Internet Sectoral Overview | Number 1 April, 2026 Year 18

Data governance in the era of Artificial Intelligence

Artificial Intelligence in the Federal Executive Branch: Governance, capacity building, and digital sovereignty

Renan Mendes Gaya Lopes dos Santos,[1] *Thaciana Guimarães de Oliveira Cerqueira,*[2] *Camila Falchetto Romero,*[3] *and Alex Lopes Pereira*[4]

Introduction

Artificial Intelligence (AI) is a technological force capable of altering governmental processes and positively impacting public policies and services provided to society. As highlighted by the Organisation for Economic Co-operation and Development (OECD), governments can use AI to "design better policies and make better decisions, improve communication and engagement with citizens and residents, and improve the speed and quality of public services" (Berryhill et al., 2019, p. 3). In the context of the Brazilian State, the formulation of guidelines for this digital transformation finds its reference point in the Brazilian Artificial Intelligence Plan (PBIA),[5] a public policy instrument designed to guide the use of technology to promote the improvement of people's lives, social inclusion, and national technological sovereignty. The increasing adoption of these technologies requires systematic observation, given that they permeate a complex ecosystem in which gains in technical efficiency must be balanced with ethical responsibility and legal compliance.

NIA

The Artificial Intelligence Center (NIA)[6] is an initiative coordinated by the Digital Government Secretariat (SGD) of the Ministry of Management

[1] Postgraduate in software engineering and director of data structuring and AI at SGD/MGI.
[2] PhD in AI and general coordinator of management of the AI Center at SGD/MGI.
[3] General coordinator of AI governance at the Directorate of Data Structuring and Artificial Intelligence of SGD/MGI.
[4] PhD in electronic and computer engineering and advisor to the Directorate of Data Structuring and Artificial Intelligence of SGD/MGI.
[5] Find out more: https://www.gov.br/lncc/pt-br/assuntos/noticias/ultimas-noticias-1/plano-brasileiro-de-inteligencia-artificial-pbia-2024-2028
[6] Find out more: https://www.gov.br/governodigital/pt-br/infraestrutura-nacional-de-dados/inteligencia-artificial-1

(...) the NIA's mission is to transform public administration through the intelligent application of AI technologies, contributing to efficiency, service personalization, and the generation of public value.

and Innovation in Public Services (MGI) to promote technological advancement with responsibility and accelerate the adoption of technologies in the public sector (Figure 1). It is one of the actions of the National Data Infrastructure (IND),[7] established by Decree 12.198/2024, which constitutes a set of rules, policies, architectures, standards, technological tools, and information assets, aimed at promoting the strategic use of data held by organizations and entities of the Federal Executive Branch.

Figure 1 – NIA

What is it?

An initiative coordinated by the SGD that works in partnership with various institutions for the collaborative development of AI initiatives in the public sector.

What is its purpose?

To promote and accelerate the use of AI in the public sector ethically and responsibly.

Brazilian Artificial Intelligence Plan

Integrated into Axis 3 of the PBIA, the NIA has five goals, with the mission of transforming public administration through the intelligent application of AI technologies and contributing to:

- administrative efficiency
- personalizing public services
- generating public value

Source: Prepared by the authors.

Integrated into Axis 3 of the PBIA, the NIA's mission is to transform public administration through the intelligent application of AI technologies, contributing to efficiency, service personalization, and the generation of public value. The overall objective of Axis 3 of the PBIA is clear and ambitious: to make Brazil a global model of efficiency and innovation in the use of AI in the public sector, developing solutions that significantly improve the supply of services and people's satisfaction with them, and with an impact on development and social inclusion (Ministry of Science, Technology, and Innovation [MCTI] & Center for Management and Strategic Studies [CGEE], 2025).

The NIA operates as a coordination ecosystem, working with multiple partners to leverage each institution's existing and specialized skills (Figure 2).

[7] Find out more: https://www.gov.br/governodigital/pt-br/infraestrutura-nacional-de-dados

Figure 2 – NIA PARTNERS

Source: Prepared by the authors.

To ensure governmental cohesion, the Chief of Staff Office (Casa Civil) supports political and institutional coordination among the ministries. The technical and scientific basis, as well as the provision of some financial resources, is coordinated by the MCTI, with the support of the Financing Agency for Studies and Projects (Finep) in promoting and financing research and development. Data infrastructure and nationwide technological implementation are enabled by the expertise of the Federal Data Processing Service (Serpro) and the Social Security Technology and Information Company (Dataprev), while the National School of Public Administration (Enap) plays a central role in training public servants, ensuring that technological advancement is accompanied by human development, in addition to acting through its Artificial Intelligence Innovation Laboratory (LIIA),[8] as a central hub for prospecting, designing, and experimenting with ethical solutions focused on innovation and solving complex public problems.

Furthermore, the NIA works in partnership with the University of Brasília (UnB), to carry out experiments on projects involving AI; the Regional Center for Studies on the Development of the Information Society (Cetic.br), for advisory services regarding research on the adoption of AI in the Self-Diagnosis of the Information Technology Resource Management System (Sistema de Administração dos Recursos de Tecnologia da Informação [SISP])[9] and also regarding its publication in the Brazilian Observatory of Artificial Intelligence (OBIA), an interactive platform that aggregates sectoral indicators to facilitate access by society and researchers to statistics in the state's AI ecosystem; with the World Bank, to support for the development of strategies and systems for ethical impact assessment; and the Center for Research

NIA works in partnership with (...) the Regional Center for Studies on the Development of the Information Society (Cetic.br), for advisory services regarding research on the adoption of AI in the Self-Diagnosis of the Information Technology Resource Management System (Sistema de Administração dos Recursos de Tecnologia da Informação [SISP]) and also regarding its publication in the Brazilian Observatory of Artificial Intelligence (OBIA) (...).

[8] Find out more: https://www.enap.gov.br/inovacao/liia/
[9] Find out more: https://www.gov.br/governodigital/pt-br/estrategias-e-governanca-digital/sisp

and Development in Telecommunications (CPQD), for research and development on the use of AI in the public sector.

NIA's work is organized across different complementary fronts that aim to act, in a systemic and responsible manner, to accelerate the adoption of AI in the public sector. These actions include the development of centralized platforms, project prospecting, experimentation in controlled environments, systemic monitoring, training of state employees, and the promotion of the ethical and responsible use of AI, encompassing everything from digital literacy to the development of regulatory instruments, as well as the identification and mitigation of social and technological risks through its own government self-assessment framework (Box 1).

Box 1 – SISP SELF-DIAGNOSIS

Figure 3 - IMPLEMENTED SOLUTIONS

Implemented solutions (2025)	Organizations with implemented solutions (2025)
182	58
vs. **117** in 2024	vs. **41** in 2024

Chart 1 - TOTAL NUMBER OF FEDERAL GOVERNMENT ORGANIZATIONS, BY BARRIER OR CHALLENGE IN DEVELOPING APPLICATIONS INVOLVING AI (2025)

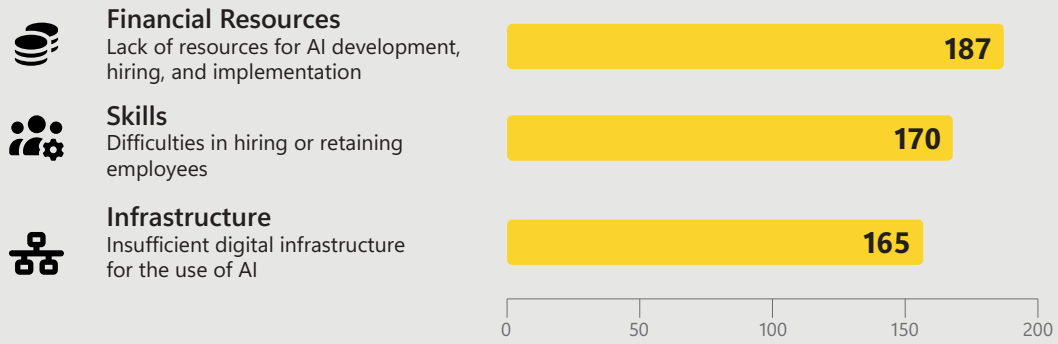
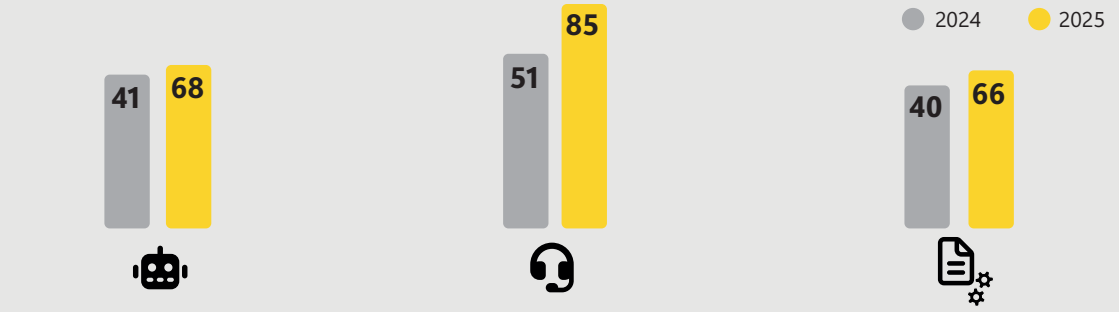


Chart 2 - TOTAL NUMBER OF FEDERAL PUBLIC AGENCIES, BY CATEGORIES OF AI TOOLS USED OR UNDER IMPLEMENTATION (2024 AND 2025)



CONTINUES ▶

Chart 3 - TOTAL NUMBER OF FEDERAL GOVERNMENT ORGANIZATIONS, BY AI COMPETENCIES MOST CRITICAL OR IN SHORTEST SUPPLY (2025)

