistemas/sigpc-acesso-publico>. Accessed on: 18 sep. 2019.

ARAUJO, Vladimir B. Technical Survey. 2017. Available at: <https://www.fnde.gov.br/sei/controlador.php?acao=procedimento_trabalhar&acao_origem=protocolo_pesquisa_rapida&id_protocolo=845651&infra_sistema=100000100&infra_unidade_atual=110000077&infra_hash=8672d85846942c24e15377f6cf7d7e52024c37dfdb9fcb9416a2a8110e8dc5e6>. Accessed on: 20 ago. 2021.

ARAUJO, Vladimir B.; ANANIAS, Mateus C.; MACIEL, Andre L.A. Techinical Note 04.2018-DIFIN/FNDE. 2018.

ARNOLD, Roger A. Economics. 9e [ed.]. Australia ; Mason, OH: South-Western Cengage Learning, 2010.

AZEVEDO, Ana; SANTOS, Manuel. KDD, semma and CRISP-DM: A parallel overview. In: [s.l.: s.n.], 2008, p. 182–185.

BOSE, Tarun. Critical success factors of SME internationalization. v. 26, p. 87–109, 2016.

CHAPMAN, Pete; CLINTON, Julian; KERBER, Randy; et al. Step-by-step data mining guide. p. 76, .

CHEN, Tianqi; GUESTRIN, Carlos. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16, p. 785–794, 2016. Available at: <http://arxiv.org/abs/1603.02754>. Accessed on: 18 sep. 2019.

COOK, Nancy R. Use and Misuse of the Receiver Operating Characteristic Curve in Risk Prediction. Circulation, v. 115, n. 7, p. 928–935, 2007. Available at: <https://www.ahajournals.org/doi/full/10.1161/CIRCULATIONAHA.106.672402>. Accessed on: 12 sep. 2021.

EDUCAÇÃO, Assessoria de Comunicação Social do FNDE com informações do Ministério da. Mapa Estratégico - Portal do FNDE. Available at: <http://www.fnde.gov.br/index.php/acoes/fnde-estrategico/plano-estrategico/mapa-estrategico>. Accessed on: 18 ago. 2021.

EDUCAÇÃO, Assessoria de Comunicação Social do FNDE com informações do Ministério da. Resolução/CD/FNDE no 2, de 18 de janeiro de 2012 - Portal do FNDE. Available at: <http://www.fnde.gov.br/acesso-a-informacao/institucional/legislacao/item/3512-resolu%C3%A7%C3%A3o-cd-fnde-n%C2%BA-2-de-18-de-janeiro-de-2012>. Accessed on: 9 ago. 2021.

EDUCAÇÃO, Assessoria de Comunicação Social do FNDE com informações do Ministério da. Sobre o FNDE Estratégico - Portal do FNDE. Available at: <http://www.fnde.gov.br/index.php/acoes/fnde-estrategico/sobre/sobre>. Accessed on: 6 abr. 2020.

ELOGROUP, Consultoria. FNDE Accountability Process Cost Analysis Report. [s.l.]: Ministry of Education / Office of Management of Processes and Strategic Projects / National Fund for the Development of Education, 2017.

FAWCETT, Tom. An introduction to ROC analysis. Pattern Recognition Letters, v. 27, n. 8, p. 861–874, 2006. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S016786550500303X>. Accessed on: 23 ago. 2021.

FNDE, Assessoria de Comunicação Social do. Mapa Estratégico - Portal do FNDE. Available at: <http://www.fnde.gov.br/index.php/acoes/fnde-estrategico/plano-estrategico/mapa-estrategico>. Accessed on: 23 ouc. 2019.

FREEMAN, Susan; DELIGONUL, Seyda; CAVUSGIL, S. Strategic re-structuring by born-globals using outward and inward-oriented activity. International Marketing Review, v. 30, 2013.

GROLLMAN, Daniel; BILLARD, Aude. Learning Algorithms. In: SEEL, Norbert M. (Org.). Encyclopedia of the Sciences of Learning. Boston, MA: Springer

US, 2012, p. 1766–1769. Available at: <https://doi.org/10.1007/978-1-4419-1428-6_759>. Accessed on: 9 sep. 2021.

