Implementing AI for Enhanced Public Services Gov.br: A Methodology for the Brazilian Federal Government

Maísa Kely de Melo[1,6,*][a], Silvia Araújo dos Reis[1,7][b], Vinícius Di Oliveira[2,11,12][c], Allan Victor Almeida Faria[1,4][d], Ricardo de Lima[3,12][e], Li Weigang[2,12,*][f], Jose Francisco Salm Junior[1,13][g], Joao Gabriel de Moraes Souza[1,9][h], Vérica Freitas [8][i], Pedro Carvalho Brom[2,10,12][j], Herbert Kimura[1,7][k], Daniel Oliveira Cajueiro[1,5][l], Gladston Luiz da Silva [1,4][m] and Victor Rafael R. Celestino[1,7,*][n]

[1]*LAMFO - Lab. of ML in Finance and Organizations, University of Brasília, Campus Darcy Ribeiro, Brasília, Brazil*
[2]*TransLab, Department of Computer Science, University of Brasília, Campus Darcy Ribeiro, Brasília, Brazil*
[3]*Ministry of Management and Innovation in Public Services, Federal District, Brazil*
[4]*Department of Statistics, University of Brasília, Campus Darcy Ribeiro, Brasília, Brazil*
[5]*Department of Economics, University of Brasília, Campus Darcy Ribeiro, Brasília, Brazil*
[6]*Department of Mathematics, Federal Institute of Education, Science and Technology of Minas Gerais, Formiga, Brazil*
[7]*Department of Business Administration, University of Brasília, Campus Darcy Ribeiro, Brasília, Brazil*
[8]*School of Business and Management, Uberlandia Federal University, Uberlândia, Brazil*
[9]*Department of Economics and Business Administration, Brazilian Institute of Education Development and Research - IDP, Brasília, Brazil*
[10]*Department of Mathematics, Federal Institute of Education, Science and Technology of Brasília, Campus Estrutural, Brasília, Brazil*
[11]*Federal District Secretariat of Economy, Brasília, Brazil*
[12]*Department of Computer Science, University of Brasília, Campus Darcy Ribeiro, Brasília, Brazil*
[13]*University of the State of Santa Catarina, Florianópolis, Santa Catarina, Brazil*

Keywords: Brazil, Design Sprint, Public Administration, AI, Machine Learning, LLM.

Abstract: The website portal of the Brazilian federal government (Gov.br) consists of pages from almost 40 ministries, 180 public agencies and up to 5000 public services for all citizens, posing a significant challenge in improving service quality. This article presents an innovative methodology to implement artificial intelligence (AI) to address these challenges, to enhance the efficiency, accessibility, and quality of services to the population. The methodology combines elements of Lean Office, Design Sprint, Analytic Hierarchy Process (AHP), and advanced AI techniques, particularly Large Language Models (LLMs), making it flexible and adaptable to the needs of government entities. Developed in collaboration with project managers, public servants, and stakeholders, the methodology includes a survey of demands, selection, and prototyping of AI projects in a complex government context. The practical application selected the Gov.br portal for prototyping, involving the development of an advanced generative agent to interact with citizens, clarify doubts, direct to the requested services, and provide human interaction when necessary. The recommended practices offer a valuable contribution to other developing countries seeking to integrate AI solutions into their public services.

[a] https://orcid.org/0000-0001-8120-9778
[b] https://orcid.org/0000-0002-1646-4454
[c] https://orcid.org/0000-0002-1295-5221
[d] https://orcid.org/0000-0002-4300-9334
[e] https://orcid.org/0000-0002-9269-3558
[f] https://orcid.org/0000-0003-1826-1850
[g] https://orcid.org/0000-0002-8492-1645
[h] https://orcid.org/0000-0003-0685-3082
[i] https://orcid.org/0000-0003-3035-9738
[j] https://orcid.org/0000-0002-1288-7695
[k] https://orcid.org/0000-0001-6772-1863
[l] https://orcid.org/0000-0001-5898-1655
[m] https://orcid.org/0000-0001-9650-2993
[n] https://orcid.org/0000-0001-5913-2997
*Corresponding author

Kely de Melo, M., Reis, S., Di Oliveira, V., Faria, A., de Lima, R., Weigang, L., Salm Junior, J., Souza, J., Freitas, V., Brom, P., Kimura, H., Cajueiro, D., Luiz da Silva, G. and Celestino, V.
Implementing AI for Enhanced Public Services Gov.br: A Methodology for the Brazilian Federal Government.
DOI: 10.5220/0012997000003825

In *Proceedings of the 20th International Conference on Web Information Systems and Technologies (WEBIST 2024)*, pages 90-101
ISBN: 978-989-758-718-4; ISSN: 2184-3252

1 INTRODUCTION

In recent years, the public sector has increasingly explored artificial intelligence (AI) to enhance operations and improve service delivery (Maragno et al., 2023). There are at least 142 AI applications in federal administrations (Engstrom et al., 2020). Some focus on making government information more accessible, while others aim to improve citizen interaction through Information and Communication Technologies (ICTs) (Kitsios et al., 2023). AI technologies, such as machine learning (ML), natural language processing (NLP), and computer vision, hold great potential for automating tasks, analyzing large datasets, and offering more efficient, personalized services.

The disruptive potential of AI in the public sector is evident in three main areas: enhancing internal efficiency, improving decision-making, and strengthening citizen-government interaction. AI applications in the public sector include facial recognition in policing (Blount, 2024), recidivism prediction in criminal justice (Wang et al., 2023), virtual agents for process automation (Scutella et al., 2024), forecasting social service needs (Dwivedi et al., 2023), chatbots (Chen et al., 2024), healthcare (Yu et al., 2018), public transport (Jevinger et al., 2024), education (Cohen et al., 2023), environmental management (Fan et al., 2023), and agribusiness (Kutyauripo et al., 2023). These initiatives promote more inclusive services and enhance citizen participation in public sector activities (Samoili et al., 2020).

