Government Data Science Teams: A Framework for Implementing Strategic Monitoring Solutions

Hugo Augusto Vasconcelos Medeiros [a]*, *André Leite Wanderley* [b], *Diogo de Carvalho Bezerra* [c], *Carlos Alberto Gomes de Amorim Filho* [d], *Rafael Zimmerle da Nóbrega* [e], *Felipe Gustavo de Moraes Ferreira* [f].

[a] Secretaria Estadual de Planejamento de Pernambuco, Governo de Pernambuco, Recife, Pernambuco, Brasil, and Programa de Pós-graduação Profissional em Políticas Públicas, Universidade Federal de Pernambuco, Recife, Pernambuco, Brasil, hugoavmedeiros@gmail.com, 0000-0001-6249-0920.

[b] Secretaria Estadual de Projetos Estratégicos, Governo de Pernambuco, Recife, Pernambuco, Brasil, and Departamento de Estatística, Universidade Federal de Pernambuco, Recife, Pernambuco, Brasil, andre.leite@sepe.pe.gov.br, 0000-0002-4718-9766.

[c] Secretaria Estadual de Mobilidade e Infraestrutura, Governo de Pernambuco, Recife, Pernambuco, Brasil, and Núcleo de Gestão, Universidade Federal de Pernambuco, Caruaru, Pernambuco, Brasil, diogo.bezerra@semobi.pe.gov.br, 0000-0002-1216-8674.

[d] Secretaria Estadual de Projetos Estratégicos, Governo de Pernambuco, Recife, Pernambuco, Brasil, and Departamento de Economia, Universidade Federal de Pernambuco, Recife, Pernambuco, Brasil, carlos.agaf@gmail.com, 0000-0001-6315-8305.

[e] Secretaria Estadual de Projetos Estratégicos, Governo de Pernambuco, Recife, Pernambuco, Brasil, and Programa de Pós-graduação em Estatística, Universidade Federal de Pernambuco, Recife, Pernambuco, rafael.zimmerle@ufpe.br, 0009-0003-0800-9042.

[f] Secretaria Estadual de Projetos Estratégicos, Governo de Pernambuco, Recife, Pernambuco, Brasil, felipeferreira.govpe@gmail.com, 0009-0004-3086-849X.

Submitted: 31 January 2025, Revised: 26 March 2025, Accepted: 21 April 2025, Published: 19 May 2025
Policy paper, DOI: https://doi.org/10.59490/dgo.2025.925

Abstract. This paper presents a framework developed by Pernambuco's Strategic Monitoring Data Science Team to design and implement data-driven solutions for monitoring public policies. Using an action research methodology, the study integrates data science, iterative development, and stakeholder engagement. Despite being major producers and consumers of data, governments still face significant challenges in applying data science for policy monitoring, including data quality issues, legal constraints, and institutional silos. Existing frameworks such as CRISP-DM, Scrum, and Kanban are either too technical or primarily focused on software development rather than the policy-driven decision-making required in government settings. The Strategic Monitoring Team was established within Pernambuco's Secretariat of Strategic Projects, comprising a Chief Data Scientist as team leader, a Project Manager, three Data Scientists specializing in modeling, engineering, and visualization, and a Trainee. The team operates through an iterative five-step process: Diagnose, which involves meetings with stakeholders to identify policy issues; Plan, where internal discussions define solutions; Act, which includes the development of dashboards, reports, and applications; Evaluate, to review whether the solutions address policy needs; and learn, focusing on documenting findings and improving tools. To enhance their workflow, the team adapted Scrum methodology by incorporating policy research alongside software development, tracking projects via Notion, and deploying solutions using R and Shiny Proxy. The study highlights that traditional frameworks such as Scrum and CRISP-DM require adaptations to effectively integrate research aspects and government governance structures. By bridging data insights with decision-making processes, the team successfully balances software development, policy research, and institutional needs. The findings emphasize the necessity of specialized data science frameworks tailored for government applications, ensuring a structured yet flexible approach to strategic policy monitoring through data-driven solutions.

Keywords. Government Data Science, Strategic Policy Monitoring, Agile Frameworks for Public Sector.

1. Introduction

This paper aims to report the construction of the framework used by Pernambuco's Strategic Monitoring Data Science Team (SMT) to develop and implement data solutions to monitor strategic public policies. The structure of the paper follows a variation of the proposition in Young and Quinn (2002), with introduction, methodology, results, and conclusion.

The introduction discusses the problem situation and the available alternatives and describes the Pernambuco's Strategic Monitoring Data Science Team. The methodology presents the problem statement, the method and the materials used by the team to create the solution. The results present the framework itself, and the products created by the team. The conclusion summarizes main points into recommendations for those who want no use the framework, discusses limitations and suggests follow-up steps.

1.1 The Problem

First, we must address the question concerning terminology and concepts. Analyzing the question, Saltz and Grady (2017) emphasize the absence of concepts and terms to describe data science and its process, team roles, etc. Aho et al (2021) indicates that there is also a confusion between data science and related concepts (like big data, data science, analytics, machine learning, artificial intelligence, etc.) due to a shared space. Provost and Fawcett (2013) emphasize that these other concepts tend to have a clearer connection with daily practice, resulting, many times, in a confusion between data science and its materials (like big data) or goals (like data-driven decision making).

One good summary of the question is offered by Brady (2019): big data and data science have different meanings but are frequently related to a new volume of data that comes with new methods, new questions, new ethical issues and societal change. Specifically, the author says that Data Science comprehends any process of gathering, representing, preparing, transforming, exploring, modeling, visualizing or presenting data, either using big or small data. Thus, even if big data and data science are different aspects, whit their own elements, the practical use and discussion about any type of data, especially in government, needs to be done along with the debate of data science, since the last governs the use of first to generate value.

Although Data Science is already producing societal changes and increasingly being used as an important tool for business and science (Klievink et al, 2016), according to Nielsen, Persson and Madsen (2019), there are still some barriers to a more complete implementation, which we can classify as technical barriers (like quality, storage and availability) and management barriers (like culture, perception and training).

