

ARTIFICIAL INTELLIGENCE IN PUBLIC ADMINISTRATION: ADMINISTRATIVE EFFICIENCY AND CITIZEN PARTICIPATION

INTELIGÊNCIA ARTIFICIAL NA GESTÃO PÚBLICA: EFICIÊNCIA ADMINISTRATIVA E PARTICIPAÇÃO CIDADÃ

INTELIGENCIA ARTIFICIAL EN LA GESTIÓN PÚBLICA: EFICIENCIA ADMINISTRATIVA Y PARTICIPACIÓN CIUDADANA

EDVALDO PINTO MANINHO JÚNIOR

Undergraduate student in Public Administration, Instituto Federal de Educação, Ciência e Tecnologia de Rondônia, Brazil

E-mail: epmaninho@gmail.com

MARINEIDE MARTINIANO DO NASCIMENTO

Thesis Supervisor

E-mail: profneidemn@gmail.com

Abstract

The growing incorporation of algorithmic technologies in public administration places artificial intelligence (AI) at the centre of the debate on state modernisation and its democratic risks. This study analyses the contributions, risks, and limits of AI adoption in Brazilian public administration, focusing on its effects on administrative efficiency and citizen participation. Through an analytical narrative literature review, documentary analysis of national and international regulatory frameworks, and multiple case studies — Sapiens/AGU, Victor/STF, Alice/TCU, Gov.br, and COMPAS/USA — the research identifies adoption patterns, success conditions, and risk factors. The bibliographic corpus comprised 28 references selected from Web of Science, Scopus, and Google Scholar, supplemented by institutional documentary sources. The results show that successful Brazilian cases share three structural characteristics: formalised human oversight, delimited functional scope, and integration with pre-existing control processes. In contrast, the absence of algorithmic governance exposes public authorities to risks of decisional opacity, discriminatory bias, and digital exclusion — all incompatible with art. 37, caput, of the 1988 Federal Constitution. The proposed five-axis governance model — legality and proportionality; transparency and explainability; human oversight and accountability; non-discrimination and data protection; and training and inclusion — emerges from the joint analysis of cases and the national and international legal framework, offering operational parameters for the evaluation, deployment, and monitoring of public AI systems.

Keywords: Artificial intelligence; Public administration; Administrative efficiency; Citizen participation; Algorithmic governance.

Resumo

A crescente incorporação de tecnologias algorítmicas na administração pública coloca a inteligência artificial (IA) no centro do debate sobre modernização do Estado e seus riscos democráticos. Este estudo analisa as contribuições, os riscos e os limites da adoção de IA na administração pública brasileira, com foco em seus efeitos sobre a eficiência administrativa e a participação cidadã. Por meio de revisão narrativa analítica da literatura, análise documental de marcos normativos nacionais e internacionais e estudo de casos múltiplos — Sapiens/AGU, Victor/STF, Alice/TCU, Gov.br e COMPAS/EUA —, a pesquisa identifica padrões de adoção, condições de sucesso e fatores de risco. O modelo orientador em cinco eixos proposto emerge da análise conjunta dos casos e do ordenamento jurídico nacional e internacional, oferecendo parâmetros operacionais para a avaliação, implantação e monitoramento de sistemas públicos de IA.

Palavras-chave: Inteligência artificial; Gestão pública; Eficiência administrativa; Participação cidadã; Governança algorítmica.

Resumen

La creciente incorporación de tecnologías algorítmicas en la administración pública sitúa la inteligencia artificial (IA) en el centro del debate sobre modernización del Estado y sus riesgos democráticos. Este estudio analiza las contribuciones, los riesgos y los límites de la adopción de IA en la administración pública brasileña. El modelo de gobernanza en cinco ejes propuesto emerge del análisis conjunto de los casos y del ordenamiento jurídico nacional e internacional, y ofrece parámetros operativos para la evaluación, implementación y monitoreo de sistemas públicos de IA.

Palabras clave: Inteligencia artificial; Gestión pública; Eficiencia administrativa; Participación ciudadana; Gobernanza algorítmica.