Integrating artificial intelligence into the public sector is evolving, with much still unknown about its full potential and optimal strategies. Existing research on AI adoption highlights three key values: transparency, effectiveness, and efficiency (Rocha et al., 2022; Faria et al., 2022; Chen et al., 2023). Recognizing the need for clear guidelines, the Brazilian federal government aims to identify the best practices for selecting AI projects in public agencies.

This paper presents the methodology developed for the project "Artificial Intelligence Research Applied to the Prototyping of Solutions for the Federal Government's Artificial Intelligence Office." The methodology focuses on mapping, prototyping, and documenting AI solutions for public administration issues, aligning with the responsibilities of the Secretariat of Digital Government (SGD) (BRASIL, 2023), under the Ministry of Management and Innovation in Public Services. The proposed solution supports AI governance and the development of data models, processes, and standards, including data mining, analysis, and visualization technologies. These practices enhance the public policy management cycle and the provision of services. Given that these responsibilities are part of other countries' digital transformation strategies, the methodology offers a valuable contribution to governmental agencies developing AI solutions for public services.

It also details the step-by-step methodology for selecting and prototyping projects in AI Decision-Making in the Public Sector. This methodology was developed and applied in collaboration with AI and project management expert researchers, members of the SGD, and other bodies comprising the Brazilian Government's Artificial Intelligence Hub (NIA). Figure 1 illustrates NIA, highlighting partner institutions' diversity. The results of the methodology's application in federal government agencies are reported, highlighting recommended practices and obtained outcomes. The project selected for prototyping, arising from the demands identified by the public agencies participating in the research, was the Gov.br portal service (BRASIL, 2024). This portal comprises pages from nearly 40 ministries and 180 public agencies, integrating up to 5000 public services. Over 150 million Brazilians access it and records approximately four thousand accesses per second (Gov, 2024). The prototyping involved the development of an Advanced Generative Agent to interact with citizens, assist them, clarify doubts, and direct them to the requested service pages, with human overflow in necessary cases.

2 RELATED WORK

We present the following works that explore various aspects of AI solutions for public services and the techniques used in the proposed methodology. Additionally, we contextualize the application of AI in public services within the Brazilian framework.

2.1 Public Service with AI

Implementing AI in the public sector requires understanding its capabilities and addressing challenges like explainability, data governance, and infrastructure. Human-AI collaboration can improve tasks and create new jobs, but caution is essential to avoid mistrust. Engaging stakeholders, such as citizens and suppliers, foster innovation but demands effective ecosystem management and clear regulations (Maragno et al., 2023). Currently, there is no unified global approach to determining AI's legal status (Atabekov, 2023), and most countries focus on the process rather than the final result generated by AI (Atabekov, 2023).

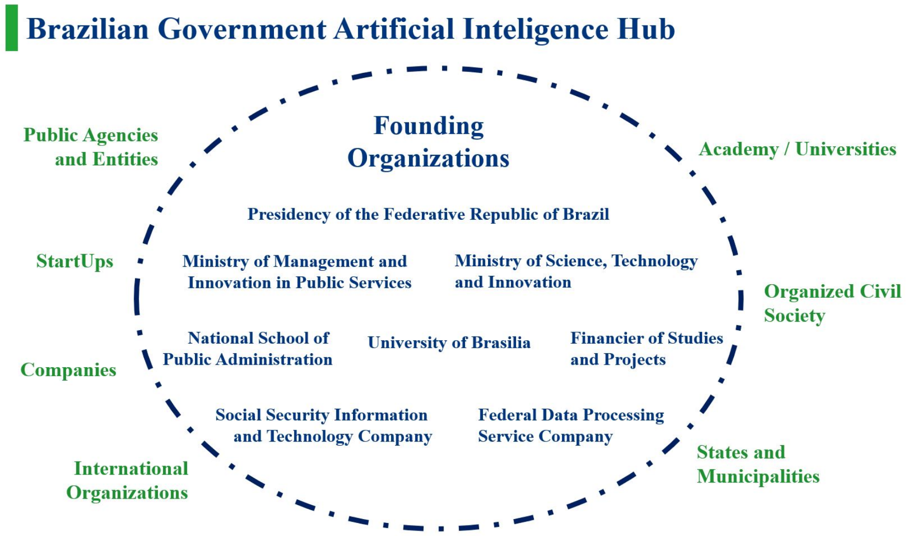

Figure 1: Brazilian Government Artificial Intelligence Hub (NIA).

There is a knowledge gap in academic research concerning AI's influence on public values and the governance challenges that arise from this interaction (Chen et al., 2023; Kitsios et al., 2023; Maragno et al., 2023; Neumann et al., 2024). Insights into how government employees perceive AI's impact on public values and governance challenges remain limited. The debate over whether AI will transform the public sector and its operational environments continues (Kitsios et al., 2023).

The rising interest in AI within government spans various scholarly fields. While this diversity is enriching, it can cause fragmentation in theoretical and terminological approaches as researchers from different disciplines work independently (Straub et al., 2023). Technical research often neglects broader perspectives and real-world contexts. Additionally, the increasing number of typologies and taxonomies to interpret AI's role in government sometimes needs to be more interconnected, reinforcing a fragmented knowledge base (Straub et al., 2023).

Despite growing debate, AI diffusion in the public sector remains limited compared to the private sector (Neumann et al., 2024). Interdisciplinary AI scholarship in government is still emerging, and theory has yet to offer significant value to practitioners (Birhane et al., 2022). Humans remain crucial as decision-makers and rule enforcers, including public safety and fraud investigations (Chen et al., 2023). The abundance of conceptual terms and lack of integration between social sciences and AI's technical aspects leave governments uncertain about which principles to prioritize in AI governance (Straub et al., 2023). Globally, AI in government is still in early implementation, with key decisions shaping future expectations (Straub et al., 2023).