The government is still starting to utilize data to make better decisions, improve efficiency or better understand social demands to solve public problems. This is a relevant issue, since in the landscape of data, governments are a major player in terms of owning and distributing data, and changes in public sector use of data have the potential to generate even greater impact (Filgueiras and Lui, 2023; Giest, 2017; Nielsen, Persson and Madsen, 2019).

When we think about public sector, all previous barriers also apply and we can add some others like data and information silos (Scott and Gong, 2021), lack of study reports (Nielsen, Persson and Madsen, 2019) and legislation about the use *per si* (Giest, 2017), and more compelling concerns about ethics, fairness and equality, which generates uncertainties and holds back a more complete use (Klievink et al, 2017).

One key element to promote and institute data science as a value approach is the construction of data science teams. Kim et al (2016) states that data scientists are an emerging role on traditional software development teams, but other authors like Aho et al (2021) contribute to show that data science teams multidisciplinary with data scientists coming from different backgrounds, with a recent trend to train data scientists on specific skills.

Both lines agree that, whilst there are some homogeneous aspects like experimentation and iterative development, the practical use of data science presents a large variation in variables like team composition, organization, roles and size; position in organization (internal / external); management framework and process; use of big or small data; project scope (from simply organizing data to advanced uses of machine learning).

Another common ground is that education and lack of definition about roles in data science teams are a key component. Saltz and Grady (2017) argue that in this context the formation of a data science team is still an issue, since it is difficult not only to train people, but to attribute team roles and to create team framework. Considering the Brazilian case, De Toni and Dorneles (2022) say that the efforts to train public servants on data science is still recent and that there is a large variability on the use. Kim et al (2016) reinforces that, in many situations, unlike traditional development teams, Data Science teams focus more on skills than on roles and static attributions.

However, another relevant point, however, is that, even with different backgrounds and skill sets on Data Science Teams, the available frameworks are still focused or emerge from software engineering backgrounds, resulting on a lack of orientations regarding other Data Science tasks, such as research and study reports.

This scenario of low definition, high variability, and few specific studies on government data science teams leads to a problem-situation in which government data science teams need to create their own framework, comprehending team roles, scope of projects and development process, expanding the available ones to include missing parts, like research and study reports.

1.2 Alternatives

Saltz and Grady (2017) produced an interesting cross section study about different cases of frameworks for data science team organization, skills and roles, compared with the types of data scientists job positions on dice.com. The findings showed that the main roles are researcher, scientist, architect, analyst, programmer and engineer.

Expanding the first study, Saltz and Hotz (2020) conducted a new survey-based study, this time focused on projects frameworks. They found out that CRISP-DM was still the most used, reaching 49% of the respondents, followed by scrum (18%), kanban (12%), TDSP (4%) and SEMMA (1%). Customized frameworks (12%), others (3%) and none (2%) were also mentioned.

Resuming the two studies, we may say that: (i) available project frameworks focus more on methods / steps (CRISP-DM, TDSP, SEMMA) or governance / coordination (Scrum, Agile, Kanban) (i) the main framework, CRISP-DM was built in the late 1990s, before some of the actual technologies were even available; and (iii) of the organization frameworks only EDISON focused on researches data scientists.

Considering that available frameworks (i) comprehend only the methods or the management, (ii) are not suited for government, (iii) don't address the role of research on data science projects and teams, and that (iv) is common that data science teams feel the need to develop their own frameworks because of the range of skills, tasks and data need to develop their projects, after the research for alternatives, we reinforced our belief on needing to develop our own framework.

1.3 Problem Statement

The problem statement was done using the How Might We? – HWE (Brown, 2020) to produce a statement capable of solving and opened enough to allow innovative solutions: How might we create a data science framework comprehending team roles, scope of projects and development process, to increase the efficiency of data projects for strategic monitoring?

1.4 The Strategic Monitoring Team

The Strategic Monitoring is a Data Science Team focused on creating data solutions to extract value from data, In Pernambuco State (a Brazilian subnational government), the team is a Sub-secretariat of the State Secretariat of Strategic Projects, responsible for the implementation and monitoring of strategic public investments.

The team is composed of five public servants and one trainee, who gives support for different activities. Three members are data scientists that divide the tasks according to project thematic; one member is the project manager; and the other is the chief data scientist, also occupying the position of sub-secretary of strategic monitoring. Except for the project manager, who has a strong background in business and government, the members have strong academic backgrounds, with three PhDs.

Different tasks are performed according to specific project needs, but the main are data engineering and wrangling, exploratory data analysis, data modeling to predict and classify, DataViz, optimization to create prescription alternatives, construction of study reports, dashboards and presentations. In some cases, the team also develops data input software.

The team uses R as its main technology. R is used to tasks from data engineering to data modelling, and to create the templates used to develop apps with shiny framework, and study reports in .pdf or .html or presentations using reveal.js through quarto framework. The deploy is done using shiny proxy, which offers better load balance, routing and access control.

It is important no say that, since the team is responsible for strategic monitoring, the portfolio construction and project selection criteria is mainly top-down. The team receives demands from the secretary of strategic projects, the secretary of mobility and infrastructure or directly from the center of government.

The figure resumes the team's organization and tasks.