1. Introduction

Brazilian public administration simultaneously faces three structural challenges that reinforce each other: persistent inefficiency in service delivery, the growing distance between the state and citizens in decision-making processes, and increasing pressure for transparency and accountability in a context of severe budgetary constraints. In this scenario, artificial intelligence (AI) stands out as one of the most promising technological responses, with the potential to automate repetitive tasks, optimise the use of public resources, and expand citizen participation channels — in line with the constitutional imperative of efficiency set out in art. 37, caput, of the 1988 Federal Constitution.

The adoption of AI by public authorities is not, however, without risks. Algorithmic opacity conflicts with the principles of publicity and mandatory justification of administrative acts. Systems trained on historical data may reproduce and amplify discriminatory biases, deepening pre-existing inequalities. Accelerated digitalisation, unaccompanied by inclusion and digital literacy policies, may convert old bureaucratic barriers into new forms of exclusion, equally harmful to the most vulnerable groups (CAMPILLO; HONORATO, 2021). Added to these challenges is the fact that AI's potential to strengthen citizen participation remains largely underexplored both in the literature and in Brazilian governmental practice.

Despite the regulatory advances achieved — from the Brazilian Digital Transformation Strategy (E-Digital, 2018) to Decree 10,332/2020 — much of the public service is still perceived as costly, slow, and distant from citizens' real priorities. The COVID-19 pandemic simultaneously revealed the potential and fragilities of e-government: in forcing the mass adoption of digital solutions, it

exposed deficiencies in data infrastructure, heterogeneity across federal entities, and the digital exclusion of vulnerable populations (CEPAL; CAF, 2020).

Against this backdrop, the following research question is posed: under what conditions can the adoption of artificial intelligence solutions expand administrative efficiency and strengthen citizen participation in Brazilian public administration without violating fundamental rights, deepening structural inequalities, or generating a democratic deficit?

The general objective of this study is to identify the conditions for responsible AI adoption in the Brazilian public sector, analysing comparative evidence of efficiency and citizen participation in light of the constitutional principle of efficiency. The specific objectives are: (i) to identify the main AI applications in the Brazilian public sector and the patterns that distinguish successful cases from problematic ones; (ii) to examine the relationship between AI adoption and the constitutional principle of efficiency from a post-positivist perspective; (iii) to assess ethical, social, and democratic risks in light of the national legal order and international regulatory framework; and (iv) to propose a governance model grounded in the empirical and normative analysis conducted.

The relevance of this study manifests on two levels. At the theoretical level, it articulates constitutional administrative law with the computer science and public policy literature, filling a gap regarding the conditions for legitimate AI adoption in the Brazilian public sector. At the practical level, it offers guidance for managers and policy-makers in a context of budgetary constraints and growing social demand for transparency — a context in which the National Association of Public Policy and Government Management Specialists (ANESP, 2023) has already identified AI as a strategic priority for state modernisation.

The article is structured in five sections. Section 2 presents the literature review, covering the conceptual foundations of AI, the trajectory of digital transformation in Brazil, the efficiency principle from a post-positivist perspective, concrete public sector applications, the international regulatory landscape, and the main identified risks. Section 3 describes the methodological procedures. Section 4 presents the results and discussion, including a comparative case analysis using a structured matrix. Section 5 presents the conclusions and future research agenda.

2. Literature Review

2.1 Artificial Intelligence: concepts and foundations applied to the public sector

The term artificial intelligence was coined by John McCarthy in 1956 to designate the science and engineering dedicated to creating intelligent machines, especially computer programs. According to McCarthy (2007), AI corresponds to the computational part of the ability to achieve goals in the world, encompassing mechanisms that transcend biologically observable methods. The Organisation for Economic Cooperation and Development (OECD, 2023) defines AI as a machine-based system that, from the information it receives, infers how to generate outputs — predictions, content, recommendations, or decisions — capable of influencing physical or virtual environments, with varying degrees of autonomy and adaptability.