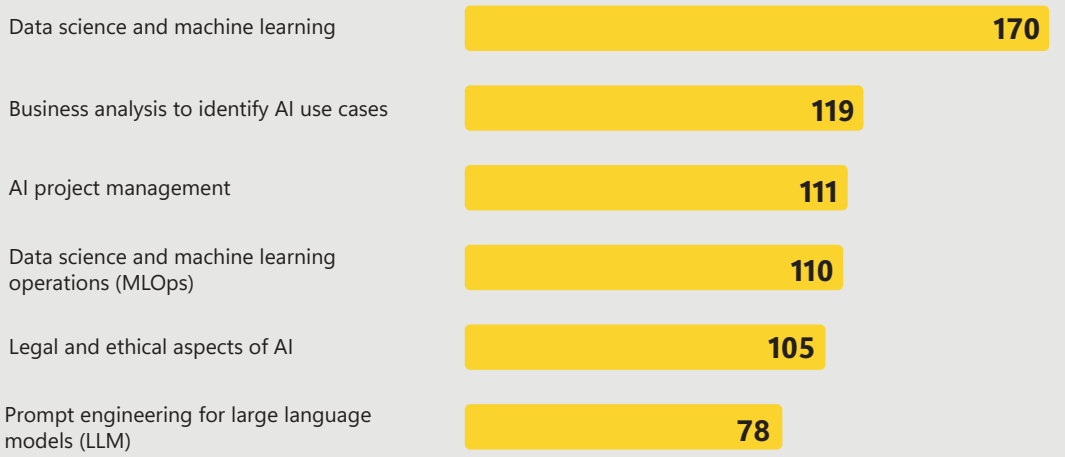
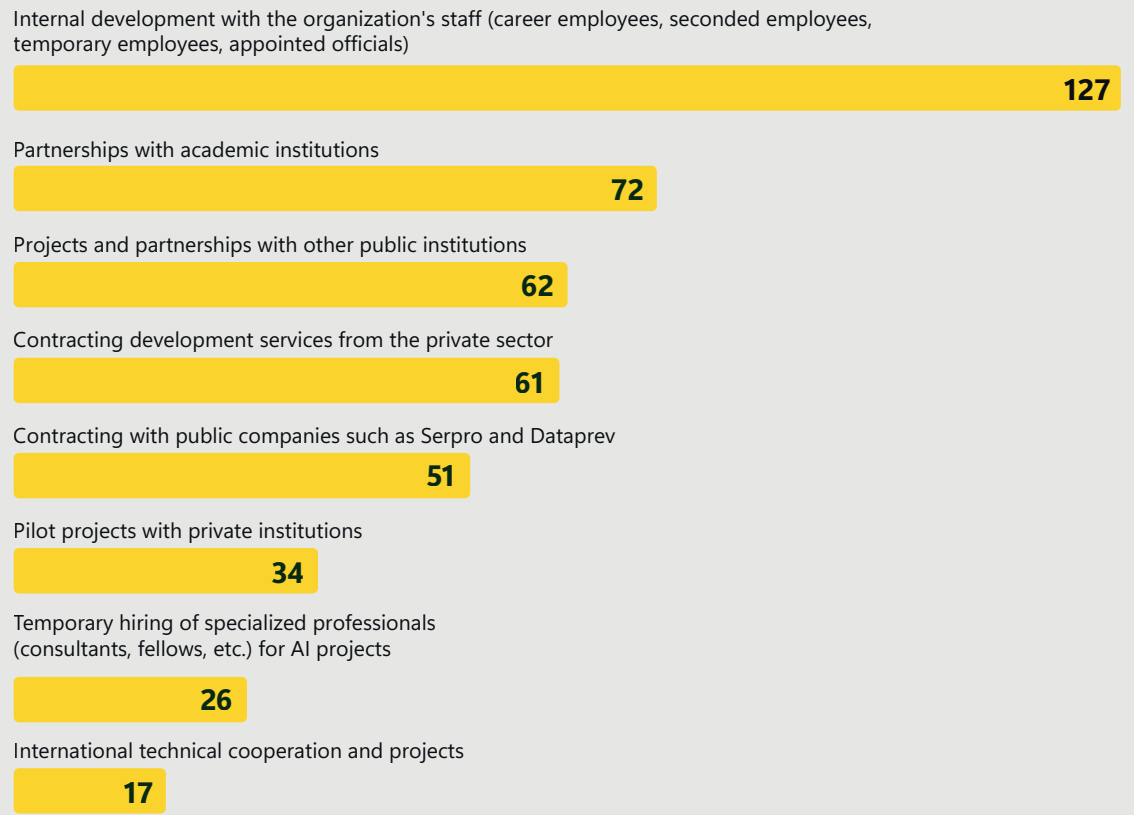


Chart 4 - TOTAL NUMBER OF FEDERAL GOVERNMENT ORGANIZATIONS, BY TYPE OF HUMAN RESOURCES OR SUPPLIER MANAGEMENT STRATEGY USED TO DEVELOP OR CONTRACT AI SOLUTIONS TAILORED TO THE ORGANIZATION'S NEEDS (2025)



Source: Prepared by the authors, based on SGD/MGI (2026).

The data from the SISP Self-Diagnosis, applied in November 2025 (based on responses from 234 Federal Executive organizations), detail the internal priorities and stages of technological maturity (...).

Monitoring the adoption of AI in the Federal Executive Branch

Formulating and correcting the course in public technology policies requires monitoring indicators. The Brazilian scenario for AI adoption in the public sector is measured using two tools: the Self-Diagnosis from SISP, established by Ordinance SGD/MGI 4.339/2023, an annual mandatory reporting instrument managed by the SGD; and the independent ICT Government[10] survey, conducted by Cetic.br, with a methodology based on samples and censuses.

The sixth edition of the ICT Electronic Government[11] survey indicated that the adoption of AI applications in federal and state government organizations increased from 24% in 2021 (NIC.br, 2022) to 30% in 2023 (NIC.br, 2024). The survey revealed a difference in institutional capacity between the spheres and branches of government: while 49% of federal organizations already use AI, the rate among state organizations is 28% (NIC.br, 2024). The judiciary accounts for the largest volume of initiatives, with 68% of its organizations reporting the use of AI (NIC.br, 2024), driven by procedural triage systems and jurisprudential analysis. For context, within civil society and the market, the ICT Households 2025 survey (NIC.br, 2025a) recorded that 32% of Internet users (approximately 50 million people) have already used generative AI tools, although there is strong socioeconomic inequality (69% in class A versus 16% in classes D/E), while the ICT Enterprises 2024 survey (NIC.br, 2025b) indicated that only 13% of Brazilian companies with 10 or more employed persons use AI solutions.

The data from the SISP Self-Diagnosis, applied in November 2025 (based on responses from 234 Federal Executive organizations), detail the internal priorities and stages of technological maturity (SGD/MGI, 2026). The survey identified 341 projects in the prospecting phase, 144 in pilot project format, and 182 operational or implemented systems (SGD/MGI, 2026). The most frequently reported motivation for investing in technology was the expectation of faster delivery of public services, as indicated by 181 organizations. In addition to this agility, the reduction in the cost of providing services (cited by 114 organizations) and the development of new public services (101 organizations) are the most anticipated impacts for the federal administration (SGD/MGI, 2026).

In terms of practical application, the core functional areas of the institutions lead in receiving these technological solutions (74 organizations) (SGD/MGI, 2026). A well-established example in this category is the Super Sapiens system, from the Attorney General's Office (AGU), which uses AI for the automated screening of cases and assistance in drafting legal documents. The second focus of adoption is the management of administrative processes (49 organizations) (SGD/MGI, 2026), illustrated by initiatives such as MentorIA, a tool launched by MGI to optimize the routine of public servants in planning public purchases and contracts. Finally, there are direct services to citizens (40 organizations).

[10] Find out more: https://www.cetic.br/en/pesquisa/governo/

[11] Available at: https://www.cetic.br/en/publicacao/pesquisa-sobre-o-uso-das-tecnologias-de-informacao-e-comunicacao-no-setor-publico-brasileiro-tic-governo-eletronico-2023/

Obstacles to AI adoption

The process of modernizing the state involves operational and human resource obstacles inherent to public administration. An analysis of data from the SISP Self-Diagnosis reveals that the main barrier to the development of AI projects is the restriction of financial resources, reported as a challenge by 187 federal organizations (SGD/MGI, 2026). Although facing budgetary limitations and challenges caused by the complexity of acquiring external solutions, the predominant adoption strategy in the government has been internalization: 127 organizations indicated internal development, with career, temporary, or commissioned civil servants as the main way to create or customize solutions (SGD/MGI, 2026).

This internal development strategy highlights a second bottleneck documented by the ICT Government[12] survey regarding the SISP Self-Diagnosis: the shortage of qualified professionals. The federal survey recorded that 170 organizations face difficulties in hiring or retaining employees with technology skills (SGD/MGI, 2026). The skill identified as the most critical and in greatest short supply in government was data science and machine learning, cited by 170 organizations (SGD/MGI, 2026). Another relevant technical limitation mentioned was the digital infrastructure, which was deemed insufficient for the use of AI by 165 governmental institutions (SGD/MGI, 2026).

In order to overcome these obstacles identified in the SISP Self-Diagnosis, the NIA, together with its institutional partners, adopted several initiatives that will be detailed below.

An analysis of data from the SISP Self-Diagnosis reveals that the main barrier to the development of AI projects is the restriction of financial resources, reported as a challenge by 187 federal organizations (...).

Experiments with AI solutions and national startups

The adoption of AI in the public sector involves risks that could compromise the generation of public value if not mitigated. A study by Gartner indicates that at least 50% of projects of generative AI are abandoned after the proof of concept (the initial phase of a project) due to data quality problems, escalating costs, and governance weaknesses (Chandrasekaran, 2026). To reduce this risk of failure, before a government organization invests large amounts of human and financial resources in AI solutions, it is recommended to conduct experiments aimed at testing, on a small scale, the real positive impact of adopting AI technology, assessing possible negative consequences, and analyzing the feasibility and sustainability of the solution.

With a focus on reducing risks and aiming to foster the national AI market, the MCTI, in partnership with the NIA, the LIIA, and Finep, has promoted, since 2023, public calls for national startups to develop AI solutions aimed at solving problems in the public sector. The calls for proposals, divided into three rounds, resulted in 25 projects funded through economic subsidies, totaling approximately R$ 80 million in investments from the National Fund for Scientific and Technological Development (FNDCT).

[12] Available at: https://www.cetic.br/en/tics/governo/2023/orgaos/H3B/

(...) the NIA partnered with the UnB to conduct more AI experiments, in which human-in-the-loop prototypes and complete architectural documentation were developed (...).

The first round involved organizations, such as the National Regulatory Agency for Private Health Insurance and Plans (ANS), the Brazilian Health Regulatory Agency (Anvisa), and the Ministry of Agriculture and Livestock (Mapa), with projects focused on a semiautomated analysis of complaints and authorization documents, as well as virtual customer service assistants (chatbots); the second round incorporated the Federal Revenue Service and the Hospital de Clínicas de Porto Alegre, focusing on tax risks and hospital optimization; the third round included the Securities and Exchange Commission (CVM), the National Social Security Institute (INSS), and the National Institute of Metrology, Quality and Technology (Inmetro), with projects related to assistants for regulatory impact analysis and investments.[13]

It is noteworthy that the legacy of these funding rounds goes beyond the systems delivered: by financing national companies through economic subsidies, the projects strengthen competencies in natural language processing, computer vision, and data analysis, affirming that digital sovereignty arises from the collaboration between government, funding agencies, and the national market.

Furthermore, the NIA partnered with the UnB to conduct more AI experiments, in which human-in-the-loop prototypes and complete architectural documentation were developed, encompassing the following:

- **Chatbot for Gov.br account support (SGD):** digital support via Retrieval-Augmented Generation (RAG) architecture.[14]
- **Gov.br text assistant (SGD):** text review and simplification, with a multi-agent approach.
- **Applications assistant (Secretariat of Federal Assets [Secretaria de Patrimônio da União – SPU]):** summarization and legal classification of administrative processes.
- **Coordination of flows and supervision procedures assistant (CPROC) and certification of charitable social assistance entities (CEBAS) from the Ministry of Education (MEC):** document screening for higher education supervision and certifications.