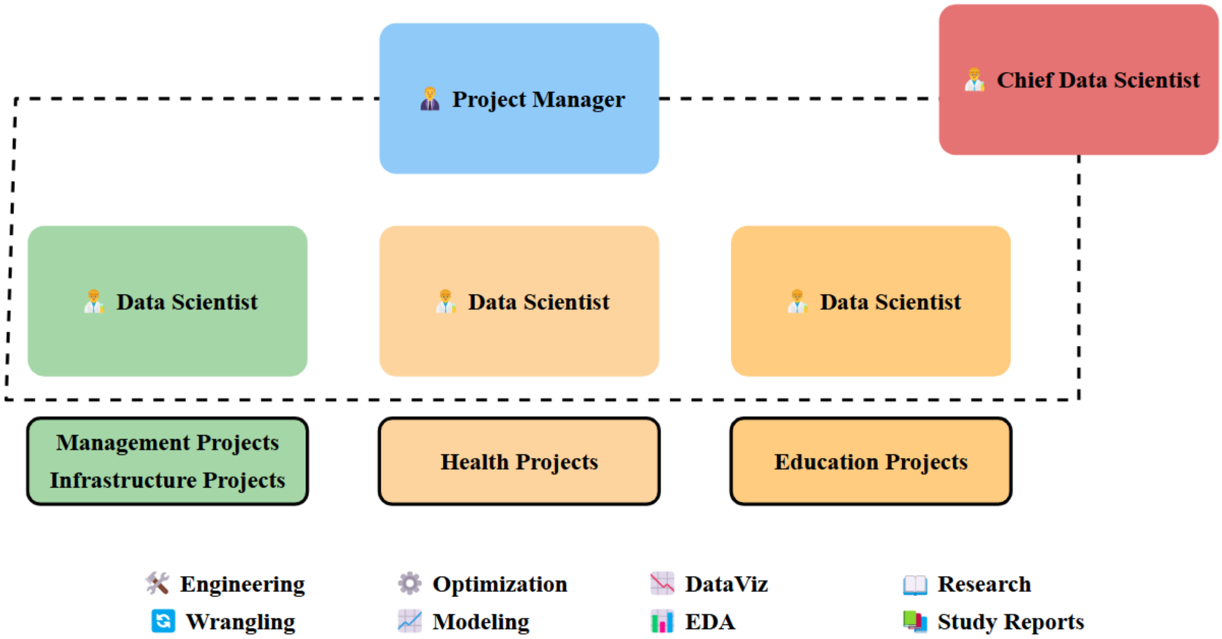


Fig 1 – Strategic Monitoring Team.

2. Methodology

2.1 Method

The Strategic Monitoring Team framework was created considering the literature of alternatives to manage data science team, the analysis of the team's empirical practice and the elements (artifacts, ceremonies and products) of scrum (Schwaber and Sutherland, 2020), using action research methodology as the conductor. The results are the framework design and details, a table comparing the artifacts, ceremonies and products of scrum with those used by the team, and a summary of the data products build so far. The discussion resumes the alternatives limitations to government data science teams and the learned lessons from the Strategic Monitoring Team.

For example, during the development of one of the products, the app Hortensias, it was always important to focus on problem-solving and practitioners' involvement, but also to document the innovation process and the software development. So, the leading methodology used was action research, which can be defined as a cycle process of diagnosis, plan, act, evaluate and document to improve practice or solve a problem, combining acting and inquiring about it (Tripp, 2005).

As argument by Ollila and Yström (2020), action research is an important method to conduct innovations because it is more capable to get practitioners involved, which tends to approximate solutions and problems, while generating richer insights and more significant knowledge.

Some authors as Saarikallio and Tyrväinen (2022) and Natarajan and Pichai (2024) have achieved success on utilizing action research to inquire about agile software development process, since the cycle nature of agile has an adequate fit to also cycle perspective of action research. But Action Research can offer contributions to Data Science Teams because it has a more robust approach to dealing with the research aspects of Data Science tasks.

Regarding data science specific techniques, the data engineering process uses the extract, transform and load – ETL framework (Kimball, and Ross, 2013), storing the final data with proper dimensions and metrics in a corporate GreenPlum distribution; the monitoring process uses exploratory analysis statistics and data visualization such as time series, key performance indicators, and data summaries (Tukey, 1980; Wickham, and Wickham, 2016; Staniak, and Biecek, 2019). Finally, in some cases the team uses machine learning and optimization techniques.

Thus, our methodology uses action research to review team roles, artifacts and ceremonies, considering integrating research.

2.2 Materials

The team uses materials to perform its action research cycle the notion platform with a custom workspace, divided in:

- **Projects**. Page with data and information linked to current projects, including backlog, reunions, project chartet, etc.
- **Reunions**. Page with data and information linked to reunions, with focus on calendar, tasks and follow-up taks.
- **Notes**. Page with annotation and miscelania information that is note directly linked to any of the other pages.
- **Cerimonies**. Page with description of the cerimonies adopted by the team.
- **Databases**. Page with the databases used on other pages.

Other relevant material is the team landpage, called box (https://box.pe.gov.br/), in honor of George E. P. Box. The landpage concentrates the products developed by the team, offering a summary, a point of access and the link to the team repository on github, which has the codes and proof of concepts (POC) of all products.

3. Results

3.1 Action Research Cycle

Following literature indications, we created our own action research cycle, adapted to our DS team routines. But, before the cycle enters operation, it is necessary that a project be presented by the secretary or prospected by the team, and the project charter being written in general lines.

- **Step 1 – Diagnose**. Meetings with practitioners that live and understand the problem context, the pain points and the reasons why their occurs. The meetings were translated into high level requirements and follow-up tasks using notion. After the meeting, the team also concudcts research tasks (simple literature revision and bechmarking) to get a more general understand of the problem and to prospect what are the possible solutions or best practices. The meetings usualy take 2h to 3h, and is scheduled by the team's project manager.
- **Step 2 – Plan**. Meetings with the DS team to define discuss research findings, to define products design and features, to estimate time and to build the backlog. Notion is used to register and monitor tasks.
- **Step 3 – Act**. Hands-on, effectively developing the products based on the diagnoses and the research findings. The development is done using the templates available on the team package called vialactea. At this moment, the data scientist have individual daily meetings with the chief data scientist and the project manager.
- **Step 4 – Evalute**. Meetings with the same practitioners in step 1 with the goal to evaluate if the producuts solve the pain points. This steps gets a little confused with step 1, in the sense that practitioners tend to address new pain points continuasly as the previsou are solved. Once the project is deployed and delivered, it enters on a continuous process of strategic monitoring, being used in decision-making. At this point, the evauation is incorpored in monitoring reunions.
- **Step 5 – Learn**. Document the findings and lessons during the DS team meetings. Github repo is updated and the main lessons go to the R package vialactea to become default features in further projects and to be back implemented in previous solutions. Also, the products are deployed and their information is made available in the team landapge. This is a team moment, lead by the chief data scientist and the project manager.