2.2 AI in Public Service in Brazil

Brazil has advanced its digital agenda since the 1990s, starting with the 1991 Information Technology Program, which offered tax incentives for R&D in hardware and automation and promoted company-university collaboration (Filgueiras and Junquilho, 2023). In 2018, the Brazilian Strategy for Digital Transformation (E-Digital) (BRASIL, 2018) outlined actions for advancing public policies, emphasizing AI, big data, 5G, and modernizing production. It also highlighted the need for ICT training and assessing the impacts of disruptive technologies. In 2021, the Brazilian Strategy for Artificial Intelligence (EBIA) (BRASIL, 2021) was introduced, focusing on AI research, public security, legislation, and international governance (Anna, 2023). EBIA promotes innovation by fostering knowledge exchange through councils and networks (Nonato et al., 2024).

Brazil has made significant strides in innovation potential (Anna, 2023), with growing social interest in enhancing public sector innovation to improve efficiency, quality, and accessibility (Guedes and Júnior, 2024). The Brazilian government has established AI centers of excellence, funded university research, created collaborative networks, and developed ecosystems while ensuring privacy and data protection (Filgueiras and Junquilho, 2023). The public sector leads in policy formulation, promotes data transparency, fosters innovation, supports professional training, and invests in emerging technologies to build effective governance ecosystems (BRASIL, 2021).

Despite various initiatives to promote digital development, Brazil's efforts need coherence and alignment with its social reality (Filgueiras and Junquilho, 2023). Key challenges hindering EBIA's implementation include a clear timeline for strategic goals (Anna, 2023) and unclear governance structures (Gaspar and Mendonça, 2021). As emphasized in the public consultation, regulatory bodies, specific authorities, and existing structures are needed. The lack of policy integration leads to non-cooperation among actors, making EBIA more of an ineffective document than a collaborative strategy for problem-solving (Gaspar and Mendonça, 2021).

Several AI initiatives have been identified in Brazil's public agencies. The Conecte SUS project aims to modernize public health management (Lemes and Lemos, 2020), while the EB S@úde Project uses AI to optimize military health databases (Lemes and Lemos, 2020). Regulatory sandboxes are advanced in the financial sector, with institutions like the Ministry of Economy and the Central Bank of Brazil implementing models since 2019 (Kubota and Rosa, 2024). The Brazilian Federal Revenue Service uses SISAM for real-time processing of over 2.5 million import declarations (Guedes and Júnior, 2024). Various agencies, including the Federal Court of Accounts and the Federal Police, employ AI robots to boost productivity (Desordi and Della Bona, 2020). Additional AI use cases are documented in an OECD.AI inventory (OECD.AI, 2022).

Although broad social participation contributed to the national AI strategy, it has yet to be fully implemented nationwide. To address this, the Federal Government established the AI Hub (NIA), tasked with creating a list of Priority AI Projects, defining projects eligible for funding, and promoting academic studies to solve public sector challenges. The project discussed in this paper is an initiative of the NIA SubHub for Prospecting Strategic Projects, executed by the University of Brasília (UnB) through LAMFO under a Decentralized Execution Agreement (DEA). This SubHub comprises professionals from various fields, as described in Figure 1.

3 EMPOWERING SERVICES: AI FRAMEWORK

A standardized methodology is necessary for implementing AI in public services. To meet the goals of the National Artificial Intelligence Strategy, this paper proposes a comprehensive process for deploying AI solutions. The following sections outline steps for identifying demands, selecting projects, designing solutions, prototyping, and validation. This methodology was developed through a literature review, highlighting key findings and limitations of AI in public administration. By exploring best practices and emerging technologies, we applied the concept of Meta AI—learning to learn about AI—through thorough applied research to achieve the project's objectives.

To address the challenge, the proposed methodology aimed to answer the following question: **What is the best method for selecting AI projects in government with subjective issues to be judged and more than ten alternative options?** Initially, a literature review was conducted on methods for selecting AI projects in government with subjective judgment criteria, highlighting Multicriteria Decision Analysis (MCDA) methods, particularly the Analytic Hierarchy Process (AHP) (Saaty, 1977), due to its widespread use for project selection. AHP Express was selected for cases with many alternatives (Leal, 2020).

In addition to the method for selecting AI projects, after the evaluation was completed, Lean Office practices were incorporated to improve public administration services and reduce bureaucratic overhead. Two literature reviews were conducted to advance this research: one focused on digitalization and the other on Lean practices.

This resulted in the development of the AI solution selection and prototyping process for public services, which was designed based on the proposed methods in the literature and enriched through discussions among the AI scientists, who are the authors of this paper, and members of the SGD, the institution responsible for promoting AI project development in the Brazilian federal government. Consequently, a methodology was proposed that reflects cutting-edge knowledge and addresses the specific and complex peculiarities of public administration in a country as large as Brazil. The steps of the proposed methodology are illustrated in Figure 2, with each phase's objective described below.

Step 1: Survey. Conduct a comprehensive survey of federal public administration agencies' AI needs. This step investigates departmental challenges and identifies areas where AI can enhance public service efficiency and effectiveness. The research should be thorough, incorporating interviews, questionnaires, and analysis of existing data.

Step 2: Elimination. Following the application of the Survey (Step 1), various project proposals for applying AI are identified. At this step, an initial filter is used to ensure that the proposed projects can indeed be addressed through AI. This filter eliminates

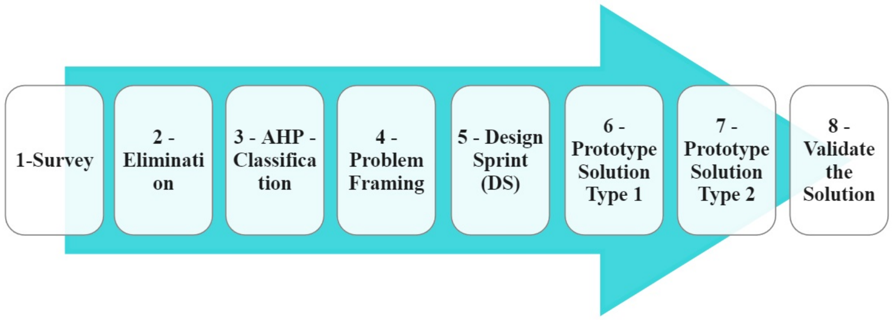

Figure 2: Steps of the Project Selection Process for Prototyping.

proposals failing to meet the established criteria, ensuring that only viable projects proceed to the next phase.