GUMEDE, Vusi; RASMUSSEN, VKamilla. Small manufacturing enterprises and exporting in South Africa: A preliminary assessment of key export success factors. Journal of Small Business and Enterprise Development, v. 9, p. 162–171, 2002.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J.H. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. [s.l.]: Springer, 2009. (Springer series in statistics). Available at: <https://books.google.com.br/books?id=eBSgoAEACAAJ>.

HUBER-, Angelika; SCHÜLLER, Dr. Marcus; MÜLLER, Nils; et al. Rethinking the value chain. KPMG | Rethinking the value chain, p. 82, 2018.

IBEH, Kevin. On the internal drivers of export performance among Nigerian firms: Empirical findings and implications. Management Decision - MANAGE DECISION, v. 41, p. 217–225, 2003.

KAPLAN, ROBERT S; NORTON, DAVID P. Organização orientada para a estratégia : como as empresas que adotam o balanced scorecard prosperam no novo ambiente de negócios / Robert S. Kaplan , David P. Norton ; [tradução Afonso Celso da Cunha Serra]. 17. ed. Rio de Janeiro : Elsevier : Campus, 2001.: [s.n., s.d.].

MARTINS, Eliseu. Contabilidade de Custos. 10. ed. [s.l.]: Atlas, 2010.

MCLEAN, R. Robert S. Kaplan and David P. Norton-Alignment: Using the Balanced Scorecard to Create Corporate Synergies. AUSTRALIAN JOURNAL OF MANAGEMENT, n. 2, p. 367, 2006. Available at: <http://search.ebscohost.com/login.aspx?direct=true&db=edsbl&AN=RN203779446&site=eds-live>.

MICHNA, Anna. The relationship between organizational learning and SME performance in Poland. Journal of European Industrial Training, v. 33, p. 356–370, 2009.

NACIONAL, Imprensa. INSTRUÇÃO NORMATIVA CONJUNTA No 1, DE 10 DE MAIO DE 2016 - Imprensa Nacional. Available at: <https://www.in.gov.br/materia>. Accessed on: 20 ago. 2021.

NAGEL, Kurt; FAIX, Werner G.; DJALALI, Ardin; et al (Orgs.). General management tools. 3. Aufl. Stuttgart: Steinbeis-Edition, 2013.

NEVES, Clara V.B.; DE JESUS, Victor L.F.; CONCEIÇÃO, Shirley G. Technical Note No. 731/2017 / DIV1 / CGEDUB / DS / SFC. 2017.

PEREIRA, Fernando S.; MENKE, Wagner B.; ALMEIDA, Aldo S.; et al. Technical Note n° 1357/2018/CGEDUB/DS/SFC. 2018.

RASCHKA, Sebastian. How do I evaluate a model? Dr. Sebastian Raschka. Available at: <https://sebastianraschka.com/faq/docs/evaluate-a-model.html>. Accessed on: 3 mar. 2020.

RÊGO, Bergson Lopes. GESTAO E GOVERNANCA DE DADOS : promovendo dados como ativo de valor nas empresas. Brazil: BRASPORT; Edition: 1, 2013. Available at: <https://www.amazon.com.br/Gest%C3%A3o-Governan%C3%A7a-Dados-Promovendo-empresas-ebook/dp/B00ZQ0EDES>.

SASI, Viveca; DARLING, John; GABRIELSSON, Mika. Finance Strategies of Rapidly-Growing Finnish Smes: Born Internationals and Born Globals. European Business Review, v. 16, 2004.

UNIÃO (CGU), Brasil Controladoria-Geral da; EDUCAÇÃO (FNDE), Brasil Fundo Nacional de Desenvolvimento da. Acordo de Cooperação Técnica n. 19. de 18 de maio de 2020. Diário Oficial da União n. 95, de 20 de maio de 2020, seção 3, p. 93, 2020. Available at: <https://repositorio.cgu.gov.br/handle/1/45258>. Accessed on: 24 ago. 2021.

WEERAWARDENA, Jay; SULLIVAN MORT, Gillian. Investigating Social Entrepreneurship: A Multidimensional Model. Journal of World Business, v. 41, p. 21–35, 2006.