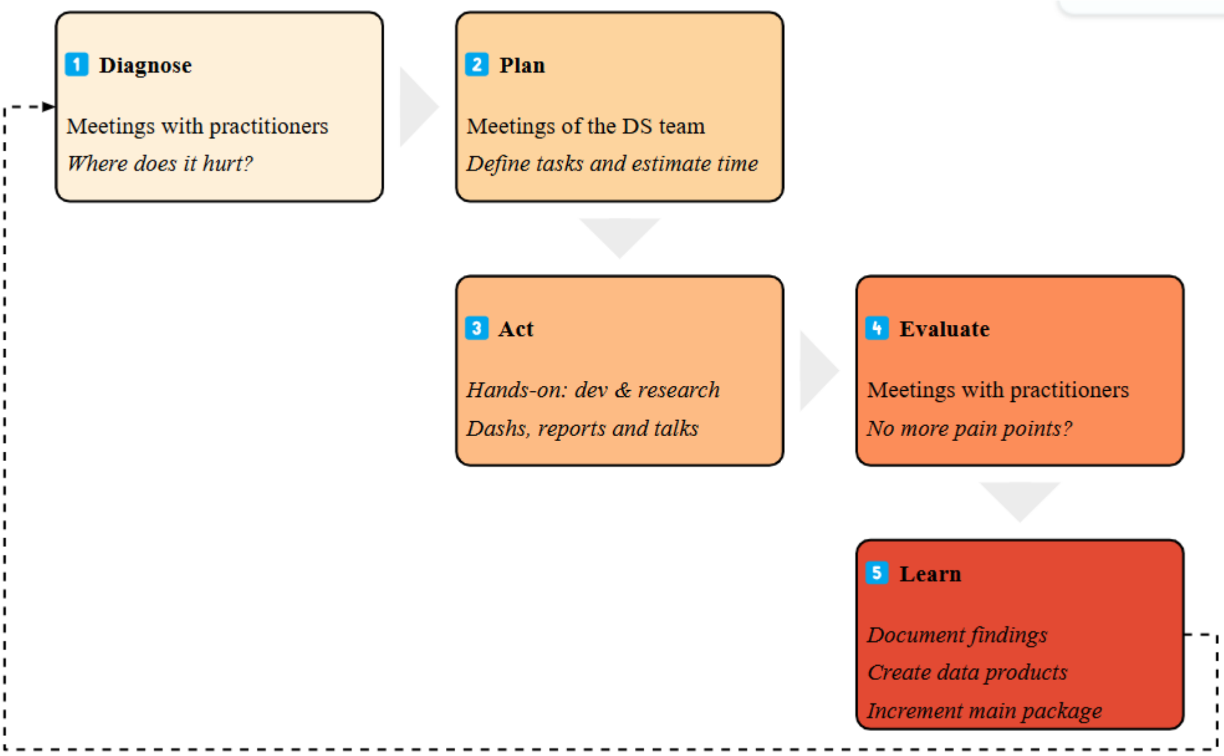


Fig 2 – Summary of the DS Team Action Research Cycle.

3.2 Artifacts, ceremonies and roles

Project management is done using a variation of scrum, in which the project manager acts as scrum master, and the data scientists act like product owners and development team. As POs, data scientists are responsible for interactions with demandant practitioners and business teams, and also for performing academic research to better perform data tasks. The chief data scientist deals with administrative demands but also as leader developer, being responsible for establishing the software developing standards, study report guides and research methods.

Thinking about the artifacts and the ceremonies, we may say that product backlog and sprint backlog are used, as also product increment. These artifacts are adapted to be used not only for software products, but also for research products, resulting, in this case in shorter backlogs with topics to be explored. The team has a weekly sprint, with planning and revision meetings occurring on the same day. Retrospective meetings are done once per month for software, and, for research, when the first summary of findings is done and when the presentation and report are done.

The following table compares scrum elements and the ones used by the team, detaching with * those that do not exist on Scrum.

Tab. 1 – Comparison of artifacts, ceremonies and roles on scrum and on Strategic Monitoring Team.

ELEMENT TYPE	SCRUM	STRATEGIC MONITORING TEAM
ARTIFACT	Product Backlog, Sprint Backlog, Product Increment	Used in software and research products
CEREMONIES	Backlog Refinement, Planning and Revision	Weekly in the same day
CEREMONIES	Daily	The same
CEREMONIES	Retrospective	Once per month
ROLE	Scrum Master	Project Manager
ROLE	Product Owner	Data Scientist
ROLE	Team member	Data Scientist and Chief Data

		Scientist
ROLE	Administration*	Chief Data Scientist

3.3 Products

The team products are divided into two main different axes: software products (apps, chatbot and packages) and research products (presentations and study reports).

Apps. Data Science Softwares that offer dynamic dashboards to interacte with data and automate analytics tasks like prediction and optimization. The apps have acces control and collect usage statistics. Some exemples are: (i) *Amparo*, an app to monitor highways construnction, focused on data engineering of 4 different data silos and dataviz; and (ii) *Hortensias*, an app to monitor public spendings, that also has a heavy use of data engineering and dataviz, but also focus on more complex tasks such as spending predictions and budget optimization.