Training

To address the staffing shortage, training was established as a central goal in the PBIA, which aims to train 115,000 federal public servants by the end of 2026 (MCTI & CGEE, 2025). Through a partnership between the NIA, Enap, and Serpro, the digital literacy program aims to encompass technical fundamentals, governance, and ethical use. According to the 2025 SISP Self-Diagnosis (SGD/MGI, 2026), 136 organizations encourage the participation of civil servants in government school courses, with a primary focus on information technology (IT) teams, cited by 130 organizations.

[13] The documentation for the rounds is available on Finep's official website: http://www.finep.gov.br/chamadas-publicas/chamadapublica/705 (round 1), http://www.finep.gov.br/chamadas-publicas/chamadapublica/718 (round 2), and http://www.finep.gov.br/chamadas-publicas/chamadapublica/735 (round 3).

[14] An AI technique that optimizes LLM by connecting them to external databases and documents. Before formulating a response, the system retrieves specific and up-to-date information from these sources, which increases accuracy and reduces "hallucinations" (false information) without requiring model retraining.

The training strategy is organized into learning paths designed for five distinct profiles of public servants: public officials, who learn to use AI tools in their daily work; public managers, who work on process optimization and decision-making; information and communication technology (ICT) managers, who deepen their knowledge of data engineering and infrastructure; data executives, who develop skills for evidence-based decision-making and legal compliance; and senior leaders, who are trained for the strategic use of AI in public policy formulation (Figure 4). Figure 5 presents the training programs offered for each profile. For everyday use, the government has published guidelines, in the form of guides, that teach AI research techniques and the understanding and responsible use of generative AI tools. By the beginning of 2026, more than 82,000 training sessions had been conducted as part of these initiatives, reaching a total of 28,000 public servants.[15]

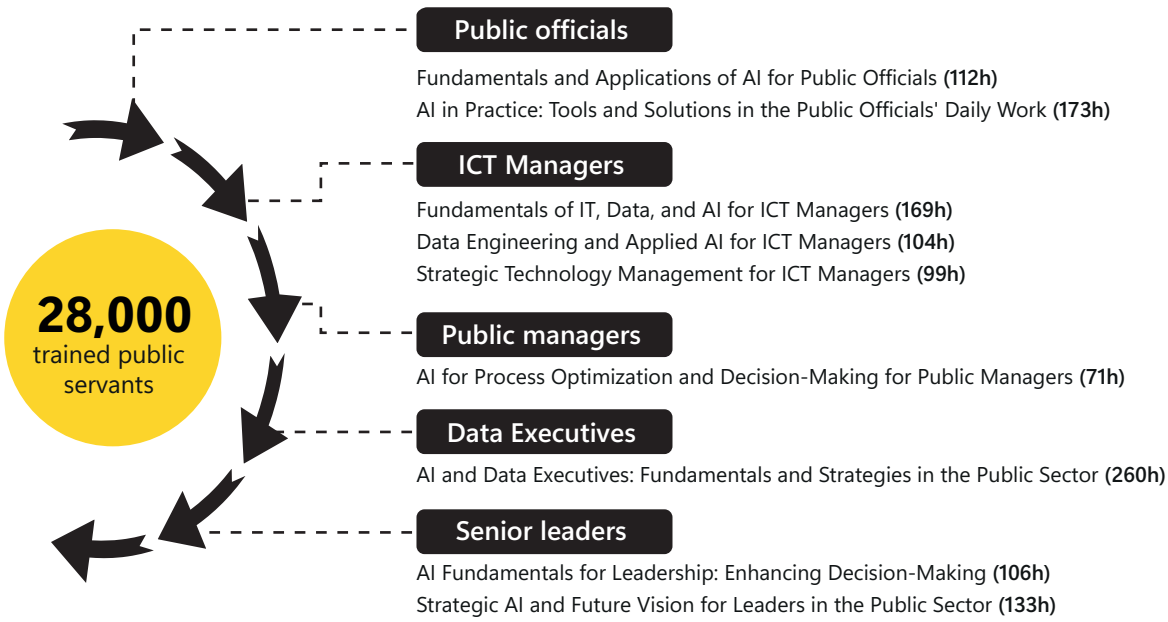
By the beginning of 2026, more than 82,000 training sessions had been conducted as part of these initiatives, reaching a total of 28,000 public servants.

Figure 4 – CAPACITY BUILDING AXIS: SKILLS DEVELOPMENT STRATEGY

Source: Prepared by the authors.

[15] The courses are available at: gov.br/ind/amplia

Figure 5 – CAPACITY-BUILDING AXIS: PROGRAMS AND COURSES FOR SKILLS DEVELOPMENT, BY TARGET AUDIENCE



Source: Prepared by the authors.

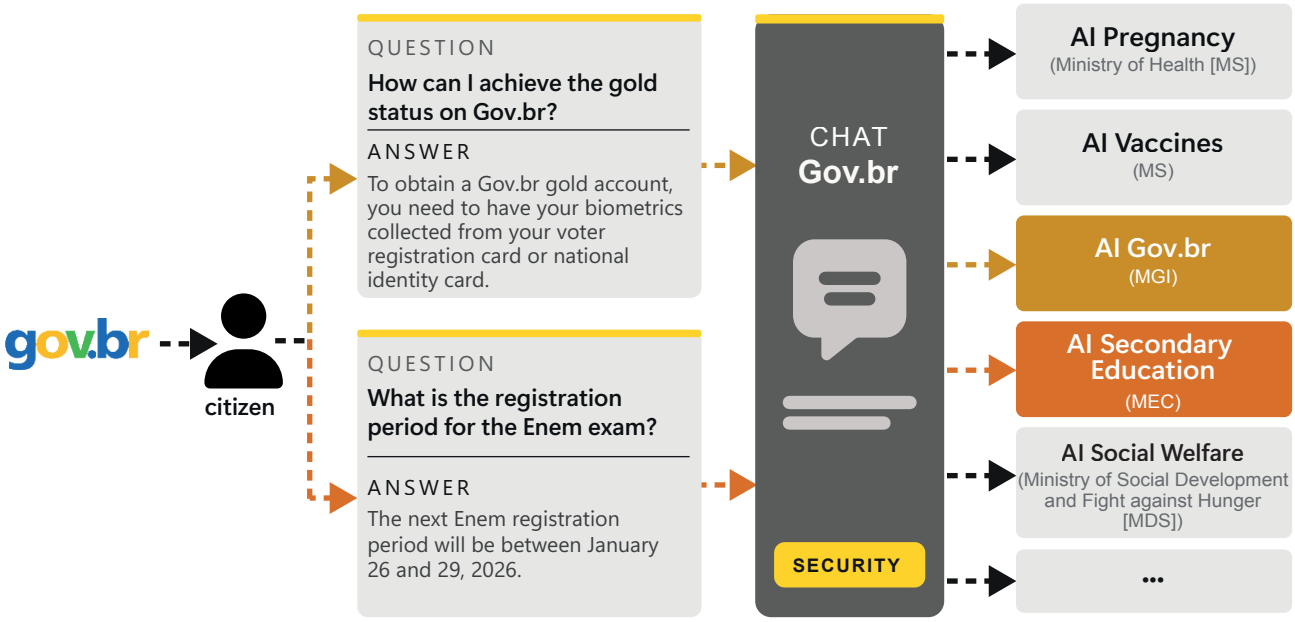
AI Platform

The SISP 2025 Self-Diagnosis survey (SGD/MGI, 2026) revealed a significant structural challenge: 165 organizations reported that their current digital infrastructure was insufficient to support the implementation and continued use of new AI-based technologies. In response to this limitation, the NIA established a strategy of providing a centralized, commonly used platform, not only to accelerate the adoption of the technology by the state but also to mitigate the redundancy of technical and financial efforts. Centralization optimizes the time each organization would otherwise spend on architectural decisions, fragmented bidding processes, and the development of repetitive modules, ensuring, from the outset, the standardized application of security and monitoring mechanisms.

Given the vast universe of technologies encompassed by AI, the federal government needed to establish development priorities guided by the real demands of the administration. The Self-Diagnosis indicated that the most desired tools in the short term (one to two years) are chatbots focused on citizen service, cited by 144 organizations (SGD/MGI, 2026), followed by the provision of LLM to assist in the internal routines of civil servants. Considering this strong demand for language processing and the governance gaps exposed by the diagnosis, the strategic decision converged on prioritizing generative AI.

In implementing this guideline, the NIA and Dataprev launched Chat Gov.br in January 2026. The solution centralizes the creation of virtual assistants for public policies of the Federal Executive Branch, offering citizens a unified and intuitive channel to interact with the government. In this initial phase, the service operates exclusively through WhatsApp, and its expansion will occur gradually. The first use case, developed with the MEC, focuses on students taking the National High School Exam (Enem) who use the Unified Selection System (Sisu), the Student Financing Fund (Fies), and the University for All Program (Prouni). The journey is proactive: the citizen who is part of the target audience receives a notification in the app containing a call to action to start the conversation. Upon accepting, they interact with an assistant powered by a RAG knowledge base, capable of answering precise questions about the programs. Meanwhile, the assistant for supporting the Gov.br account itself is already in the pilot phase (Figure 6).

Figure 6 – AI PLATFORM (CHAT GOV.BR)



Source: Prepared by the authors.

From a technical standpoint, the architecture under development envisions a model-agnostic routing layer, allowing for the flexible combination of different foundational models, be they LLM, small language models (SLM), or multimodal models. The system relies heavily on advanced RAG techniques. This means that each specialized chat consults an official and restricted government database to formulate its responses, ensuring accuracy and mitigating the risk of algorithmic bias. To meet governance demands, the platform integrates security guardrails, human-in-the-loop mechanisms for the continuous evolution of applications, and

(...) the NIA developed the Self-Assessment framework for the Ethical Impact of AI in the Public Sector (AIE), in alignment with international benchmarks, such as the OECD and the United Nations Educational, Scientific and Cultural Organization (UNESCO).

rigorous protocols for observability, quality assessment, and compliance with the Brazilian General Data Protection Law (LGPD).[16] Complementing managerial control, the MGI provides for the implementation of granular cost management mechanisms, which will allow for precise monitoring of AI operational expenses across the federal public administration.