Step 3: AHP - Classifier. The proposals filtered in the elimination step proceed to prioritization. In this step, the proposals are ranked in order of priority and preference, considering the subjective aspects of stakeholders, such as expected impact, urgency, and feasibility. The AHP is utilized here to structure and facilitate decision-making, ensuring that the most promising and strategic projects are selected.

Step 4: Problem Framing. In the proposed approach presented in this paper, Problem Framing was used - for (Bardwell, 1991), Problem Framing helps structure and provide information organization to understand the problem, which people do not do daily. Hence, it helps create a data management and analysis environment that aims to solve problems. A lack of understanding of the problem can lead to solutions that do not bring the desired effects (Robertson, 2016).

Step 5: Design Sprint. Afterward, an adapted version of the Design Sprint was used (Knapp et al., 2016). A study (Mendonça de Sá Araújo et al., 2019) mentions that the computer systems industry has widely adopted Design Sprint to minimize the problems in defining requirements in software projects. Another study (Larusdottir et al., 2021) highlights the focus on adapting the design of the product, usually a computational application, to the users' needs and adds that, in the Design Sprint proposal, a relatively small team, which initially has a superficial idea of the users' requirements, intends to, at the end of the process, have a prototype solution.

Step 6: Prototyping the Solution - Type 1. Aiming at the prototyping of the solution selected in Step 5, a theoretical and summarized prototype of the solution is conceptualized. This conceptualization occurs through collaborative brainstorming sessions between the project researchers and the team from the SGD. In this step, the existing computational and scientific methods in the literature that are most suitable for implementing the solution are identified and discussed. The definition of requirements and the initial system architecture are outlined.

Step 7: Prototyping the Solution - Type 2. The researchers, specialists in artificial intelligence, implement the computational framework conceptualized in Step 6 as a Minimum Viable Product (MVP). The MVP is a functional version of the solution with basic functionalities sufficient for initial tests and validations. This step includes selecting and identifying databases, developing AI algorithms, system integration, and preliminary testing.

4 DETAILED EXECUTION: PHASE GUIDE

A detailed guide on the applicability of the methodology defined in Section 3 is presented below.

4.1 Survey

The development of the survey was based on the academic experience of the project team, as no specific survey was identified in the literature. At the grantor's request, a preamble to the survey was prepared, with guidelines for the application and use of AI models in public administration. The preamble includes content on the premises for preparing guidelines, the definition of generative AI, applications of generative AI in government, and the limitations of generative AI and large language models (LLMs). This survey was carefully designed to accommodate varying levels of AI maturity in agencies, aiming to collect specific knowledge about public sector processes related to problems and needs that can be framed as strategic projects and addressed with AI-based solutions.

4.2 AHP Express

The Analytic Hierarchy Process (AHP) is a multicriteria decision-making method that uses hierarchical or network structures to represent a decision problem and develops priorities for the alternatives based on the decision maker's judgments (Saaty, 1987). AHP should be understood as a facilitator, a thought structuring process, rather than an algorithm that solves problems (Colin, 2007). According to (Colin, 2007), the application of the AHP comprises four phases:
1. Hierarchy representation: involves the development of the decision hierarchy that relates the various levels of interrelated elements. **2. Pairwise comparisons:** includes evaluating preferences for each decision element at a given hierarchy level. **3. Eigenvector method:** encompasses using the eigenvector method to estimate the relative weights of decision elements at a particular level and assess the consistency of preferences established in pairwise comparisons. **4. Aggregation of priorities:** concerns the aggregation of relative priorities to evaluate the outcome related to the objective.

In addition to the formula for calculating the inconsistency of judgments and the maximum acceptable level, (Saaty, 1977) indicates that when there are comparisons involving more than nine elements, the expectation is that consistency will be low. In line with the proposal of (Saaty, 2006), Step 2 - Elimination was devised by the authors of this paper for cases involving more than nine alternatives.

The step-by-step application of the AHP Express method is detailed in (Leal, 2020). It follows the principles of traditional AHP, whose framework can be seen in (Saaty, 1987; Colin, 2007). The main difference is that, as the AHP Express method assumes judgment consistency, there is no need to calculate inconsistency. From the judgment of the preferred alternative, obtaining the values of the other pairwise comparisons through the relationship established in the initial judgment line is possible. The AHP Express method was chosen for being suitable for the agile project management model.

4.2.1 Elimination

This phase was developed to address two main objectives: firstly, to eliminate AI project proposals identified during the Survey phase that are deemed infeasible, and secondly, to classify the remaining proposals according to their maturity levels in descending order. A set of criteria and sub-criteria were established to achieve these goals, comprising a suitable model for an AI solution. AI specialists, including the authors of this study, initially formulated these criteria and subsequently refined them through input from stakeholders within the NIA. Notable contributors included leaders from SGD/MGI, MCTI, Enap, Finep, Serpro, Dataprev and other entities.

Table 1 (Appendix) outlines the specific criteria and sub-criteria utilized to assess the feasibility of each project proposal for AI implementation. These criteria encompass technical feasibility, data requirements, proposed solutions, societal impact, and strategic alignment. Possible answers to the questions are:

YES NO I DON'T KNOW

- If any proposal receives a "NO" response for any sub-criteria, it is automatically eliminated from the process.
- "YES" responses are scored as "1".
- "I DON'T KNOW" responses are scored as "0".
- All proposals must pass through this elimination phase.
- At the end, the scores of the proposals will be summed-up and placed in descending order.

Initially, the nine highest-scoring projects will be considered for the AHP-Classifier step, as these indicate higher maturity levels. If there is a tie among the top nine, the criteria weights can be adjusted based on stakeholder input. It is recommended that weights be adjusted only in the case of a tie, with the most relevant criterion assigned a weight of 2.