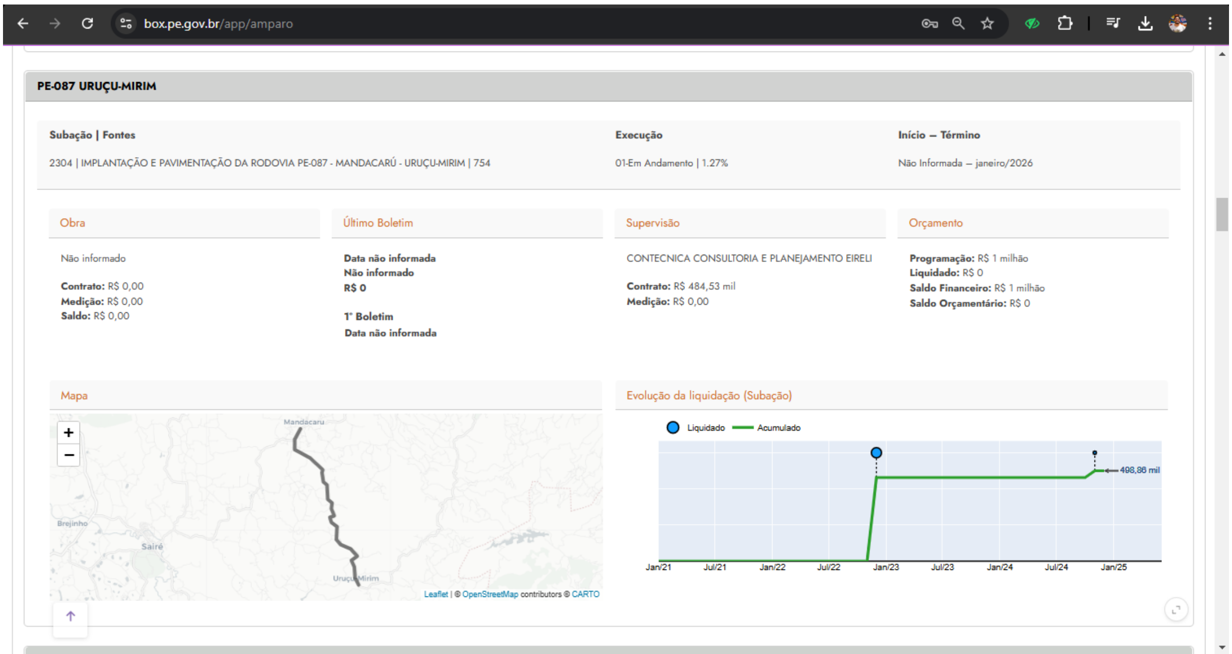


Fig 3 – Amparo App.

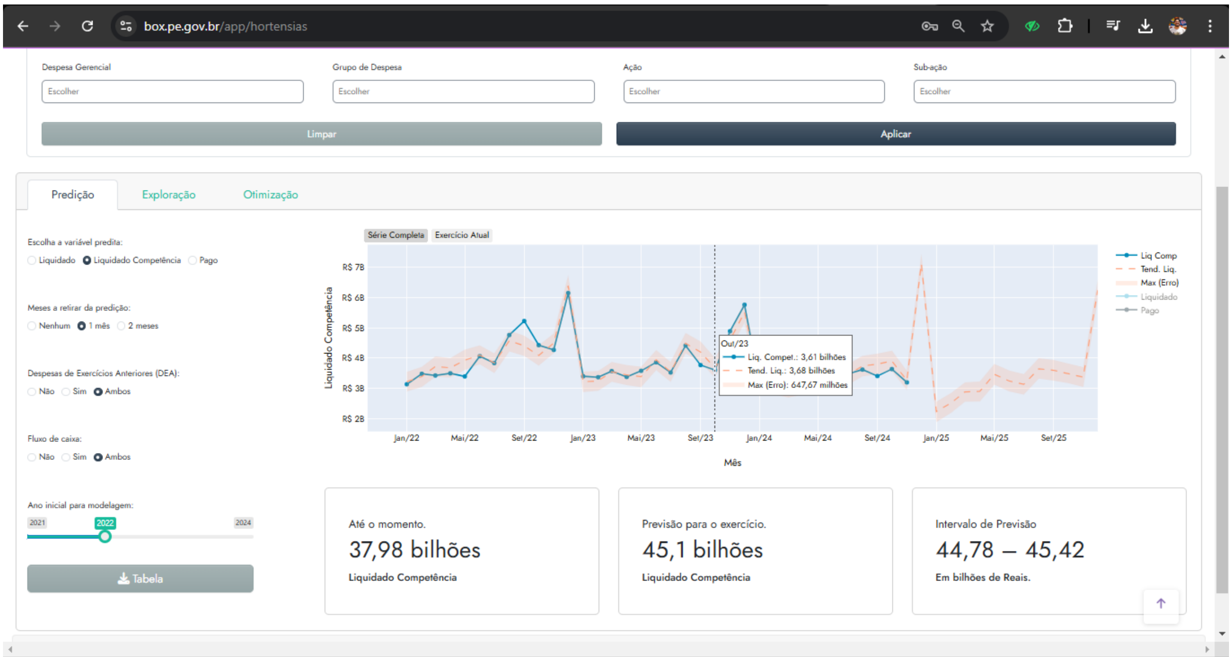


Fig 4 – Hortensia App.

Argos. Argos is our WhatsApp chatbot that offers just in time information to strategic users via text messages. Currently, Argos is capable of reading most of the databases used on apps and has commands to summarize and present data in text format.

Fig 5 – Argos chatbot.

Packages. When the data science team creates solutions that may be used in other contexts, they are delivered in the form of a R package. The example is the BigDataPE an R package that provides a secure and intuitive way to access datasets from the BigDataPE platform.