Ethical and responsible use of AI

As explained, the Federal Executive Branch has been advancing in the adoption of AI through the development of solutions, the training of civil servants, and the provision of commonly used platforms. This progress is accompanied by an effort to structure ethical governance, for which the NIA developed the Self-Assessment framework for the Ethical Impact of AI in the Public Sector (AIE), in alignment with international benchmarks, such as the OECD and the United Nations Educational, Scientific and Cultural Organization (UNESCO).[17]

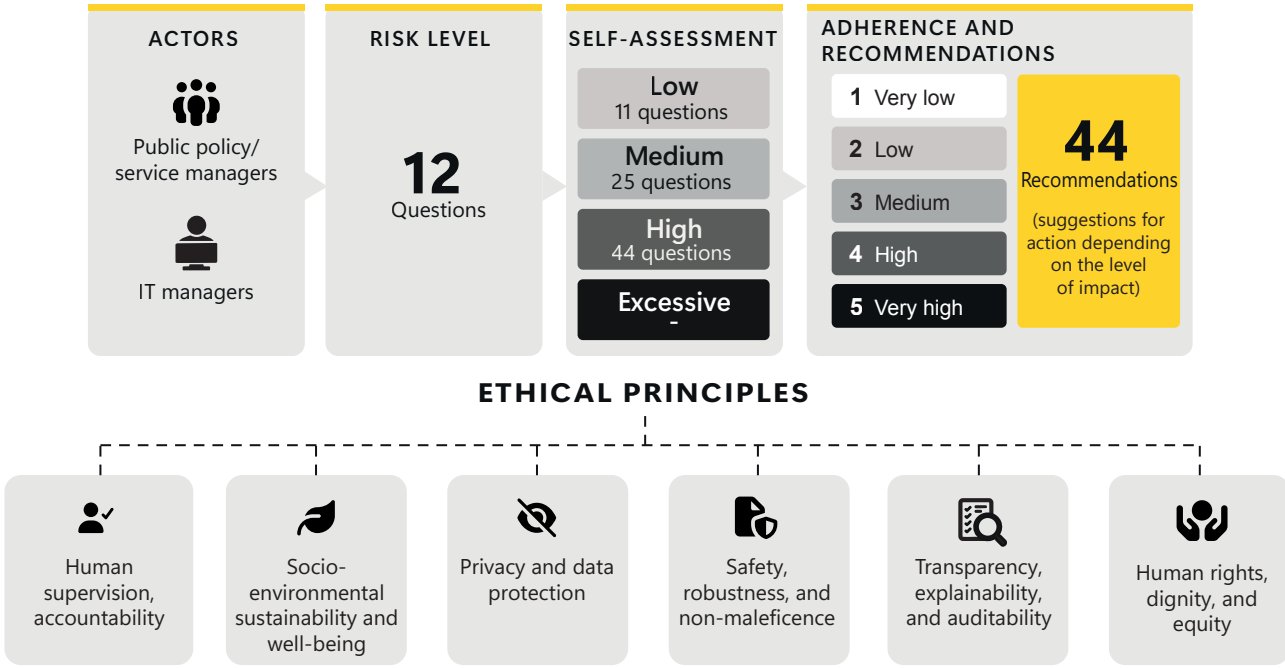
The SISP Self-Diagnosis of 2025 recorded progress in the ethical governance indicators tracked since 2024: the number of organizations with institutional guidelines or policies increased from 8 to 41, ethics committees were created in 23 organizations (compared to 6), and training on the subject reached the same level. Furthermore, the incorporation of ethical considerations into the lifecycle of solutions increased from 18 to 22 organizations, and specific measures of ethical AI behavior were recorded by 11 organizations, compared to 6 in 2024. Nevertheless, 118 organizations reported that there was no defined governance body for AI projects, and 214 indicated an absence of an institutionalized policy or strategy. To support organizations in structuring ethical governance, the NIA developed the AIE throughout 2025.

Based on the mapping of four risk categories (algorithmic bias and discrimination; technical and security risks; legal vulnerabilities; and social and reputational risks), the framework, whose main function is to guide teams in identifying and mitigating risks before they cause harm to citizens, aims to classify AI solutions into four risk levels (low, medium, high, or excessive) and is structured around six fundamental principles: human rights, transparency, human-in-the-loop, security, privacy, and sustainability.

The self-assessment completion process has two stages. The first is a 12-question questionnaire to classify the solution according to the indicated risk levels. In the second phase, the team answers the questionnaire in one of three versions (simplified, standard, or complete), according to the level of risk identified (solutions with excessive risk receive a recommendation for project reassessment). Finally, the framework generates a report with an ethical adherence score from one to five, charts by section, and up to 44 recommendations, functioning as a list of suggestions for improvement for the development team (Figure 7).

[16] Find out more: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm
[17] Find out more: gov.br/ind/amplia

Figure 7 – AXIS OF ETHICAL AND RESPONSIBLE USE OF AI (ETHICAL IMPACT SELF-ASSESSMENT FRAMEWORK)



Source: Prepared by the authors.

The instrument underwent validation in alpha phases (in partnership with Dataprev, Serpro, Enap, the MCTI, and the National Fund for the Development of Education [FNDE]) and beta phases (with MDS, the Comptroller General of the Union [CGU], the CVM, the National Telecommunications Agency [Anatel], and the Federal Revenue Service), in addition to receiving contributions from academia and organized civil society. Furthermore, in partnership with the World Bank, SGD has developed an open-source web application that will be available in the first half of 2026.

Vision for the future and the INSPIRE project

In addition to the products developed by the NIA with its partners over the past two years, there are new products and developments underway in the areas of data, security, and AI, which relate to (a) making available new educational resources for the retraining of the public sector; (b) developing and accelerating new AI solutions to improve and personalize public services; (c) making available new AI technology platforms that bring more sovereignty to the country; and (d) detailing procedural guidelines to increase adherence to the ethical principles of AI solutions.

The analysis of indicators collected by the SISP Self-Diagnosis, together with the statistics systematized by Cetic.br, demonstrates that the adoption of AI technologies by the Federal Executive Branch is already a reality.

To accelerate this evolution, the NIA, in partnership with the MCTI, launched the Artificial Intelligence in Public Service with Innovation, Responsibility, and Ethics (INSPIRE) project in December 2025. Conceived from the propositions of Axis 3 of the PBIA, the initiative integrates the national digital transformation strategy and results from a technological order using resources from the FNDCT, managed by Finep and executed by the CPQD. With a projected duration of four years, the program will have a total investment of R$ 390 million. Divided into six initiatives, INSPIRE aims to enhance, through research and development (R&D), the actions of the NIA and the IND:

1. **AI for data cataloging and interoperability:** to facilitate the integration, use, and reuse of information across the various interactions people have with the government;
2. **Platform for generative AI:** development of tools that enable the creation of generative AI applications to make people's lives easier in their interactions with public administration;
3. **AI applications center:** a centralized environment for experimentation and a repository of AI-based models, methodologies, and solutions to transform public services;
4. **Cybersecurity in the context of AI:** a hub of technologies to ensure the privacy and security of data used in AI applications, especially those involving sensitive data;
5. **AI for qualifying registration data:** infrastructure to validate, enrich, and ensure the accuracy of registration information, improving the identification of individuals, the personalization of services, and the effectiveness of public policies; and
6. **Personalized AI for disseminating public policies and services:** using semantic AI for cataloging, tagging, archiving, and thematically personalizing government information tracking.

Conclusion

The analysis of indicators collected by the SISP Self-Diagnosis, together with the statistics systematized by Cetic.br, demonstrates that the adoption of AI technologies by the Federal Executive Branch is already a reality. It is also possible to observe a continuous and accelerated increase in the maturity of organizations and entities over time, both in terms of the technologies used and their ethical and responsible use.

However, this scenario coexists with institutional disparities, a shortage of specialized professionals, and funding challenges. In order to further amplify the potential of technology to improve the delivery of public services, a set of coordinated actions is needed, combining the updating of technological infrastructure, the retraining of public servants, and the observance of fundamental rights.

In this direction, the products (already developed or under development) made available by the NIA of the SGD, built collaboratively with various partners, indicate the establishment of structuring bases and accelerators that have gradually increased the State's capacity to improve people's lives through more assertive, easily accessible, and agile public policies and services, expanding the exercise of citizenship in the country.

References

Chandrasekaran, A. (2026). *Why 50% of GenAI projects fail — And how to beat the odds.* Gartner. https://www.gartner.com/en/articles/genai-project-failure

Berryhill, J., Heang, K. K., Clogher, R. & McBride, K. (2019). Hello, World: Artificial intelligence and its use in the public sector. *OECD Working Papers on Public Governance*, 36, OECD Publishing. https://doi.org/10.1787/726fd39d-en

Brazilian Network Information Center. (2022). *H3 - Federal and state government organizations by use of Artificial Intelligence technologies in the last 12 months. Survey on the use of information and communication technologies in the Brazilian public sector: ICT Electronic Government Survey 2021* [Table]. https://www.cetic.br/pt/tics/governo/2021/orgaos/H3/

Brazilian Network Information Center. (2024). *H3 - Federal and state government organizations by use of Artificial Intelligence technologies in the last 12 months. Survey on the use of information and communication technologies in the Brazilian public sector: ICT Electronic Government Survey 2023* [Table]. https://www.cetic.br/pt/tics/governo/2023/orgaos/H3/

Brazilian Network Information Center. (2025a). *M1 – Internet users by use of generative Artificial Intelligence (AI) tools. Survey on the use of information and communication technologies in Brazilian households: ICT Households Survey 2025.* https://www.cetic.br/pt/tics/domicilios/2025/individuos/M1/

Brazilian Network Information Center. (2025b). H9 - Enterprises that used Artificial Intelligence technologies. *Survey on the use of information and communication technologies in Brazilian enterprises: ICT Enterprises Survey 2024.* https://www.cetic.br/pt/tics/pesquisa/2024/empresas/H9/

Decree No. 12.198, of September 24, 2024. (2024). Establishes the Federal Digital Government Strategy for the period 2024 to 2027 and the National Data Infrastructure, within the scope of the bodies and entities of the direct, autonomous, and foundational federal public administration. Official Gazette of the Union. https://www2.camara.leg.br/legin/fed/decret/2024/decreto-12198-24-setembro-2024-796286-publicacaooriginal-173095-pe.html

Digital Government Secretariat of the Ministry of Management and Innovation in Public Services. (2026). *Autodiagnóstico SISP 2025*. https://www.gov.br/governodigital/pt-br/estrategias-e-governanca-digital/sisp/autodiagnostico-igovsisp

Ministry of Science, Technology and Innovation, & Center for Management and Strategic Studies. (2025). *IA para o bem de todos - Plano Brasileiro de Inteligência Artificial*. https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/plano-brasileiro-de-inteligencia-artificial-pbia-_vf.pdf

SDG/MGI Ordinance no. 4.339, of August 10, 2023. (2023). Provides for the completion of the Self-Diagnosis within the Information Technology Resources Management System (SISP) and the Information and Communication Technology Governance Maturity Index (iGOVSISP). https://www.in.gov.br/en/web/dou/-/portaria-sgd/mgi-n-4.339-de-10-de-agosto-de-2023-502727051

Article II

Company neglect of data governance during transitions to embedded AI

By Daniel Insulza,[18] Zachary Nichols,[19] and Georgina Núñez[20]

Abstract

Many organizations are already rapidly integrating Artificial Intelligence (AI) into their business processes. The new goal of organizations is to generate a different way of operating through deeper AI integration. Often, this transition prioritizes model deployment and the speed of feature development over the fundamental controls that make AI secure, reliable, and auditable, i.e., data governance.