4.2.2 AHP Express Classifier

The criteria used in the AHP-Express method to classify/select AI projects from each Government area are described below. Although only the criteria were scored, since the methodology needed to be adapted for agile project management, they were described through sub-criteria presented to stakeholders during the AHP application interviews.
1. Urgency of Solution: Time to solve the problem without AI: How long will it take without AI? Impact of delay: What are the consequences of not addressing the problem? Future urgency: Will the problem become more urgent?
2. Availability of Data Sources: Internal data quality: Are internal data reliable? Cost of external data: How much will external data acquisition cost? Time to acquire external data: How long will it take?
3. Broader Population Reach: People impacted: How many people will be affected? Geographical reach: Will the AI solution be available nationwide? Accessibility: Will it be accessible across income levels and digital skills?

4. Greater User Value: User benefits: What advantages does the AI provide? User satisfaction: Is it user-friendly and does it meet needs? Time saved: How much time will users save?
5. Projected Access Volume - Cost Estimate: Estimate per period: What is the estimated access volume? Growth curve: How is access expected to grow? Seasonality: Is demand seasonal?
6. Lower Technical Effort: Problem complexity: Does the problem need advanced AI techniques? Tools and libraries: Are there tools available for development? Expertise: Is there AI expertise? Replicability: Can the solution be used for other problems? Maintenance: Is maintenance challenging?
7. Ease of Implementation: System integration: Does the solution integrate with existing systems? Regulation: Is it compliant with laws?
8. Risks, Ethics, and LGPD: Are risks mapped? Are there ethical concerns? How does LGPD impact the solution?

Meetings with process stakeholders are essential for applying AHP Express. It is recommended to hold three meetings for each segment (Area, Department, Ministry): one with expert researchers, one with segment members, and one with the SGD group, each consisting of two to four members.

The process starts with presenting up to nine project proposals from Step 2 - Elimination, the eight criteria, the Saaty Scoring Table, and a brief explanation of the dynamics. Participants first assess the importance of each criterion in response to the question "Which AI project should be selected?" and then evaluate each project based on these criteria. AHP Express uses the Saaty scale (1-9), where one indicates equal preference, and nine indicates extreme preference.

4.3 Problem Framing

To carry out this step, a meeting was held with the individuals most directly involved with the identified and prioritized problem, including people from various areas since the problem is cross-functional. These participants were instructed beforehand to identify the causes and characteristics of the problem, enabling a fruitful conversation and ensuring everyone understood its context and relevant aspects. Considering the project's objective, the availability of data is essential. Databases with relevant data were analyzed to illustrate the problem, and the currently implemented solution was examined to identify strengths and areas for improvement.

4.4 Design Sprint

Three of the five traditional Design Sprint steps were carried out in two days, with a multidisciplinary team (the Squad) - made up of people involved in the problem: Understand/Map, Sketch, and Decide. In "Understand/Map", the long-term objectives were defined, the questions established for the Sprint were answered, the user map was created, interviews with experts were carried out, and the target user was defined. In "Sketch", there was a division into two groups, which analyzed all available information and came up with proposals – the entire group analyzed this in "Decide", arriving at a solution definition.

4.5 Prototyping

This step pertains to the mapping and development of prototypes considering the agile approach in a demonstration environment and the documentation of AI solutions obtained in Step 5. The experimentation through prototyping aims to evaluate the feasibility of an MVP solution. The requirements and specifications will be developed closely with the requesting entity during the design sprint phase. The Type 2 prototype, conceived in Step 6, is delivered through Python code, developed in a controlled environment that simulates the user's real-world problem. This prototype accompanies a detailed technical report and a presentation for the main stakeholders and interested parties.

4.6 Validation

Validating AI solutions in the public sector demands a rigorous approach to ensure effectiveness and efficiency. This paper describes the validation process of an MVP prototype for an interactive chatbot providing information and virtual addresses for public agency services. MVP validation for AI applications, like chatbots, requires a systematic process that combines functionality tests, user evaluations in controlled settings, and real-world deployments. Robust development, testing, and monitoring tools are essential for successful validation and final implementation, ensuring the solution effectively meets user needs.

The process begins by defining the chatbot's objectives and success criteria, measured by metrics such as correct response rate, interaction time, and user feedback. Internal tests validate functionality, including conversation flow and database integration. This is followed by user testing in controlled environments, gathering qualitative and quantitative feedback for improvements. Adjustments are made before large-scale deployment. The chatbot is then beta-tested in a real environment, where performance is monitored through continuous user feedback. After validation, a crucial monitoring plan is established to ensure the chatbot adapts to new demands and updates, maintaining its effectiveness.

Based on its specifications, the prototype's user interface is also designed for evaluation. This structured approach ensures AI solutions meet technical and operational standards and are validated under near-real conditions, providing a foundation for large-scale implementations.

5 RESULTS

This study proposes an innovative methodology for improving Brazilian public services using AI, as outlined earlier. Developed in collaboration with SGD members, AI researchers, the authors, and NIA collaborators, this methodology follows the steps detailed in Section 3. Below are the results from each step of its application.

Initially, discussions between project executors and stakeholders determined that the Executive Secretaries of the Ministries would select the agencies for the survey. To ensure the survey was comprehensive and precise, we meticulously crafted 16 open-ended questions, focusing on the Service Context, Problem Characterization, AI Needs, and AI Solution.

The survey was conducted in the Department of Digital Public Service Platforms (DESER) of SGD, and three project proposals were raised. Subsequently, four people from the General Coordination of Automation and Service (CGATE), two AI professors, a facilitator professor, and SGD staff carried out the elimination step, in which all three projects passed the filter. After the filter, the process proceeded to the AHP Express. Three groups of respondents were interviewed for the application of AHP Express: 1) professors specialized in AI, 2) SGD public servants, and 3) public servants from the originating area of the proposal/demand. The respondent teams consisted of 2 to 4 participants. Each stakeholder from each group had their response recorded during an interview, in which the questions contained in Table 1 were asked. In the AHP Classification, the three projects selected in Step 2 were analyzed:

- Automated Recovery of the Gov.br Account access
- CGATE citizen support channels
- Citizen support channel for public services on the Gov.BR Portal

These projects focus on the need for high citizen interaction through digital platforms, particularly the Gov.br Portal. Consequently, the classification process results highlight the importance of impacting a large portion of the population. The three groups evaluated the projects based on the criteria in Section 4.2.2. The final AHP Express results, detailed in Figure 3, represent the geometric means of the three scores for each criterion. The project "Citizen service channel for public services on the Gov.br Portal" was selected for the prototyping phase.