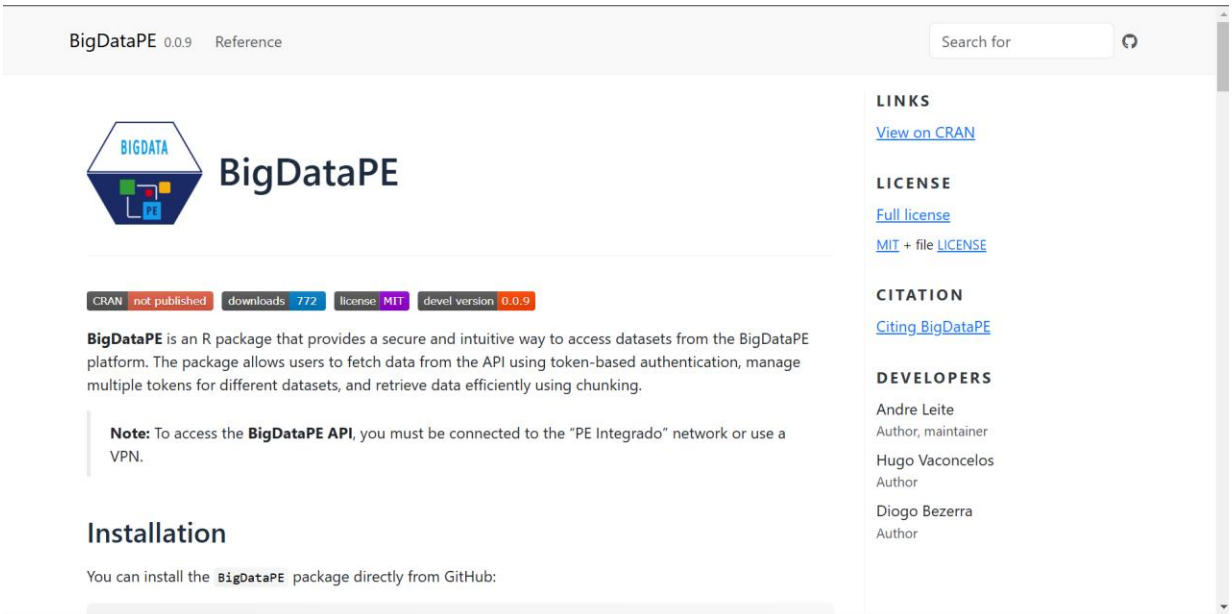


Fig 6 – Maternity study report.

Presentations. Data Science presentations based on research and data analysis. Less dynamic but with defined storytelling. One example is REDIPE presentation, which focuses on present the research and benchmarking of alternatives to manager data of public investments.

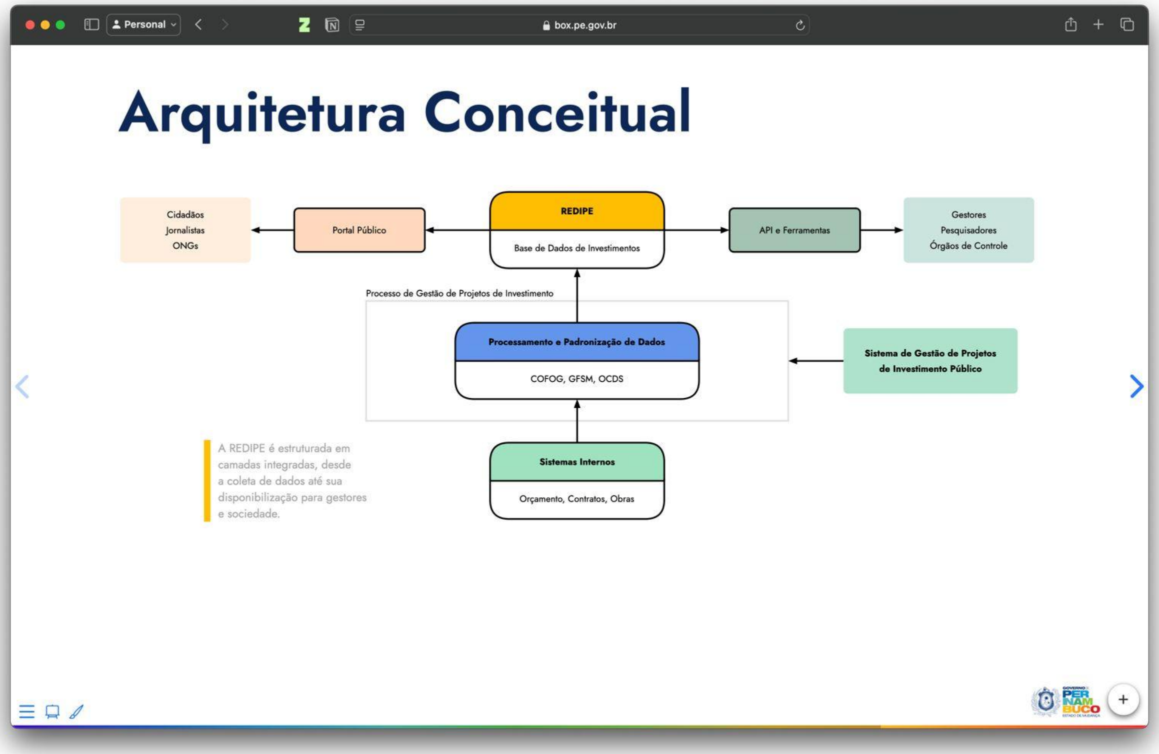
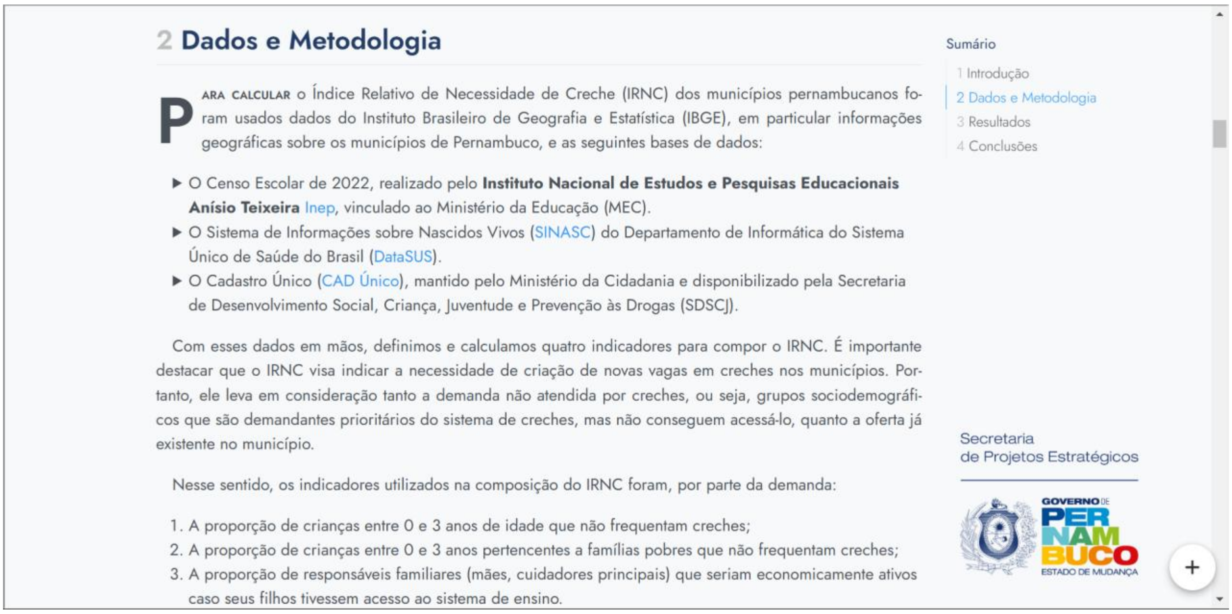