This article focuses on the current trend of underinvestment in data governance and explains why it is one of the main reasons AI projects fail, increase cyber risk, and hinder environmental, social, and governance (ESG) reporting as well as rule/regulatory compliance. It will also explain how AI offers a set of prescriptive governance, security, and organizational measures to solve the problems that arise.

AI implementation in organizations requires solid governance that guarantees responsible, ethical, and efficient use of data. Corporate governance establishes the rules, processes, and practices for decision-making and internal controls, ensuring that the company reaches its goals while also fulfilling its legal obligations.

AI data governance focuses on defining policies and mechanisms to manage the quality, privacy, and security of data used for smart systems. This requires establishing clear roles, sensitive information management guidelines, and mitigating risks associated with algorithm bias and improper use of information. Proper data governance promotes transparency, trust, and responsibility in the development and operation of AI-based solutions.

[18] Attorney licensed to practice law in the State of Colorado, United States. Certified in Sustainability Climate-Related Risks (SCR) by Global Association of Risk Professionals (GARP). Master's degree in compliance in fraud and laundering. Master's degree in innovation management and entrepreneurship. Former competitiveness and innovation consultant at the Inter-American Development Bank (IDB). Compliance consultant at the Development Bank of Latin America and the Caribbean (CAF).

[19] Attorney licensed to practice law in the State of Colorado, United States. https://www.linkedin.com/in/zachary-nichols-0465a773

[20] Economist, former economic affairs officer at the United Nations (UN) Economic Commission for Latin America and the Caribbean (ECLAC). Member of the Board of Directors of the Center for Regional Studies University of Tarapacá (Centro de Estudios Regionales Universidad de Tarapacá [CEUTA]). https://www.linkedin.com/in/georgina-n%C3%BA%C3%B1ez-reyes-20489a40

The main challenges presented by integrating AI into business strategies occur in corporate governance and the decision-making process. Integrating AI into corporate governance presents several pressing challenges (Díaz & Núñez, 2023). One major issue is the lack of clear regulatory frameworks to guide ethical AI deployment and risk management, leading to uncertainties about accountability when AI-driven decisions impact stakeholders. Another concern is transparency: AI systems often function as "black boxes," making it difficult for boards and executives to fully understand or audit their operations and outcomes.

Additionally, biases embedded in AI algorithms can perpetuate or even amplify existing inequalities, raising questions about fairness and compliance. There is also a growing need for specialized expertise among board members to effectively oversee AI strategies and mitigate potential risks. Finally, balancing innovation with privacy and data security remains a constant challenge, as corporations must navigate complex legal and ethical considerations in their pursuit of AI-driven growth. Leadership committed to ethics and best practices will undoubtedly contribute to leveraging the integration of AI into company business models.

The main challenges presented by integrating AI into business strategies occur in corporate governance and the decision-making process.

Why does data governance matter for AI?

AI systems depend on the availability of data and how data governance may affect their uses and management. The development in AI becomes even more complex considering the absence of clear data ownership, definitive origins, precise classifications, quality controls, access rules, and established lifecycle policies. AI may create organizations with fragile systems that are vulnerable to bias, data manipulation, privacy violations, and non-auditable decisions. Leading auditors and information technology (IT) sector studies have shown that organizations have identified these deficiencies. Many have pointed out a lack of data governance as one of the main obstacles to scaling AI responsibly within enterprises (Colombo & Flynn, 2025).

According to *2025 Readiness report: The readiness tipping point*, in 2024, the companies surveyed increased their investment in AI by 33% (Kyndryl, 2025). In part because of constant pressure from shareholders on company leadership to deliver advances in AI integration to their businesses, it is viewed as the next step to reach efficiency and optimization. However, at the same time, leaders admit their concern regarding their workforce, which they do not consider to be prepared for the changes that AI integration represents, and do not have the skills to maximize the impact of these new ways of operating. The lack of data governance within enterprises—which would manage how data is stored, processed, accessed, and protected—only makes the potential negative effects of increased embedded AI even worse (Kyndryl, 2025). Enterprises view data governance as a compliance exercise to fulfill regulatory requirements. However, data governance is a project success enabler. Data governance aids in ensuring data quality, improving decision-making, increasing operational efficiency, and accelerating enterprises' digital transformation.

AI public policy advances are limited, as are investments in these technologies. According to numbers from the AI Latin American Index (ILIA) 2025, Brazil and Costa Rica show significant increments in commitment to AI investment. Nevertheless, no country in the region is above the world average in terms of AI investment over GDP per capita (...).

Having clean and structured data, with a clear origin and flows, is of vital importance to AI, digital transformation, automation, and analytics. Otherwise, projects will underperform if it is not clear where the data originated from and whether it is accurate or even complete.

WHAT'S HAPPENING WITH GOVERNANCE IN LATIN AMERICA AND AI PENETRATION?

Latin America is not immune to this reality. According to the ECLAC, one of the main challenges in the region associated with the development of AI in Latin America and the Caribbean is the low institutional and governance capabilities, along with their low effectiveness.[21] Under these conditions, it becomes difficult to guarantee that the use of AI is ethical, responsible, and properly guided for the common good.

For Latin America, complexity also stems from the very advances in AI regulation, compared to other regions. The most advanced country in this aspect is Brazil, where a bill establishing a governance framework for Artificial Intelligence is currently under consideration,[22] emulating Europe's risk-based model (Whitecase, 2024).

Some countries in the region are more advanced than others, which only further complicates the situation, considering how common it is for data to be stored in servers located in another country, generating the need for homologation and cooperation that goes beyond borders. Initiatives such as the IDB's fAIr LAC+, which look to promote the responsible use of Artificial Intelligence and collaboration, are not adopted evenly. Therefore, the need for a regional framework remains (Access Now, 2024).

AI public policy advances are limited, as are investments in these technologies. According to numbers from the AI Latin American Index (ILIA) 2025,[23] Brazil and Costa Rica show significant increments in commitment to AI investment. Nevertheless, no country in the region is above the world average in terms of AI investment over GDP per capita, with the regional average being one-sixth of that threshold.

The region only represents 1.12% of the world's AI investment. The interoperability and availability of data in the region are slow, often prioritizing regulatory debates over the country's technological development. Progress in both regulatory frameworks and technological advances is needed and should be prioritized. However, one should not hinder the other, and the regulatory debate should look to benefit from private sector advances and experience.

[21] The Latin American Artificial Intelligence Index, launched in 2023, it is a joint effort between the ECLAC and the National Center for Artificial Intelligence (CENIA) of Chile, with the support of various academic, public, and private organizations. It aims to measure the progress of AI in 19 countries of the region. This initiative is part of ECLAC's Digital Development Observatory, which since 2024 has produced, compiled, and analyzed more than 85 indicators in 12 key areas for digital transformation in the region. Available at: https://repositorio.cepal.org/server/api/core/bitstreams/58abbd61-7c47-4208-8e8f-44bc1d14b894/content

[22] Available at: https://www25.senado.leg.br/web/atividade/materias/-/materia/157233

[23] Available at: https://www.cepal.org/es/publicaciones/82514-indice-latinoamericano-inteligencia-artificial-ilia-2025

Evidence of enterprises' carelessness in data governance in AI adoption: Industry surveys and reports

Recent research in the sector has shown a persistent trend; the increase in the adoption of AI is often not accompanied by an increase in the maturity of governance. Reports from the big four consulting firms (Deloitte, PricewaterhouseCoopers [PWC], Ernst & Young [EY], and KPMG) and other industry bodies have highlighted that unifying data governance and AI remains the exception rather than the rule, and that fragmented governance is an obstacle to the secure implementation of AI in various processes (KPMG+1). Further, a survey conducted by Precisely and Drexel University (2025) found that only 12% of the companies surveyed rate their data as of sufficient quality and accessibility for the adoption of AI. This is surprising, considering that, in the same survey, more than half of the respondents answered that they use AI for their data analysis. The survey highlighted that companies are moving forward with implementing AI solutions despite knowing that they lack effective data governance.

The causes of this disconnect seem to be associated with deficiencies in the operational management of AI. Some analysts have estimated that a significant portion of AI projects fail because of poor data management due to issues such as a lack of data traceability, inconsistent labeling, insecure pipelines, and ad hoc use of third-party models with unclear data provenance. These operational deficiencies translate directly into AI projects that fall short of goals or create new legal and security risks that demand attention (Bansal, 2025).

Further, technical vulnerabilities are linked to poor data governance. Researchers and experts warn that AI-specific attack vectors—data poisoning, model inversion, and code injection—are amplified when organizations lack formal policies for data collection, validation, and access control. Continuous learning systems and training datasets with lax governance are especially vulnerable (Innovative Routines International [IRI], 2025).

Recent research in the sector has shown a persistent trend; the increase in the adoption of AI is often not accompanied by an increase in the maturity of governance.

Why do enterprises neglect data governance?

Some factors that result in weak and inefficient governance are:

1. The contrast between pressure from business leaders for speed and product development that reward rapid feature delivery and data governance work, which is rather slow and "invisible."
2. Compartmentalized responsibilities also contribute to data isolation, as security, legal, compliance, and AI product leaders often operate separately with conflicting incentives.

(...) organizations still need unified roadmaps that combine data governance, AI, and cybersecurity (...).

3. The fragmentation and continued use of legacy systems and unstructured data make it difficult to classify and track legacy data expansion.
4. Over-reliance on tools can be self-defeating as vendors sell "AI automation" that promises to solve governance problems without changing underlying organizational processes.
5. Limited oversight at the board of directors level, as boards and senior management may be aware of AI adoption, but not of the data governance mechanisms needed to protect data and models.

According to industry reports, organizations still need unified roadmaps that combine data governance, AI, and cybersecurity (Colombo & Flynn, 2025).

Risks and consequences

Significant risks are associated with poor data governance. These threats are not hypothetical, and technical evidence and guidelines from the IT sector already exist. Real incidents (and red team exercises) have repeatedly demonstrated how deficiencies in governance facilitate cyberattacks, threaten reputations, generate incomplete or inaccurate reports, violate regulations, and ultimately generate poor business performance.

CYBERSECURITY RISKS

This type of risk includes: data poisoning and model sabotage (small portions of poisoned data can significantly reduce model performance or alter its behavior); data exfiltration or theft through models facilitated by poorly governed data pools; and ad hoc connectors, shadow AI tools, and unmanaged third-party models that increase attack spread and expand the organization's attack surface.