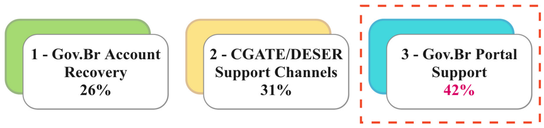

Figure 3: AHP Express Result.

The Problem Framing and Design Sprint steps were conducted as described in Sections 4.3 and 4.4. As a result, after the Problem Framing of Problem 3, the following answers were reached for Problem 3:

- **Where do we want to be in 1 year?** Increase public service engagement rate by 100% through Gov.Br.
- **What is the product?** Generative AI chatbot.
- **What does the product do?** Assists and answers citizens' questions about the public service they need and directs them to the service page.
- **What problem is the team trying to solve with the product?** Better direct citizens to the service they are looking for.
- **Who currently uses the product, and who would we like to use it?** Citizens (Gov.br users).
- **What is the best phrase to communicate the long-term goal?** Reduce requests to support channels: Gov account support and ombudsman by 25%.

The prototyping will involve developing an Advanced Generative Agent to interact with citizens on the Gov.br Portal, assisting, clarifying doubts, and directing them to the appropriate service pages. Prototyping will occur in two steps, as outlined in Section 4.5. Below are the results of the strategy and architecture developed by AI experts and stakeholders for this generative agent:

Data Collection and Structuring: Identify and organize data communication. Attach and restructure additional information. Summarize documents. Propose a FAQ document.

Document Processing: Extract text, topics, keywords, and structure using document intelligence tools (e.g., Azure).

Vector Database: Generate semantic chunking with layout. Create embeddings. Organize the database by themes and problems. Define indexes in the vector database.

Prompt Embedding: Define settings for the generative agent to communicate inclusively with polite, accessible language tailored to different user profiles. Specify topics the agent won't respond to, such as political matters or prejudices. Provide examples for In-context Learning, like step-by-step guides for public services, responses about losing Gov.br access, processes needing human intervention, and complaint instructions. Enable human overflow by offering support links, operating hours, and contact info. Ensure system status is informed to users before starting a conversation.

Advanced Semantic Search: To enhance the generative agent's precision, a specific LLM, LLaMA, trained on a Portuguese-language database, will be used. This model will optimize Retrieval-Augmented Generation with Fine-Tuning (RAFT) (Zhang and et al., 2024), combining supervised fine-tuning with RAG (Lewis and et al., 2020). RAFT adjusts training data for better question responses using Chain-of-Thought reasoning from specific documents, enhancing RAG by integrating external database retrieval with pre-trained models.

The proposed RAFT solution, which combines supervised fine-tuning with RAG, introduces an innovative approach to improving the accuracy of interactions between generative AI and citizens. Validating a chat intelligent agent with an LLM includes testing its functionality, performance, user experience, domain-specific knowledge, ethical considerations, benchmark comparisons, and ongoing monitoring to ensure it meets requirements and provides satisfactory interactions (Xi et al., 2023). At the time of writing, Step 8 is under development, as described in Section 4.6.

6 DISCUSSION

This paper presents a comprehensive methodology for deploying AI solutions in public services, incorporating Lean Office waste elimination methods, process optimization, and standardization before evaluating workflow automation. Once automated, a continuous improvement (kaizen) loop optimizes the NLP chatbot's workflow. Digitization enables the automation of standardized processes, eliminating waste and embedding intelligence through smarter control algorithms. Key questions addressed include the link between operational excellence in public administration and AI in workflows and the driving forces behind improved performance and efficiency in public services through AI.

Beyond portal access benefits, the Brazilian government can gain additional advantages. Using Webometrics (Saeidnia et al., 2024), interactions between citizens and the generative agent will produce valuable data. Authorities can leverage this to:

- Identify patterns in citizen behavior, needs, and preferences.
- Understand popular services and common complaints.
- Create data-driven public policies.
- Personalize services and improve efficiency.
- Monitor service trends and conversion rates for successful demand resolutions.
- Enhance metrics for "Gov.br" portal usage.

Data privacy and security are crucial when applying AI in public administration. The information collected covers healthcare, smart cities, and public service forecasting at varying levels of maturity. These challenges emphasize the need for robust safeguards to protect sensitive data and ensure adherence to data protection regulations.

Other projects have been selected using the proposed methodology across different agencies and secretariats. To harmonize the selection process at a national level, one option is to apply uniform weights to all criteria for each project. This would standardize the selection process, ensuring that each criterion is equally evaluated from a national perspective rather than from an individual secretariat's viewpoint. Consequently, the process would better align with the broader objective: identifying which AI project should be prioritized for Brazil as a whole.

7 CONCLUSION AND FUTURE WORK

This paper presented an innovative methodology for implementing AI solutions in Brazilian public services, aimed at enhancing efficiency, accessibility, and service quality. Integrating Lean Office, Design Sprint, and advanced AI techniques, the methodology is both flexible and adaptable to diverse governmental needs, ensuring inclusivity. Collaboration among AI experts, public servants, and stakeholders was key to ensuring technical feasibility and alignment with public policies.

The methodology addresses a specific demand from the SGD, part of the National Digital Government Strategy. Given Brazil's numerous federal agencies, it includes the selection and prototyping of projects within a complex government structure. Its flexibility allows application in other countries with extensive governmental frameworks, promoting digital transformation in varied contexts.