Fig 7 – REDIPE presentation.

Study Reports. Data science study reports based on research and data analysis. Less dynamic and with defined storytelling like the presentations. The description is denser and aimed at detailing all research aspects and data analysis. An example is maternity location analysis, which focuses on present the research and data analysis of early childhood education.



2 Dados e Metodologia

Para calcular o Índice Relativo de Necessidade de Creche (IRNC) dos municípios pernambucanos foram usados dados do Instituto Brasileiro de Geografia e Estatística (IBGE), em particular informações geográficas sobre os municípios de Pernambuco, e as seguintes bases de dados:

- O Censo Escolar de 2022, realizado pelo **Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira** Inep, vinculado ao Ministério da Educação (MEC).
- O Sistema de Informações sobre Nascidos Vivos (SINASC) do Departamento de Informática do Sistema Único de Saúde do Brasil (DataSUS).
- O Cadastro Único (CAD Único), mantido pelo Ministério da Cidadania e disponibilizado pela Secretaria de Desenvolvimento Social, Criança, Juventude e Prevenção às Drogas (SDSCJ).

Com esses dados em mãos, definimos e calculamos quatro indicadores para compor o IRNC. É importante destacar que o IRNC visa indicar a necessidade de criação de novas vagas em creches nos municípios. Portanto, ele leva em consideração tanto a demanda não atendida por creches, ou seja, grupos sociodemográficos que são demandantes prioritários do sistema de creches, mas não conseguem acessá-lo, quanto a oferta já existente no município.

Nesse sentido, os indicadores utilizados na composição do IRNC foram, por parte da demanda:

1. A proporção de crianças entre 0 e 3 anos de idade que não frequentam creches;
2. A proporção de crianças entre 0 e 3 anos pertencentes a famílias pobres que não frequentam creches;
3. A proporção de responsáveis familiares (mães, cuidadores principais) que seriam economicamente ativos caso seus filhos tivessem acesso ao sistema de ensino.

Sumário
1 Introdução
2 Dados e Metodologia
3 Resultados
4 Conclusões

Secretaria de Projetos Estratégicos

GOVERNO DE PERNAMBUCO ESTADO DE MUDANÇA

Fig 8 – Maternity study report.

4. Conclusion

The first point to address is that, as highlighted in the literature about data science teams, data science is a broad and complex field, deeply connected to other disciplines in science, technology, and innovation. During our research, we found no specific frameworks that fully aligned with our team's primary goal: creating data-driven monitoring solutions. Existing frameworks were either too generic—such as Scrum and Kanban—or too focused on technical data science workflows, often neglecting the research phase and the structured development of necessary products. This gap led us to develop our own tailored framework.

Action research proved to be a highly effective methodological approach for customizing our artifacts, ceremonies, and roles, as it integrates both the research and iterative development aspects of data science, providing structured guidance for inquiry and reflection on practice. This dual focus allows for continuous adaptation and improvement, making it particularly well-suited for governmental contexts, where rigid bureaucratic structures often clash with the need for flexibility and responsiveness.

One relevant limitation of such approaches as action research is the selection and confirmation bias, because the researches are very close to the problems they are investigating. Thus, adaptations of the framework must consider this limitation, and procced an adaptation to the context of each organization, since aspects like regulation, time size, governance, and number and resistance of data-silos vary from institution to institution.

Also, the adaptations of traditional agile framework are necessary due to public sector constraints, like a more rigid organizational structure, strong hierarchization and segmentation of activities, as well as severe restrictions regarding resource availability and staff training. One of the most challenging points is to balance the roles, since in original Scrum framework the Scrum Master and the Product Owner have framed activities and responsibilities, which need to be done by the development team. For this to work, it was very important to have a senior team, and to focus every data scientist on a theme portfolio. Nevertheless, those aspects have been experienced in other similar implementations in Brazil, as documented by different authors, like Vacari (2015), Rosa and Pereira (2021), and Guimarães et al (2025).

All things considered, and beyond the specific needs of our team, we believe that our framework addresses fundamental challenges faced by government data science teams. These include the difficulty of coordinating traditional hierarchical structures with matrix management approaches, as well as the broader struggle to balance agility with institutional requirements. Our framework offers a foundation upon which public sector organizations can customize agile elements while ensuring alignment with governance protocols.

When considering the public sector, additional challenges further complicate the implementation of data-driven approaches. Issues such as data and information silos, the scarcity of comprehensive study reports, and legal constraints regarding data usage create structural barriers to effective data science initiatives. Furthermore, ethical concerns related to fairness, transparency, and equality add layers of complexity, fostering uncertainties that can hinder the full adoption of data-driven methodologies. Addressing these barriers will require not only technical solutions but also institutional and regulatory adaptations to create an environment where data science can thrive in the public sector.

At the moment we don't have a method to evaluate impact of the solutions regarding efficiency, efficacy, influence in decision-making or comprehension about social problems. Nevertheless, we understand that this is a general limitation of evaluating the impact of monitoring and business intelligence solutions.