ESG REPORTING AND REPUTATIONAL RISKS

Weak data governance relates to the quality of reporting as well as poorly reported ESG metrics. Increasingly, ESG reporting is based on auditable and traceable data, because poor quality control and traceability create the risk of unreliable sustainability or social impact claims.

The use of personal data in AI without strong governance may cause loss of privacy and consumer trust, besides adverse economic, social, reputational, and legal reactions, eroding stakeholder trust as well. The Organisation for Economic Co-operation and Development (OECD) and other regulatory bodies emphasize that data governance and privacy must be coordinated with AI governance to maintain public trust (OECD, 2024).

REGULATORY AND COMPLIANCE RISKS

Regulatory and compliance exposure is another risk to be considered due to the fragmentation of legal requirements. New AI and privacy rules vary widely by region and industry, and compliance requires demonstrable documentation, data

minimization, and risk mitigation. The lack of effective data governance increases the likelihood of fines, enforcement actions, and/or binding corrective measures, which in turn increases organizations' overall risk profiles (OECD, 2024).

FAILURE TO ACHIEVE BUSINESS RESULTS

In terms of financial effects, not all AI projects generate ROI. Analysts have noted that many initiatives are stalled, not because of models, but because data flows and governance frameworks are immature. As a result, organizations lose investment, time, and credibility, making it difficult for business strategies to succeed.

Additionally, incidents in which employees upload confidential documents to AI utilities (or use unapproved tools) have highlighted how poor data governance can lead to data leaks and regulatory risks. Enterprises are beginning to prohibit or strictly control these tools through formally established protocols, such as private models or contractual obligations with AI vendors.

Industry analysis predicts that many AI projects will fail due to poor data governance, rather than poor models. This is a clear sign that governance must be a priority (Chandrasekaran, 2026; Kleppe et al., 2025).

As organizations scale AI across core business functions, data governance can no longer be treated as a technical afterthought or a compliance checkbox.

Recommendations for scaling AI responsibly

As organizations scale AI across core business functions, data governance can no longer be treated as a technical afterthought or a compliance checkbox. AI systems depend on vast, dynamic, and often sensitive datasets. Without structured governance frameworks, enterprises expose themselves to regulatory penalties, cybersecurity incidents, reputational harm, and operational failures such as biased outputs or model drift.

Given this context, enterprises should adopt pragmatic and prioritized roadmaps that integrate data governance into their broader enterprise risk and strategy frameworks. This work must occur across multiple, interconnected dimensions of the AI lifecycle.

DATA GOVERNANCE AND LEADERSHIP

Organizations should establish a unified plan that aligns data governance, AI governance, and cybersecurity under their risk management. This ensures that AI-related risks—such as data misuse, privacy violations, adversarial attacks, and systemic bias—are treated as enterprise-level risks in a comprehensive way.

Responsibility for executing these plans should lie with senior executives. However, oversight and strategy approval should rest with the boards of directors. Research by firms such as KPMG has emphasized that board-level engagement is increasingly expected in AI governance, particularly as regulators focus on accountability and oversight. It is also necessary to establish a cross-functional council that includes representatives from all the relevant areas (cybersecurity and legal, among others) to mirror the complexity and amplitude of AI risks.

Product managers, engineers, and data scientists should receive training that integrates data governance principles directly into development workflows.

CONTROLS AND DATA ENGINEERING

One high-priority implementation is enterprise data catalogs that automatically register any datasets used in model training or inference. Having this mechanism in place enables traceability, auditability, and detection of data poisoning or drift. Without such visibility, organizations cannot confidently explain or reproduce model behavior.

Datasets should be formally classified (e.g., public, internal, confidential, or regulated), and there should be handling rules such as encryption standards, access controls, retention limits, and sensitivity labeling for each data class.

Provenance controls are equally critical. Before any dataset is used for training, it should have source verification and certification, validation checks, bias and representativeness assessments to prevent discrimination, and automated anomaly detection.

SAFETY AND THE AI LIFECYCLE

Governance should extend across the entire AI lifecycle—from data ingestion to model deployment and monitoring. Model governance should include signed datasets, reproducible builds, model signing, and runtime access controls. These practices enhance integrity and allow organizations to detect anomalies in model inputs and outputs.

The principle of least privilege should govern all data access. Credentials should be rotated regularly, and employees should be prohibited from uploading sensitive corporate data into public third-party AI systems unless explicitly authorized.

REGULATION, COMPLIANCE, AND REPORTING

As AI regulatory scrutiny is expanding globally, organizations must maintain clear and accessible documentation, including data source inventories, model documentation, and formal risk assessments. These records should be linked directly to compliance checklists required for regulatory, privacy, and ESG reporting.

Continuous auditing is a must as enterprises should schedule periodic red-team exercises focused on data poisoning, reverse engineering, and model extraction risks. Governance areas of opportunity identified during these exercises must be formally tracked and included in remediation reports presented to senior management.

Additionally, contractual safeguards with AI vendors are essential. Organizations must require vendors to get explicit approval before using company data to retrain or enhance general models.

PEOPLE AND CULTURE

No governance framework can succeed without cultural alignment. Product managers, engineers, and data scientists should receive training that integrates data governance principles directly into development workflows. Governance work must be reflected in performance metrics and incentives, rather than treated as secondary administrative overhead.

Before procuring third-party AI tools or models, enterprises should conduct due diligence on data provenance, data usage policies, security controls, and compliance posture. Procurement decisions should incorporate AI risk assessments alongside financial and operational criteria.

KEY METRICS AND PERFORMANCE INDICATORS (KPI)

To ensure accountability and measurable progress, organizations should track governance performance indicators to transform governance from an abstract principle into an operational discipline.

Conclusion

AI can bring transformative value, but it can also exacerbate existing gaps in organizations' data management and protection. The market and regulatory environment have shifted from prioritizing innovation over governance to viewing data governance as critical to responsible progress in AI. Enterprises that postpone or underinvest in data governance will face higher AI project failure rates, more cybersecurity incidents and regulatory breaches, and increased reputational damage. The way forward is clear: it must integrate data governance, cybersecurity, and AI governance.

References

Access Now. (2024). *Regulatory mapping on Artificial Intelligence in Latin America: Regional AI public policy report.* https://www.accessnow.org/wp-content/uploads/2024/07/TRF-LAC-Reporte-Regional-IA-JUN-2024-V3.pdf

Chandrasekaran, A. (2026). *Why 50% of GenAI projects fail — And how to beat the odds.* Gartner. https://www.gartner.com/en/articles/genai-project-failure

Colombo, M., & Flynn, G. (2025). *Data governance in the age of AI: Examining the paradigm shift to an integrated governance umbrella*. KPMG. https://kpmg.com/us/en/articles/2025/data-governance-age-ai.html

Díaz, R. M., & G. Núñez, G. (2023). *"Ciberataques a la logística y la infraestructura crítica en América Latina y el Caribe"*, Documentos de Proyectos (LC/TS.2023/93), Santiago, Economic Commission for Latin America and the Caribbean (ECLAC), 2023.

Innovative Routines International. (2025). Data governance and security in the age of AI. *Database Trends and Applications*, 13-18. https://www.iri.com/ftp9/pdf/Voracity/DBTA-2Q25-Data-Governance-and-Security-in-the-Age-of-AI.pdf

Kleppe, A., Mills, S., Broestl, N., Acenov, G., Katsov, K. & Yang, N. (2025). *What happens when AI stops asking permission?.* BCG. https://www.bcg.com/publications/2025/what-happens-ai-stops-asking-permission?utm_source=chatgpt.com

Kyndryl. (2025). *The readiness tipping point* [Readiness Report]. *https://www.kyndryl.com/us/en/insights/readiness-report-2025*

OECD. (2024). AI, data governance & privacy: Synergies and areas of international cooperation. *OECD Artificial Intelligence Papers*, *22*, 1-55. https://doi.org/10.1787/2476b1a4-en

Precisely & Drexel University's LeBow College of Business. (2025). *2025 Outlook: Data integrity trends and insights*. https://www.precisely.com/resource-center/analystreports/lebow-report-2024?utm_source=Blog-Post&utm_medium=Blog-In-Text&utm_campaign=Data-Integrity_Global_Content-Integrated-LeBow-Report_2024&utm_content=DIS_AR-LeBow-Data-Integrity-Trends-Insights_2024

Bansal, N. (2025). *AI and machine learning projects will fail without good data* [Opinion]. TechRadar. *https://www.techradar.com/pro/ai-and-machine-learning-projects-will-fail-without-good-data*

Interview I

Photo: Personal archive

Stefaan Verhulst
Co-founder of The GovLab and The DataTank and research professor at NYU

Between governed accessibility and the risk of a data winter in the era of Artificial Intelligence

In this interview, Stefaan Verhulst, co-founder of The Governance Lab (The GovLab) and The Data Tank and research professor at New York University (NYU) Tandon School of Engineering, reflects on the historical evolution of open data, the emergence of a "fourth wave" alongside a potential "data winter," and the central role of data governance as a strategic policy issue in the Artificial Intelligence (AI) era.

Internet Sectoral Overview (I.S.O.)_ In your recent work, you refer to the emergence of a "fourth wave of open data" occurring alongside a potential "data winter." Could you elaborate on these two dynamics and explain how they coexist, and possibly interact, in the current data ecosystem?

Stefaan Verhulst (S. V.)_ In a recent work,[24] we describe a paradoxical moment in the evolution of the data ecosystem: the emergence of a "fourth wave of open data"[25] alongside a potential "data winter."[26] These dynamics are not contradictory; rather, they represent two competing trajectories shaping the future of data in the AI era. The "fourth wave" should be understood as having two intertwined dimensions. First, it represents a shift from openness as default toward openness as "governed accessibility". Earlier waves of open data[27] focused on transparency (freedom of information), portal creation (open government data), and data collaboration (public-private partnerships). The current phase builds upon these past waves. It recognizes that many of the datasets most relevant for societal problem-solving and AI development are sensitive, proprietary, or distributed across institutional boundaries. It also recognizes that openness by itself is not enough and can sometimes be weaponized,[28] and that efforts to open data must be focused on data that others will meaningfully use for the public good. As a result, openness increasingly depends on governance innovations—data commons[29] and structured access models—that allow responsible reuse while safeguarding rights and managing risk.