- Mandado de segurança. Tribunal de Contas da União. Supremo Tribunal Federal STF - MANDADO DE SEGURANÇA : MS 21644 DF. Jusbrasil. Available at: <https://stf.jusbrasil.com.br/jurisprudencia/14706970/mandado-de-seguranca-ms-21644-df>. Accessed on: 10 ago. 2021.

AC 7790/2015 - PRIMEIRA CAMERA| Tribunal de Contas da União. Available at: <https://pesquisa.apps.tcu.gov.br/#/documento/acordao-completo/7790/ANOACORDAO%253A%25222015%2522/DTRELEVANCIA%2520desc%252C%2520NUMACORDAOINT%2520desc/0/%2520>. Accessed on: 30 ago. 2021.

Arnold - 2010 - Economics.pdf.

Chapman et al. - Step-by-step data mining guide.pdf. Available at: <https://the-modeling-agency.com/crisp-dm.pdf>. Accessed on: 20 ago. 2021.

Constituicao-Compilado. Available at: <http://www.planalto.gov.br/ccivil_03/constituicao/constituicaocompilado.htm>. Accessed on: 12 nov. 2019.

DEL 200. Available at: <http://www.planalto.gov.br/ccivil_03/decreto-lei/del0200.htm>. Accessed on: 10 ago. 2021.

Del0872. Available at: <http://www.planalto.gov.br/ccivil_03/Decreto-Lei/1965-1988/Del0872.htm>. Accessed on: 18 sep. 2019.

Discover MicroStrategy 2021. MicroStrategy. Available at: <https://www.microstrategy.com/en/get-started/2021>. Accessed on: 23 ago. 2021.

Gestão de Pessoas - Portal do FNDE. Available at: <https://www.fnde.gov.br/index.php/gestao-de-pessoas>. Accessed on: 9 ago. 2021.

L5537. Available at: <http://www.planalto.gov.br/ccivil_03/LEIS/L5537.htm>. Accessed on: 18 sep. 2019.

Lcp101. Available at: <http://www.planalto.gov.br/ccivil_03/leis/lcp/lcp101.htm>. Accessed on: 6 ago. 2021.

Li - A Gentle Introduction to Gradient Boosting.pdf. Disponível em: <http://www.chengli.io/tutorials/gradient_boosting.pdf>. Available at: 20 ago. 2021.

Niakšu - CRISP Data Mining Methodology Extension for Medica.pdf. Available at: <https://www.bjmc.lu.lv/fileadmin/user_upload/lu_portal/projekti/bjmc/Contents/3_2_2_Niaksu.pdf>. Accessed on: 3 mar. 2020.

Novo método moderniza análise de prestação de contas no FNDE - Portal do FNDE. Available at: <http://www.fnde.gov.br/index.php/acesso-a-informacao/institucional/area-de-imprensa/noticias/item/12815-novo-m%C3%A9todo-moderniza-an%C3%A1lise-de-presta%C3%A7%C3%A3o-de-contas-no-fnde?highlight=WyJlc2NvbGEiXQ==>. Accessed on: 3 set. 2021.

Pesquisa textual | Tribunal de Contas da União. Available at: <https://pesquisa.apps.tcu.gov.br/#/documento/ata-sessao/*/DATASESSAOORDENACAO%253A%255B20160921%2520to%252020160921%255D%2520COLEGIADO%253A%2528Plen%25C3%

25A1rio%2529/DTRELEVANCIA%2520desc/1/%2520>. Accessed on: 9 ago. 2021.

Programas. Fundo Nacional de Desenvolvimento da Educação. Available at: <https://www.gov.br/fnde/pt-br/acesso-a-informacao/acoes-e-programas/programas>. Accessed on: 12 sep. 2021.

Quem é quem - Portal do FNDE. Available at: <https://www.fnde.gov.br/index.php/acesso-a-informacao/institucional/quem-e-quem>. Accessed on: 27 abr. 2020.

R: The R Project for Statistical Computing. Available at: <https://www.r-project.org/>. Accessed on: 3 mar. 2020.

Relatório de Gestão - Estrutura. Available at: <https://sites.google.com/view/relatriodegestao/apresenta%C3%A7%C3%A3o/estrutura>. Accessed on: 12 sep. 2021.