The development of a generative agent for the Gov.br Portal meets the demand and transforms citizen interactions into valuable data. This data provides insights for personalizing services. We propose using LLaMA X, trained on Portuguese documents, with RAFT for improved accuracy. Future work includes developing a ChatGPT-based AI for comparison and exploring integrations with programs like Bolsa Família and systems such as CPF (Social Security Number), CNPJ (National Registry of Legal Entities), SEI (Electronic System of Information), and institutions payroll.

ACKNOWLEDGEMENTS

Acknowledgments and recognition are extended to the Ministry of Management and Innovation in Public Services, especially the Secretaria de Governo Digital (SGD), for financially supporting this project and enabling the practical feasibility of solution identification and prototyping.

REFERENCES

Anna, U. (2023). Artificial intelligence landscape in south america. *AARMS–Academic and Applied Research in Military and Public Management Science*, 22(1):101–114.

Atabekov, A. (2023). Artificial intelligence in contemporary societies: legal status and definition, implementation in public sector across various countries. *Social Sciences*, 12(3):178.

Bardwell, L. V. (1991). Problem-framing: A perspective on environmental problem-solving. *Environmental Management*, 15(5):603–612.

Birhane, A., Kalluri, P., Card, D., Agnew, W., Dotan, R., and Bao, M. (2022). The values encoded in machine learning research. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, pages 173–184.

Blount, K. (2024). Using artificial intelligence to prevent crime: implications for due process and criminal justice. *AI & SOCIETY*, 39(1):359–368.

BRASIL (2023). Decreto Nº 11.437, de 17 de março de 2023. https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2023/decreto/D11437.htm#art5. [Online; accessed 27-June-2024].

BRASIL, M. d. C. T. e. I. (2018). *Estratégia Brasileira para a Transformação Digital*. https://www.gov.br/mcti/pt-br/centrais-de-conteudo/comunicados-mcti/estrategia-digital-brasileira/estrategiadigital.pdf.

BRASIL, M. d. G. e. I. e. S. P. (2024). Serviços e informações do brasil. https://www.gov.br/pt-br. Accessed: 27 June 2024.

BRASIL, Ministério da Ciência, T. e. I. (2021). *Estratégia brasileira de inteligência artificial: Ebia*. https://bit.ly/3C6LgY4.

Chen, T., Gascó-Hernandez, M., and Esteve, M. (2024). The adoption and implementation of artificial intelligence chatbots in public organizations: Evidence from us state governments. *The American Review of Public Administration*, 54(3):255–270.

Chen, Y.-C., Ahn, M. J., and Wang, Y.-F. (2023). Artificial intelligence and public values: value impacts and governance in the public sector. *Sustainability*, 15(6):4796.

Cohen, S., Manes Rossi, F., and Brusca, I. (2023). Debate: Public sector accounting education and artificial intelligence. *Public Money & Management*, 43(7):725–726.

Colin, E. C. (2007). *Pesquisa operacional: 170 aplicações em estratégia, finanças, logística, produção, marketing e vendas*. Livros Técnicos e Científicos.

Desordi, D. and Della Bona, C. (2020). A inteligência artificial e a eficiência na administração pública. *Revista de Direito*, 12(02):01–22.

Dwivedi, Y. K., Sharma, A., Rana, N. P., Giannakis, M., Goel, P., and Dutot, V. (2023). Evolution of artificial intelligence research in technological forecasting and social change: Research topics, trends, and future directions. *Technological Forecasting and Social Change*, 192:122579.

Engstrom, D. F., Ho, D. E., Sharkey, C. M., and Cuéllar, M.-F. (2020). Government by algorithm: Artificial intelligence in federal administrative agencies. *NYU School of Law, Public Law Research Paper*, (20-54).

Fan, Z., Yan, Z., and Wen, S. (2023). Deep learning and artificial intelligence in sustainability: a review of sdgs, renewable energy, and environmental health. *Sustainability*, 15(18):13493.

Faria, A. V. A., de Melo, M. K., de Oliveira, F. A. R., Weigang, L., and Celestino, V. R. R. (2022). Automated slr with a few labeled papers and a fair workload metric. In *International Conference on Web Information Systems and Technologies*, pages 1–23. Springer.

Filgueiras, F. and Junquilho, T. A. (2023). The brazilian (non) perspective on national strategy for artificial intelligence. *Discover Artificial Intelligence*, 3(1):7.

Gaspar, W. B. and Mendonça, Y. C. d. (2021). A inteligência artificial no brasil ainda precisa de uma estratégia.

Gov, A. (2024). Gov.br é a página de governo mais acessada do mundo.

https://agenciagov.ebc.com.br/noticias/202403/gov-br-e-a-pagina-de-governo-mais-acessada-do-mundo. Accessed: 27 June 2024.

Guedes, L. and Júnior, M. O. (2024). Artificial intelligence adoption in public organizations: a case study. *Future Studies Research Journal: Trends and Strategies*, 16(1):e860–e860.

Jevinger, Å., Zhao, C., Persson, J. A., and Davidsson, P. (2024). Artificial intelligence for improving public transport: a mapping study. *Public Transport*, 16(1):99–158.

Kitsios, F., Kamariotou, M., and Mavromatis, A. (2023). Drivers and outcomes of digital transformation: The case of public sector services. *Information*, 14(1):43.

Knapp, J., Zeratsky, J., and Kowitz, B. (2016). *Sprint: o método usado no Google para testar e aplicar novas ideias em apenas cinco dias*. Editora Intrínseca Ltda.

Kubota, L. C. and Rosa, M. B. (2024). Inteligência artificial no brasil: Adoção, produção científica e regulamentação. digitalização e tecnologias da informação e comunicação. *Digitalização e tecnologias da informação e comunicação: oportunidades e desafios para o Brasil. Rio de Janeiro: Ipea*.

Kutyauripo, I., Rushambwa, M., and Chiwazi, L. (2023). Artificial intelligence applications in the agrifood sectors. *Journal of Agriculture and Food Research*, 11:100502.