[24] Available at: https://sverhulst.medium.com/it-was-the-best-of-times-it-was-the-worst-of-times-the-dual-realities-of-data-access-in-the-age-8732837acaa5
[25] Find out more: https://www.genai.opendatapolicylab.org/
[26] Find out more: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5660451&__cf_chl_tk=fa7jYq7vq81lbsbHr2mi3V4IicBQuSqgcV45Ag.xoxo-1772347099-1.0.1.1-83bvz4DvQOFJ055S3BPMDW8V..HfMfIfSmhp5PeG0Gk
[27] Find out more: https://dx.doi.org/10.2139/ssrn.3937638
[28] Find out more: https://sverhulst.medium.com/the-weaponisation-of-openness-toward-a-new-social-contract-for-data-in-the-ai-era-fb9f49ef6109
[29] Find out more: https://blog.thegovlab.org/the-new-commons-challenge-advancing-ai-for-public-good-through-data-commons

Second, the "fourth wave" reflects the growing interconnection between generative AI and open data. This relationship is bidirectional[30]: Open data can improve generative AI ("data for AI"), while generative AI can fundamentally transform how people access and use open data ("AI for data"). Research conducted[31] through the Open Data Policy Lab shows that open data can enhance model quality, reduce hallucinations, advance open source AI, and expand use cases when applied to training, adaptation, or retrieval-augmented architectures. Conversely, generative AI interfaces can make data more conversational and accessible, lowering barriers to use and democratizing access beyond technical specialists.

At the same time, these advances occur against the backdrop of a "data winter," a growing enclosure of data driven by risk aversion and policy confusion, generative AI–induced anxieties about extraction, research data lockdowns, geopolitical fragmentation, and increasing privatization.

The result is widening structural asymmetries where a small number of actors control high-value datasets while many researchers and public-interest actors face shrinking access.

Thus, the "fourth wave" and the "data winter" coexist because they respond to the same structural transformation: Data has become both more valuable and more contested. This wave represents an effort to redesign openness for the AI era; the "data winter" reflects the risks of failing to do so. Which trajectory prevails will depend on whether governance, incentives, and institutional innovation evolve quickly enough to support trusted and equitable reuse.

I.S.O._ You argue that this "fourth wave" could simultaneously enhance AI systems and democratize access to data. What technical, institutional, or governance-related conditions are necessary to ensure that this potential generates tangible public value?

S.V._ The promise of the "fourth wave" lies precisely in its dual ambition: enabling data to improve AI systems while using AI to expand the accessibility and usability of data. Yet this promise will only materialize if several technical, institutional, and governance conditions are met.

First, from a technical perspective, one of the key findings from our research is that open data must become "AI-ready,"[32] but readiness is contextual. Our Spectrum of scenarios[33] framework demonstrates that different intersections between open data and generative AI—pretraining, adaptation, inference, data augmentation, or open-ended exploration—require distinct standards of data quality, metadata, and provenance. In other words, data quality is purpose-dependent; what works for model training may be unsuitable for reasoning or policy inference.

Second, governance and documentation become central technical enablers.

"Thus, the 'fourth wave' and the 'data winter' coexist because they respond to the same structural transformation: Data has become both more valuable and more contested. This wave represents an effort to redesign openness for the AI era; the 'data winter' reflects the risks of failing to do so."

[30] Find out more: https://repository.opendatapolicylab.org/genai
[31] Available at: https://doi.org/10.48550/arXiv.2405.04333
[32] Available at: https://arxiv.org/html/2404.05779v1
[33] Available at: https://www.genai.opendatapolicylab.org/

"(...) institutional capacity and structures must evolve. (...) Without dedicated stewardship roles, even strong legal frameworks risk remaining aspirational rather than actionable."

Transparency about data lineage, sourcing, and intended use allows users to evaluate reliability and mitigate risks. Enhanced metadata standards, provenance tracking, and ethical oversight mechanisms are essential to ensure trust and accountability in AI-mediated data environments.

Third, institutional capacity and structures must evolve. Data stewardship[34] plays a pivotal role in bridging demand and supply, translating governance principles into operational practice, and enabling collaboration across sectors. Without dedicated stewardship roles, even strong legal frameworks risk remaining aspirational rather than actionable. Complementing this human infrastructure, data commons[35] represent a collective governance model that repositions data not as a private commodity but as a renewable societal asset, enabling shared value creation through coordinated stewardship and participatory governance.

Finally, legitimacy, validation, and participation are indispensable. If communities perceive data reuse as extractive—particularly in AI contexts—trust erodes and access closes. Governance frameworks must therefore include mechanisms for social license[36] and participatory decision-making that embed community agency. Equally important are continuous validation processes, enabled through structured feedback loops that allow affected stakeholders to evaluate outcomes, flag unintended consequences, and iteratively refine governance arrangements over time. Such feedback mechanisms transform governance from a one-off approval exercise into an adaptive system, ensuring that data reuse remains aligned with evolving societal expectations, contextual realities, and public value objectives.

Ultimately, the fourth wave generates public value when it moves beyond the binary of openness vs. restriction and instead builds systems where data flows responsibly, AI systems become more transparent and reliable, and the benefits of data-driven innovation are distributed more equitably across society.

I.S.O._ Data governance appears to be central to unlocking this next phase of open data. In the current context, marked by geopolitical fragmentation, regulatory divergence, and uneven digital capacity, what would meaningful progress in data governance look like in practice? What risks may arise if this agenda fails to advance?

S.V._ In a context marked by geopolitical fragmentation, dependency weaponisation, regulatory divergence, and uneven institutional capacity, meaningful progress in data governance should not be understood as achieving global uniformity.[37] Rather, progress lies in developing governance systems that are purpose-driven, interoperable, and institutionally adaptable—capable of supporting responsible data reuse across different political, legal, and cultural contexts.

[34] Find out more: https://doi.org/10.48550/arXiv.2502.10399
[35] Find out more: https://incubator.opendatapolicylab.org/files/data-commons-for-ai-blueprint.pdf
[36] Find out more: https://www.afd.fr/en/ressources/reimagining-data-governance-ai
[37] Find out more: https://doi.org/10.48550/arXiv.2302.13731

In recent work,[38] we defined data governance as the combination of people, processes, policies, practices, and technologies that guides decisions across the data lifecycle in alignment with a set of principles toward trust, value creation, equity, and risk mitigation.

This framing highlights an important shift: Governance is not merely compliance or control, but also a decision-making architecture that answers four questions—why data is used, how it should be governed, who is responsible, and what mechanisms operationalize those decisions. Meaningful progress, therefore, requires moving beyond fragmented, reactive governance toward structured institutional capacity.

First, governance must become operational rather than declarative. Many jurisdictions possess principles but lack institutional roles—such as data stewards—capable of translating those principles into daily decisions across the lifecycle. Capacity building, training, and institutionalization of stewardship roles are therefore central to progress.

Second, governance must reduce the "cost of distrust." Mechanisms such as interoperability standards, provenance documentation, lifecycle traceability, and transparent access criteria allow actors to collaborate without sacrificing accountability. Governance frameworks should be proportional, i.e., calibrated to risk and purpose rather than applying uniform restrictions across all data uses.

Third, progress requires participation and agency. Governance increasingly recognizes that legitimacy cannot rely solely on legal compliance. Mechanisms such as social licensing, participatory design, and digital self-determination[39] enable communities to shape how data is used and reused.

If this agenda fails to advance, the consequences are systemic. Data asymmetries will deepen, reinforcing the concentration of AI capabilities among a few actors. Cross-border collaboration on global challenges will become more difficult, and innovation may increasingly rely on low-quality or synthetic substitutes rather than trustworthy data. Most importantly, the legitimacy of data-driven systems may erode, accelerating the "data winter" dynamic where access contracts rather than expands.

> "(...) Governance is not merely compliance or control, but also a decision-making architecture that answers four questions—why data is used, how it should be governed, who is responsible, and what mechanisms operationalize those decisions."

I.S.O._ You recently co-authored the* Data Governance Toolkit: Navigating Data in the Digital Age[40]*. What specific gaps in current data governance practices does this initiative aim to address, and which core principles do you see as most critical for public and private institutions seeking to build governance frameworks that are both robust and context-sensitive?

S.V._ The *Data Governance Toolkit: Navigating Data in the Digital Age* was developed in response to a recurring observation: While many institutions recognize the importance of data governance, they often lack practical frameworks that translate principles into operational practice. Existing resources are frequently technical, siloed, or inaccessible to policymakers, particularly in contexts with limited institutional capacity.

[38] Available at: https://incubator.opendatapolicylab.org/files/what-is-data-governance-report.pdf
[39] Find out more: https://doi.org/10.1017/dap.2023.11
[40] Available at: https://unesdoc.unesco.org/ark:/48223/pf0000394518

"The toolkit addresses [the] gap by positioning governance as both a capacity-building exercise and a practical design framework. Rather than treating governance as a compliance checklist, it reframes it as purpose-driven decision-making across the data lifecycle—focused on building trust, enabling reuse, and ensuring equity."

The toolkit addresses this gap by positioning governance as both a capacity-building exercise and a practical design framework. Rather than treating governance as a compliance checklist, it reframes it as purpose-driven decision-making across the data lifecycle—focused on building trust, enabling reuse, and ensuring equity.

Several specific gaps motivated the toolkit creation. The first is the implementation gap. Organizations often articulate ethical aspirations but struggle to operationalize them. The toolkit, therefore, emphasizes clear governance functions: defining purpose, establishing guiding principles, clarifying responsibilities, and implementing concrete tools and practices.

The second is the lifecycle gap. Governance has historically emphasized protection at the point of collection, while undergoverning access and reuse, which is the stage where both societal value and risk often emerge. The toolkit promotes lifecycle thinking that balances protection with responsible access and reuse.

The third is the contextual gap. Countries operate at very different levels of governance maturity. The toolkit is intentionally modular, designed to support jurisdictions that are building foundational frameworks as well as those confronting emerging issues such as generative AI, cross-border flows, or interoperability challenges.

Several core principles are central. Purposefulness ensures that governance begins with clearly articulated societal objectives. Proportionality aligns safeguards with risk rather than imposing blanket restrictions. Stewardship emphasizes responsibility over ownership, enabling systematic and sustainable access. Digital self-determination ensures that governance reflects diverse stakeholders and reduces power asymmetries. Human-rights grounding ensures that governance protects dignity, equity, and agency across contexts.

More importantly, the toolkit reframes governance as enabling infrastructure rather than constraint. Its goal is to help policymakers and institutions move from reactive governance toward anticipatory, adaptive systems that support innovation while preserving trust and legitimacy. In this sense, the toolkit reflects a broader evolution in the field: Governance is no longer simply about preventing misuse. It is about enabling responsible, equitable, and future-ready data ecosystems.

Domain Report

Domain registration dynamics in Brazil and around the world

The Regional Center for Studies on the Development of the Information Society (Cetic.br), department of the Brazilian Network Information Center (NIC.br), carries out monthly monitoring of the number of country code top-level domains (ccTLD) registered in countries that are part of the Organisation for Economic Co-operation and Development (OECD) and the G20.[41] Considering members from both blocs, the 20 nations with the highest activity sum more than 98.13 million registrations. In March 2026, domains registered under .de (Germany) reached 17.86 million, followed by China (.cn), United Kingdom (.uk), and Russia (.ru), with 12.47 million, 8.91 million, and 6.11 million registrations, respectively. Brazil had 5.63 million registrations under .br, occupying 6th place on the list, as shown in Table 1.[42]

[41] Group composed by the 19 largest economies in the world and the European Union. More information available at: https://g20.org/

[42] The table presents the number of ccTLD domains according to the indicated sources. The figures correspond to the record published by each country, considering members from the OECD and G20. For countries that do not provide official statistics supplied by the domain name registration authority, the figures were obtained from: https://research.domaintools.com/statistics/tld-counts. It is important to note that there are variations among the date of reference, although the most up-to-date data for each country is compiled. The comparative analysis for domain name performance should also consider the different management models for ccTLD registration. In addition, when observing rankings, it is important to consider the diversity of existing business models.