SEI / FNDE. Available at: <https://www.fnde.gov.br/sip/login.php?sigla_orgao_sistema=FNDE&sigla_sistema=SEI>. Accessed on: 30 ago. 2021.

Setting SMART Objectives. p. 5, .

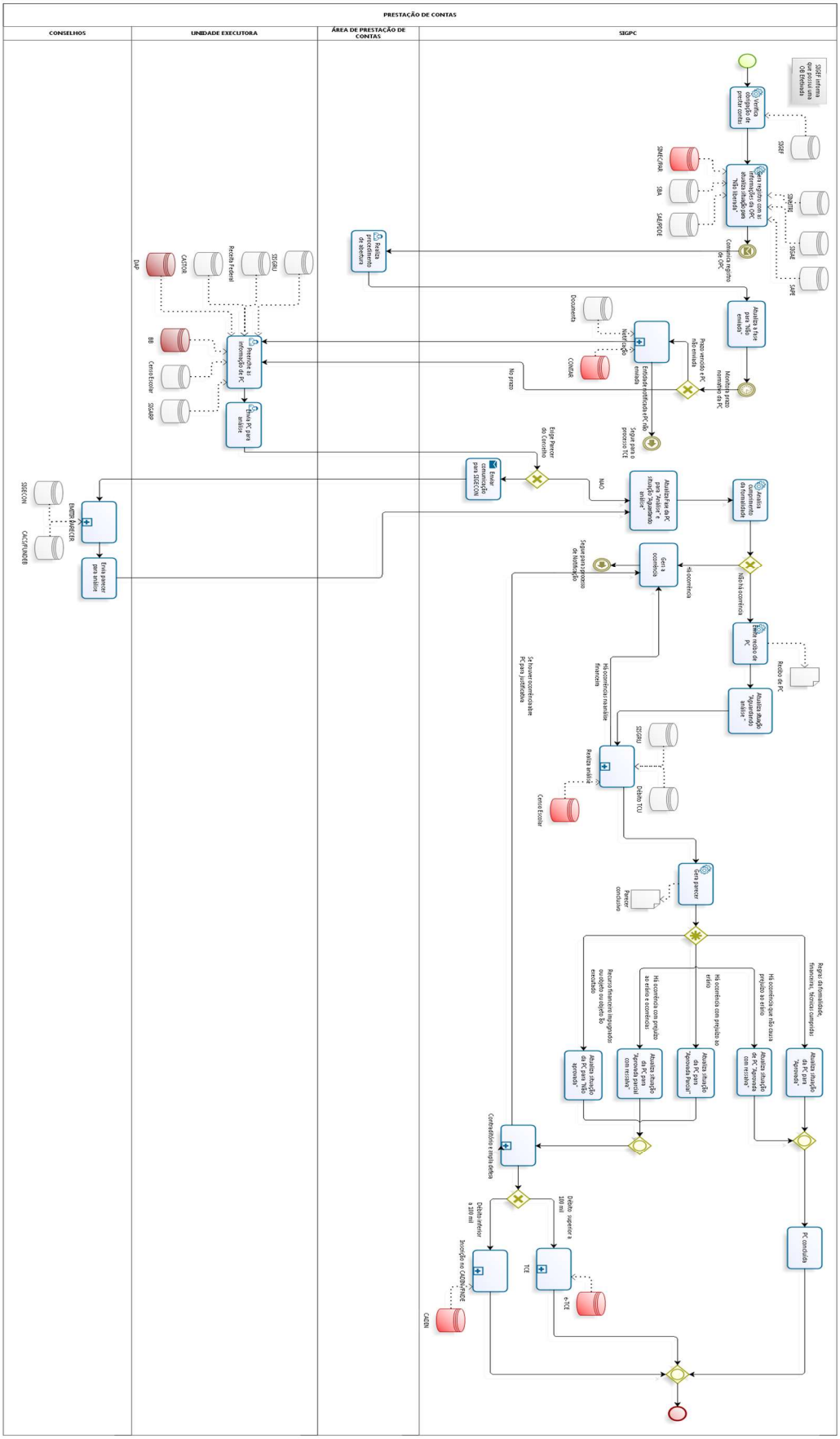
S.M.A.R.T-Way-Management-Review.pdf. Available at: <https://community.mis.temple.edu/mis0855002fall2015/files/2015/10/S.M.A.R.T-Way-Management-Review.pdf>. Accessed on: 10 sep. 2021.