Larusdottir, M., Roto, V., and Cajander, A. (2021). Introduction to user-centred design sprint. In Ardito, C. e. a., editor, *Human-Computer Interaction – INTERACT 2021*, volume 12936 of *Lecture Notes in Computer Science*, pages 283–293. Springer, Cham.

Leal, J. E. (2020). Ahp-express: A simplified version of the analytical hierarchy process method. *MethodsX*, 7.

Lemes, M. M. and Lemos, A. N. L. E. (2020). O uso da inteligência artificial na saúde pela administração pública brasileira. *Cadernos Ibero-Americanos de Direito Sanitário*, 9(3):166–182.

Lewis, P. and et al. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. In *Advances in Neural Information Processing Systems*, volume 33, pages 9459–9474.

Maragno, G., Tangi, L., Gastaldi, L., and Benedetti, M. (2023). Exploring the factors, affordances and constraints outlining the implementation of artificial intelligence in public sector organizations. *International Journal of Information Management*, 73:102686.

Mendonça de Sá Araújo, C. M., Miranda Santos, I., Dias Canedo, E., and Favacho de Araújo, A. P. (2019). Design thinking versus design sprint: A comparative study. In Marcus, A. and Wang, W., editors, *Design, User Experience, and Usability. Design Philosophy and Theory*, volume 11583 of *Lecture Notes in Computer Science*, pages 335–344. Springer, Cham.

Neumann, O., Guirguis, K., and Steiner, R. (2024). Exploring artificial intelligence adoption in public organizations: a comparative case study. *Public Management Review*, 26(1):114–141.

Nonato, S. D. O., de Souza Figueiredo, S. d. S., de Melo Alves, C. A., and Dias, C. N. (2024). Cooperação interorganizacional em planos estratégicos de inteligência artificial: uma análise comparativa. *Future Studies Research Journal: Trends and Strategies*, 16(1):e838–e838.

OECD.AI (2022). Ai use case in the public sector. Retrieved on 29-Jun-2024.

Robertson, S. I. (2016). *Problem Solving: Perspectives from Cognition and Neuroscience*. Psychology Press, 2nd edition.

Rocha, C. A. A., Weigang, L., Dib, M. V. P., Faria, A. V. A., Cajueiro, D. O., de Melo, M. K., and Celestino, V. R. R. (2022). Leveraging transfer learning for long text classification with limited data. In *International Conference on Web Information Systems and Technologies*, pages 98–120. Springer.

Saaty, R. W. (1987). The analytic hierarchy process—what it is and how it is used. *Mathematical Modelling*, 9(5):161–176.

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15.

Saaty, T. L. (2006). Rank from comparisons and from ratings in the analytic hierarchy/network processes. *European Journal of Operational Research*, 168(2):557–570.

Saeidnia, H. R., Hosseini, E., Abdoli, S., and Ausloos, M. (2024). Unleashing the power of ai: a systematic review of cutting-edge techniques in ai-enhanced scientometrics, webometrics and bibliometrics. *Library Hi Tech*.

Samoili, S., Cobo, M. L., Gómez, E., De Prato, G., Martínez-Plumed, F., and Delipetrev, B. (2020). Ai watch. defining artificial intelligence. towards an operational definition and taxonomy of artificial intelligence.

Scutella, M., Plewa, C., and Reaiche, C. (2024). Virtual agents in the public service: examining citizens' value-in-use. *Public Management Review*, 26(1):73–88.

Straub, V. J., Morgan, D., Bright, J., and Margetts, H. (2023). Artificial intelligence in government: Concepts, standards, and a unified framework. *Government Information Quarterly*, 40(4):101881.

Wang, C., Han, B., Patel, B., and Rudin, C. (2023). In pursuit of interpretable, fair and accurate machine learning for criminal recidivism prediction. *Journal of Quantitative Criminology*, 39(2):519–581.

Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., Zhang, M., Wang, J., Jin, S., Zhou, E., et al. (2023). The rise and potential of large language model based agents: A survey. *arXiv preprint arXiv:2309.07864*.

Yu, K.-H., Beam, A. L., and Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature biomedical engineering*, 2(10):719–731.

Zhang, T. and et al. (2024). Raft: Adapting language model to domain specific rag. *arXiv preprint arXiv:2403.10131*.

APPENDIX

Table 1: Criteria and Subcriteria for Feasibility of AI Implementation in Public Services.

Criteria	Subcriteria
1- Is There Technical Feasibility?	a. Cost: The AI solution's development, implementation, and maintenance are economically viable. b. Time: The solution can be implemented within the PBIA timeline; c. Maintenance: The AI solution is feasible to maintain and update; d. Operationalization: Necessary resources for implementation (human, material, infrastructure) are available;
2- Will the data not be a limiting factor for the model?	a. Data Quality: Some data patterns can be learned to help solve the problem; b. Data Quality: The available data tend to be sufficiently accurate, complete, and reliable once preprocessed so as not to impede the progress of the project; c. Data Quantity: There is sufficient data to train and test the AI solution. d. Data Diversity: The data significantly represent the target population's characteristics.
3- Can AI improve the solution?	a. Effort: The current solution to the problem is unsatisfactory, and the AI-based solution will increase process efficiency. b. Effort: The solution cannot be found using simple rules (e.g., through spreadsheets, If..else) or basic statistical analysis. c. Effort: There is a need for prediction, classification, or the use of language models (like ChatGPT) based on input data. d. Measurability: The outcome of the process and/or service is measurable.
4- Will there be a good Impact on society?	a. Benefits: The AI solution will bring real benefits to society, either directly or indirectly through improved management; b. Threats: The AI solution considers and mitigates negative impacts on society; c. Acceptance: The AI solution will be accepted by society; d. Regulation: The AI solution complies with existing laws and regulations.
5- Is it in line with Strategic Alignment?	a. Alignment with the organisation's mission: The AI solution aligns with the organisation's mission; b. Strategic priorities: The AI solution contributes to the organisation's strategic priorities; c. Reputation risk: The AI solution will not harm the organisation's reputation; d. Governance: The solution contributes to advancing specific public policies and government-established goals.