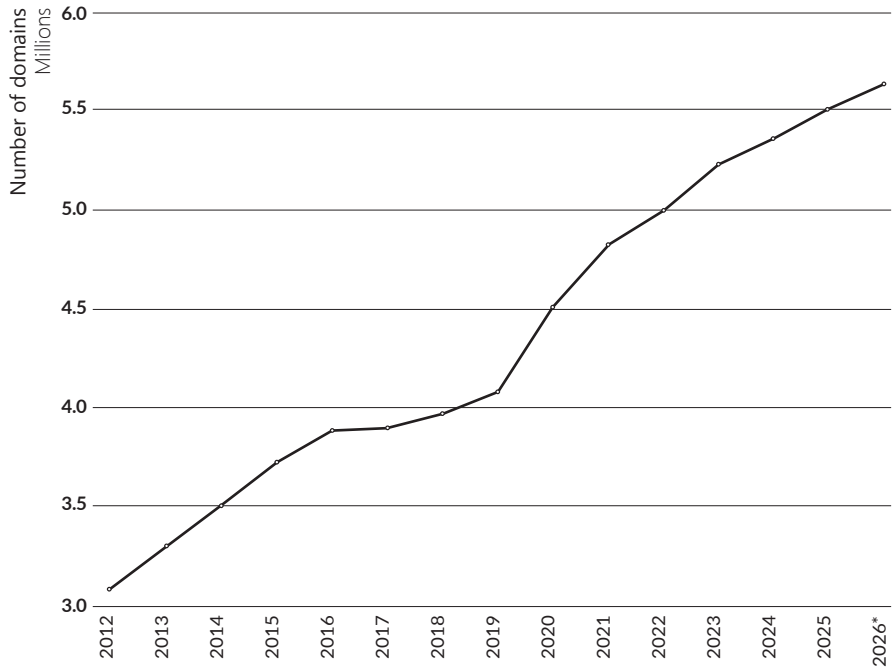
Table 1 – TOTAL REGISTRATION OF DOMAIN NAMES AMONG OECD AND G20 COUNTRIES

Position	Country	Number of domains	Date of reference	Source (website)
1	Germany (.de)	17,863,162	01/04/2026	https://www.denic.de
2	China (.cn)	12,473,901	01/04/2026	https://research.domaintools.com/statistics/tld-counts/
3	United Kingdom (.uk	8,910,553	28/02/2026	https://nominet.uk/reports-and-statistics/
4	Russia (.ru)	6,115,947	01/04/2026	https://cctld.ru
5	Netherlands (.nl)	6,063,116	28/02/2026	https://stats.sidnlabs.nl/en/registration.html
6	**Brazil (.br)**	**5,630,927**	**31/03/2026**	**https://registro.br/dominio/estatisticas/**
7	France (.fr)	4,402,660	30/03/2026	https://www.afnic.fr/en/observatory-and-resources/statistics/
8	Australia (.au)	4,336,695	01/04/2026	https://www.auda.org.au/
9	European Union (.eu)	3,693,521	01/04/2026	https://research.domaintools.com/statistics/tld-counts/
10	Italy (.it)	3,572,261	01/04/2026	http://nic.it
11	Canada (.ca)	3,539,662	01/04/2026	https://www.cira.ca
12	India (.in)	3,373,980	01/04/2026	https://research.domaintools.com/statistics/tld-counts/
13	Colombia (.co)	3,015,037	01/04/2026	https://research.domaintools.com/statistics/tld-counts/
14	Switzerland (.ch)	2,594,855	15/03/2026	https://www.nic.ch/statistics/domains/
15	Poland (.pl)	2,534,384	01/04/2026	https://research.domaintools.com/statistics/tld-counts/
16	United States (.us)	2,175,340	30/11/2025	https://www.about.us/stats-trends
17	Spain (.es)	2,164,199	28/02/2026	https://www.dominios.es/es/sobre-dominios/estadisticas
18	Portugal (.pt)	2,124,743	01/04/2026	https://www.dns.pt/en/statistics/
19	Japan (.jp)	1,844,251	01/04/2026	https://jprs.co.jp/en/stat/
20	Belgium (.be)	1,703,847	01/04/2026	https://www.dnsbelgium.be/en

Collection date: April 1, 2026.

Chart 1 shows the performance of .br since 2012.

Chart 1 – TOTAL NUMBER OF DOMAIN REGISTRATIONS FOR .BR – 2012 to 2026*



*Collection date: March 31, 2026.
Source: Registro.br
Retrieved from: https://registro.br/dominio/estatisticas

In march 2026, the five generic Top-Level Domains (gTLD) totaled more than 199.32 million registrations. With 161.19 million registrations, .com ranked first, as shown in Table 2.

Table 2 – TOTAL NUMBER OF DOMAINS AMONG MAIN gTLD

Position	gTLD	Number of domains
1	**.com**	161,191,505
2	**.net**	12,205,131
3	**.org**	11,752,093
4	**.xyz**	8,090,963
5	**.top**	6,080,506

Collection date: April 1, 2026.
Source: DomainTools.com
Retrieved from: research.domaintools.com/statistics/tld-counts

DATA GOVERNANCE:

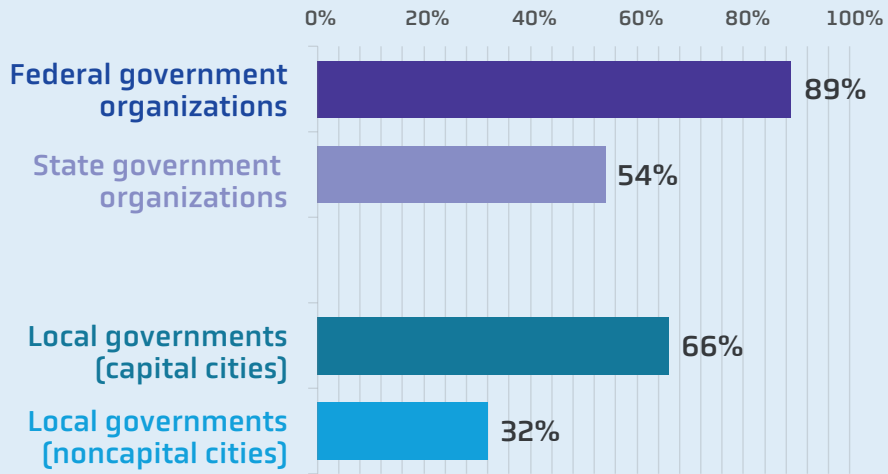
Information security plans in Brazilian government organizations

The world we live in is characterized by the growing digitization of public and private services and advances in Artificial Intelligence (AI), phenomena driven by the continuous and exponential increase in the volume of data collected, stored, and processed: the so-called era of datafication. In this environment, data governance plays a central role as associated risks grow, including the misuse of information, leaks of sensitive data, and automated decisions without clear criteria. This scenario underscores the importance of establishing guidelines that promote data security, transparency, and responsible use.

In Brazil, the adoption of formal information security measures is already a reality in parts of the public sector. According to the ICT Government 2023[43] survey, 89% of federal organizations, 54% of state government organizations, 66% of local governments in state capitals, and 32% of local governments in noncapital cities have formally established information security plans, among those with an information technology (IT) department (Chart 1).

Chart 1 – INFORMATION SECURITY PLANS IN BRAZILIAN GOVERNMENT ORGANIZATIONS

Total number of state and federal government organizations and local governments in Brazil with an information technology department (%)



[43] Data from the ICT Electronic Government 2023 Survey by Cetic.br|NIC.br, available at: https://cetic.br/pt/tics/governo/2023/orgaos/B5A/ and https://cetic.br/pt/tics/governo/2023/prefeituras/B5A/

/Credits

TEXT

DOMAIN REPORT

Thiago Meireles (Cetic.br|NIC.br)

GRAPHIC DESIGN

Thiago Planchart (Comunicação|NIC.br)

PUBLISHING

Grappa Marketing Editorial

ENGLISH REVISION AND TRANSLATION

Prioridade Consultoria Ltda.: Isabela Ayub, Lorna Simons, Luana Guedes, Luísa Caliri, and Maya Bellomo Johnson

EDITORIAL COORDINATION

Alexandre F. Barbosa, Graziela Castello, Javiera F. M. Macaya, Rodrigo Brandão, and Mariana Galhardo Oliveira (Cetic.br|NIC.br)

ACKNOWLEDGMENTS

Renan Mendes Gaya Lopes dos Santos (SGD/MGI)
Thaciana Guimarães de Oliveira Cerqueira (SGD/MGI)
Camila Falchetto (SGD/MGI)
Alex Lopes Pereira (SGD/MGI)
Daniel Insulza (CAF)
Zachary Nichols
Georgina Núñez (Ceuta)
Stefaan Verhulst (The GovLab, The Data Tank and NYU)

ABOUT CETIC.br

The Regional Center for Studies on the Development of the Information Society – Cetic.br (https://www.cetic.br/en/), a department of NIC.br, is responsible for producing studies and statistics on the access and use of the Internet in Brazil, disseminating analyzes and periodic information on the Internet development in the country. Cetic.br acts under the auspices of UNESCO.

ABOUT NIC.br

The Brazilian Network Information Center – NIC.br (http://www.nic.br/about-nic-br/) is a non-profit civil Entity in charge of operating the .br domain, distributing IP numbers, and registering Autonomous Systems in the country. It conducts initiatives and projects that bring benefits to the Internet infrastructure in Brazil.

ABOUT CGI.br

The Brazilian Internet Steering Committee – CGI.br (https://cgi.br/about/), responsible for establishing strategic guidelines related to the use and development of the Internet in Brazil, coordinates and integrates all Internet service initiatives in the country, promoting technical quality, innovation, and dissemination of the services offered.

*The ideas and opinions expressed in the texts of this publication are those of the respective authors and do not necessarily reflect those of NIC.br and CGI.br.

unesco
Centre under the auspices of UNESCO

cetic.br
Regional Center for Studies on the Development of the Information Society

nic.br
Brazilian Network information Center

cgi.br
Brazilian Internet Steering Committee

STRIVING FOR A BETTER INTERNET IN BRAZIL

CGI.br, MODEL OF MULTISTAKEHOLDER GOVERNANCE

www.cgi.br

nic.br cgi.br

PHOTOS: RICARDO MATSUKAWA AND SHUTTERSTOCK