TC 027.076/2016-4 | Tribunal de Contas da União. Available at: <https://pesquisa.apps.tcu.gov.br/#/documento/processo/027.076%252F2016-4%2520/COPIAANO%253A%25222016%2522/DTAUTUACAOORDENACAO%2520desc%252C%2520NUMEROCOMZEROS%2520desc/1/%2520>. Accessed on: 30 ago. 2021.

Vision and Mission Statements -- a Roadmap of Where You Want to Go and How to Get There | Ag Decision Maker. Available at: <https://www.extension.iastate.edu/agdm/wholefarm/html/c5-09.html>. Accessed on: 6 apr. 2020.

10. APPENDIX

ANNEX I - ACCOUNTABILITY FLOW IN SIGPC. Source: DIRTI/FNDE

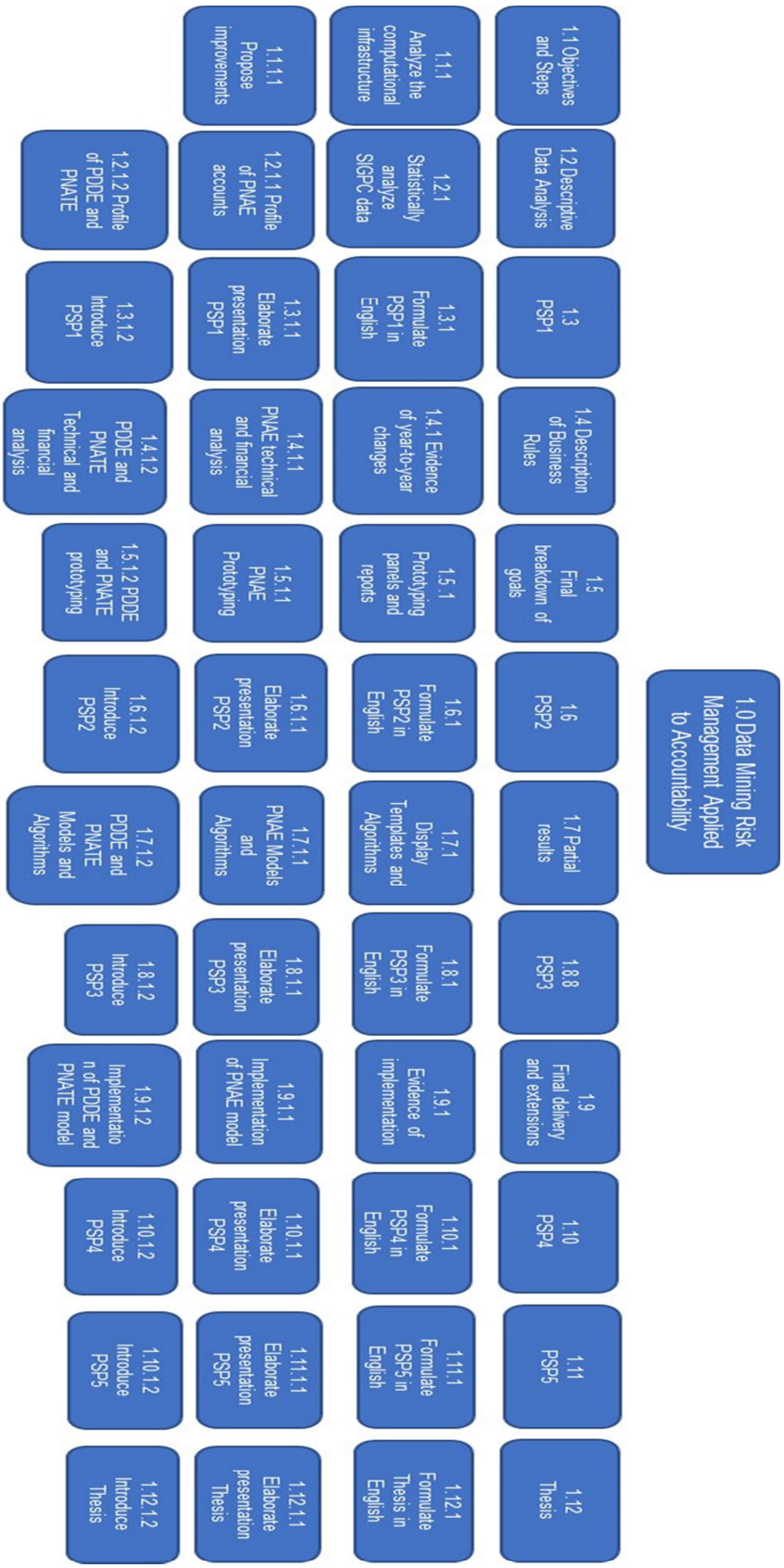


ANNEX II – PROJECT ATTRIBUTES DESCRIPTION

Attribute	Description
DANO	Response variable, with binary response: 0 - no loss to the treasury and 1 - with loss to the treasury.
SG_UF	Federation Unit
PROG	Analyzed program being: PNAE , PNATE and PDDE
AN_EXERCICIO	Fiscal year of the analyzed program
NO_TP_ENTIDADE	Type of entity
A_PERC_GINDV	Relationship between improper and proper expenses obtained through the bank statement
A_SALDO_TOTAL	Relationship between remaining balance from previous year + transfer and expense due obtained through the bank statement
A_GASTO_TRANS	Relationship between the amount transferred and the amount spent in the year obtained from the bank statement
A_SALDO_TRANSFERENCIA	Relationship between the amount transferred in the current year and the remaining balance of the previous year obtained through the bank statement
VL_SALDO_CONTA_SF	Amount remaining in the entity's current account at the end of the fiscal year obtained through the bank statement
VL_SALDO_FUNDOS_SF	Amount remaining in the fund account of the entity at the end of the fiscal year obtained through the bank statement
VL_SALDO_POUPANCA_SF	Amount remaining in the entity's savings account at the end of the fiscal year obtained through the bank statement
VL_SALDO_RDB_SF	Amount remaining in the entity's Bank Deposit Receipts type application account at the end of the fiscal year obtained through the bank statement