As a continuous agenda, with more products being created, we aim to develop mechanisms for systematically collecting organic and structured feedback, ensuring a deeper understanding of user needs and experiences. Another pending issue is the transformation of our landing page into a more informative and user-friendly platform, enhancing transparency and accessibility. Finally, we recognize the need for a sustainable workflow for technological retrofitting, enabling the incorporation of lessons learned from previous projects into future iterations. These steps will contribute to the long-term sustainability and effectiveness of our data-driven monitoring solutions in the public sector.

References

Aho, T., Kilamo, T., Lwakatare, L. E., & Yaman, S. (2021). *Managing and composing teams in data science: An empirical study. In 2021 IEEE International Conference on Big Data* (Big Data). IEEE. https://doi.org/10.1109/BigData52589.2021.9671737.

Brady, H. E. (2019). *The challenge of big data and data science. Annual Review of Political Science*, 22(1), 297-323. https://doi.org/10.1146/annurev-polisci-090216-022929.

Brown, T. (2020). *Design thinking: Uma metodologia poderosa para decretar o fim das velhas ideias*. Elsevier.

De Toni, J., & Dorneles, R. (Orgs.). (2022). *Ciência de dados em políticas públicas: Uma experiência de formação* (Colaboração de N. Koga, H. Macedo & A. Pereira). Enap. https://repositorio.enap.gov.br/handle/1/7472.

Filgueiras, F., & Lui, L. (2023). Designing data governance in Brazil: An institutional analysis. *Policy Design and Practice*, 6(1), 41-56. https://doi.org/10.1080/25741292.2022.2065065.

Giest, S. 2017. "Big Data for Policymaking: Fad or FastTrack?" *Policy Sciences* 50 (3): 367–382. doi:10.1007/s11077-017-9293-1.

Guimarães, L. S., Toledo, G. L., Rodrigues, K. C. D. G., & Silva, P. N. (2025). Práticas ágeis no setor público: Estrutura de governança do Acordo de Reparação de Brumadinho à luz do framework Scrum. *P2P & Inovação*, 11(2), 1–28. https://revista.ibict.br/p2p/article/view/6847.

Kim, M., Zimmermann, T., DeLine, R., & Begel, A. (2016). The emerging role of data scientists on software development teams. In *2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE)* (pp. 96-107). IEEE. https://doi.org/10.1145/2884781.2884783.

Kimball, R.; Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*. Indianapolis: Wiley.

Klievink, B., Romijn, B.-J., Cunningham, S., & de Bruijn, H. (2017). Big data in the public sector: Uncertainties and readiness. *Information Systems Frontiers*, 19(2), 267–283. https://doi.org/10.1007/s10796-016-9686-2.

Natarajan, T., & Pichai, S. (n.d.). Transition from Waterfall to Agile Methodology: An Action Research Study. *IEEE Xplore*. Retrieved from https://ieeexplore.ieee.org/document/10488855.

Nielsen, O. B., Persson, J. S., & Madsen, S. (2019). Why governing data is difficult: Findings from Danish local government. In A. Elbanna, Y. Dwivedi, D. Bunker, & D. Wastell (Eds.), *Smart working, living and organising* (Vol. 533, pp. XX-XX). Springer, Cham. https://doi.org/10.1007/978-3-030-04315-5_2.

Olilla, S., & Yström, A. (2020). Action research for innovation management: Three benefits, three challenges, and three spaces. *R&D Management*, 50(3), 396–411. https://doi.org/10.1111/radm.12407.

Provost, F., & Fawcett, T. (2013). *Data science and its relationship to big data and data-driven decision making. Big Data*, 1(1), 51–59. https://doi.org/10.1089/big.2013.1508.

Rosa, M. R. da, & Pereira, E. N. (2021). Metodologias ágeis no contexto da administração pública: Análise de estudos de caso de implementação ágil. *Revista do Serviço Público*, 72(2), 479–497. https://revista.enap.gov.br/index.php/RSP/article/view/4310

Saarikallio, M., & Tyrväinen, P. (2023). Quality culture boosts agile transformation—Action research in a business-to-business software business. *Journal of Software: Evolution and Process*, 35(1), e2504. https://doi.org/10.1002/smr.2504.

Saltz, J. S., & Grady, N. W. (2017). The ambiguity of data science team roles and the need for a data science workforce framework. In *2017 IEEE International Conference on Big Data (BIGDATA). IEEE*. https://doi.org/10.1109/BigData.2017.8258401.

Saltz, J. S., & Hotz, N. (2020). Identifying the most common frameworks data science teams use to structure and coordinate their projects. *2020 IEEE International Conference on Big Data (Big Data)*, 1615–1624. https://doi.org/10.1109/BigData50022.2020.9377813.

Scott, I., & Gong, T. (2021). Coordinating government silos: Challenges and opportunities. *Global Public Policy and Governance*, 1, 20–38. https://doi.org/10.1007/s43508-021-00004-z.

Schwaber, K., & Sutherland, J. (2020). The Scrum guide: The definitive guide to Scrum: The rules of the game.

Scrum.org. https://scrumguides.org.
Tripp, D. (2005). Action research: A methodological introduction. *Educação e Pesquisa*, 31, 443–466. https://doi.org/10.1590/S1517-97022005000300009.
Tukey, J. W. (1980). We need both exploratory and confirmatory. *The American Statistician*, 34(1,) 23–25. 10.1080/00031305.1980.10482706.
Vacari, I. (2015). *Um estudo empírico sobre a adoção de métodos ágeis para desenvolvimento de software em organizações públicas* (Dissertação de mestrado, Pontifícia Universidade Católica do Rio Grande do Sul). PUCRS Repositório Digital. https://tede2.pucrs.br/tede2/bitstream/tede/6189/2/471291%20-%20Texto%20Completo.pdf.
Wickham, H.; Grolemund, G. (2016). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media. .
Young, E., & Quinn, L. (2002). *Writing effective public policy papers: A guide for policy advisers in Central and Eastern Europe*. Open Society Institute. Retrieved from https://caw.ceu.edu/sites/caw.ceu.edu/files/attachment/basicpage/116/writingeffectivepublicpolicypapersyoungquinn.pdf.