C_BLOQUEIO	Amounts moved to credit in the "BLOCK" condition obtained through the bank statement
C_OUTRAS_RECEITAS	Values moved to credit in the condition of "OTHER REVENUES" obtained through the bank statement
C_RESGATE_DE_APLICAO	Values moved to credit in the condition of "APPLICATION

	RESCUE" obtained through the bank statement
C_TRANSFERENCIA_FNDE	Values moved to credit in the condition of transfer made by the FNDE obtained through the bank statement
D_APLICAO	Values moved to debit in the condition of transfer to application account obtained through the bank statement

ANNEX III – WORK BREAKDOWN STRUCTURE



ANNEX IV - 2019 PROJECT SCHEDULE

ACTIONS	DELIVERY	MANAGEMENT SCHEDULE - 2019											
		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Analyze the computational infrastructure	Proposed improvements												
Analyze SIGPC data	Profile of PNAE accounts												
Statistical analysis of SIGPC data	Profile of PDDE e PNATE accounts												
Formulate PSP1	Formulate PSP1 in English												
Formulate PSP1	Elaborate presentation PSP1												
Formulate PSP1	Introduce PSP1												
Evidence of year-by-year changes	PNAE technical & financial analysis												
Evidence of year-by-year changes	PDDE and PNATE technical & financial analysis												
Formulate PSP2	Formulate PSP2 in English												
Formulate PSP2	Elaborate presentation PSP2												
Formulate PSP2	Introducing PSP2												
Formulate PSP3	Formulate PSP3 in English												
Formulate PSP3	Elaborate presentation PSP3												
Formulate PSP3	Introducing PSP3												

Prototyping panels and reports	PNAE Prototyping												
Prototyping panels and reports	PDDE and PNATE Prototyping												
Present Models and Algorithms	PNAE Models and Algorithms												
Present Models and Algorithms	PDDE and PNATE Models and Algorithms												
Demonstrate implementation	Implementation of the PNAE model												
Demonstrate implementation	Implementation of PDDE and PNATE models												

TABLE 6. ACTION STRUCTURING MATRIX: 2019 SCHEDULE

ANNEX V - 2020 PROJECT SCHEDULE

ACTIONS	DELIVERY	MANAGEMENT SCHEDULE - 2020											
		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Demonstrate implementation	Implementation of the PNAE model												
Demonstrate implementation	Implementation of PDDE and PNATE models												
Formulate PSP4	Formulate PSP4 in English												
Formulate PSP4	Elaborate presentation PSP4												
Formulate PSP4	Introducing PSP4												
Formulate PSP5	Formulate PSP4 in English												
Formulate PSP5	Elaborate presentation PSP5												
Formulate PSP5	Introducing PSP5												
Formulate Thesis	Formulate Thesis in English												
Formulate Thesis	Elaborate Thesis presentation												
Formulate Thesis	Defend Thesis												

TABLE 7. ACTION STRUCTURING MATRIX: 2020 SCHEDULE

ANNEX VI - PROJECT COSTS

Objectives and Steps TOTAL OF THE STAGE: R$ 20.000,00		INPUT	VALUE	AMOUNT	TOTAL
Analyze the computational infrastructure	Propose improvements	Product A - Consulting	R$ 10.000,00	2	R$ 20.000,00
Descriptive Data Analysis TOTAL OF THE STAGE: R$ 30.000,00		**INPUT**	**VALUE**	**AMOUNT**	**TOTAL**
Statistically analyze SIGPC data	Profile of PNAE accounts	Product B - Consulting	R$ 15.000,00	1	R$ 15.000,00
	Profile of PDDE and PNATE accounts	Product B - Consulting	R$ 15.000,00	1	R$ 15.000,00
Description of Business Rules TOTAL OF THE STAGE: R$ 30.000,00		**INPUT**	**VALUE**	**AMOUNT**	**TOTAL**
PSP 1	Formulate PSP1 in English	No cost registered			
	Elaborate presentation PSP1	No cost registered			
	Introduce PSP1	No cost registered.			
Evidence of year-to-year changes	PNAE technical and financial analysis	Product C - Consulting	R$ 15.000,00	1	R$ 15.000,00
	PDDE and PNATE technical and financial analysis	Product C - Consulting	R$ 15.000,00	1	R$ 15.000,00
The final breakdown of goals TOTAL OF THE STAGE: R$ 32.000,00		**INPUT**	**VALUE**	**AMOUNT**	**TOTAL**
Prototyping panels and reports	PNAE Prototyping	Product D - Consulting	R$ 16.000,00	1	R$ 16.000,00
	PDDE and PNATE Prototyping	Product D - Consulting	R$ 16.000,00	1	R$ 16.000,00
PSP 2	Formulate PSP2 in English	No cost registered			
	Elaborate presentation PSP2	No cost registered			
	Introduce PSP2	No cost registered			
Partial results TOTAL OF THE STAGE: R$ 40.000,00		**INPUT**	**VALUE**	**AMOUNT**	**TOTAL**

Display Templates and Algorithms	PNAE Models and Algorithms	Product E - Consulting	R$ 20.000,00	1	R$ 20.000,00
	PDDE and PNATE Models and Algorithms	Product E - Consulting	R$ 20.000,00	1	R$ 20.000,00
PSP 3	Formulate PSP3 in English	No cost registered			
	Elaborate presentation PSP3	No cost registered			
	Introduce PSP3	No cost registered			
Final delivery and extensions TOTAL OF THE STAGE: R$ 40.000,00		**INPUT**	**VALUE**	**AMOUNT**	**TOTAL**
Evidence of implementation	Implementation of PNAE model	Product F - Consulting	R$ 20.000,00	1	R$ 20.000,00
	Implementation of PDDE and PNATE model	Product F - Consulting	R$ 20.000,00	1	R$ 20.000,00
PSP4	Formulate PSP4 in English	No cost registered			
	Elaborate presentation PSP4	No cost registered			
	Introduce PSP4	No cost registered			
PSP5	Formulate PSP5 in English	No cost registered			
	Elaborate presentation PSP5	No cost registered			
	Introduce PSP5	No cost registered			
Thesis	Formulate Thesis in English	No cost registered			
	Elaborate Thesis presentation	No cost registered			
	Introduce Thesis	No cost registered			
TOTAL PROJECT			**R$ 192.000,00**		

TABLE 8. RESOURCES AND PROJECT BUDGET