Three technical components underpin the main public AI applications. Machine learning enables machines to analyse data, identify patterns, and make decisions without direct human intervention, through algorithms that continuously improve (KAUFMAN; SANTAELLA, 2020). Artificial neural networks are computational structures inspired by the architecture of the human brain, capable of classifying large volumes of data with high accuracy. Natural language processing (NLP), in turn, focuses on developing programs capable of analysing, recognising, and generating text in human languages, forming the technical basis of virtual assistants, *chatbots*, and document triage tools (CONEGLIAN; SEGUNDO, 2022).

Combined, these components enable what Araújo, Zullo and Torres (2020) call non-human decision-making: the massive analysis of data combined with self-learning algorithms can produce more efficient results than those generated by individual decision-making processes. This phenomenon constitutes what the international literature calls the algorithmic society — a context in which algorithms permeate everyday decisions, from consumer recommendation systems to public policies, with frequently invisible impacts on citizens' lives.

The implications for the public sector are significant. Da Silva et al. (2021) argue that the main effect of AI on the public workforce will be the complementation

and requalification of civil servants, with the potential to free up to 50% of functional time for activities that machines have not yet mastered. For Bollotti and Wachowicz (2024), when AI demonstrates superior and verifiable results compared to traditional methods, maintaining more costly analogue protocols may constitute a misuse of purpose and a violation of the constitutional duty of efficiency. It is important, however, to distinguish three analytically distinct dimensions: the legal possibility of adopting AI, the administrative convenience of doing so in a given context, and the eventual legal duty of adoption — the latter requiring objective demonstration of the technological system's superiority, sufficient state capacity to sustain it, and compatibility with the fundamental rights of those affected.

2.2 Digital transformation and e-government in Brazil

Digital transformation in the public sector cannot be confused with the mere digitalisation of forms or the migration of services to the internet. It is a deeper organisational change, supported by data, technologies, and new institutional arrangements, which redesigns processes, roles, and decision-making models with the objective of producing public value with greater agility, quality, and transparency (ARAÚJO et al., 2025). In Brazil, this agenda assumed state policy status with the Brazilian Digital Transformation Strategy (E-Digital/MCTIC, 2018) and Decree 10,332/2020, which established the Digital Government Strategy.

The trajectory of Brazilian e-government runs from the historic Rede Governo portal, launched in the late 1990s, to the current Gov.br, conceived as a "platform of platforms": a single access point that accelerates digitalisation, broadens service reach, and reduces transaction costs for both the state and citizens (SPANÓ; SILVA, 2022). Diniz et al. (2009) record that this journey unfolded in well-defined phases, from basic computerisation to user-oriented digital government. Studies on Brazilian smart cities during the COVID-19 pandemic (FARINIUK, 2020) show that municipalities with more robust digital infrastructure responded more swiftly to emergency demands — evidence that prior digital transformation constitutes a factor of institutional resilience.

Constitutional Amendment 85/2015 elevated science, technology, and innovation to the status of a constitutional mission of the state, reinforcing the

strategic character of these agendas. This normative decision did not, however, eliminate the practical obstacles to digital transformation. The absence of interoperability standards, federal fragmentation, and technological dependence on private suppliers generate *lock-in* effects that progressively increase maintenance and integration costs (ARAÚJO et al., 2025). Capacity asymmetries between municipalities are particularly relevant: while large federal entities have the technical infrastructure, resources, and specialised teams to adopt and sustain AI systems, smaller municipalities often lack the minimum conditions of data governance, technology procurement, and civil servant training. This structural heterogeneity determines that the algorithmic governance model proposed in this study must necessarily be graduated according to available state capacity, rather than treated as a uniform standard applicable to all entities.

2.3 The efficiency principle and AI as an instrument of good administration

The efficiency principle, constitutionalised by art. 37, caput, of the 1988 Federal Constitution through Constitutional Amendment 19/1998, transcends the managerial dimension to become a mandatory legal precept. In light of post-positivist hermeneutics, Hachem and Gabardo (2018) argue that efficiency imposes on Public Administration the duty to exercise administrative power with speed, promptness, economy, and productivity, always reconciling the pursuit of results with the observance of fundamental rights and other constitutional principles. It is not, therefore, a financial-economic calculation, but a normative mandate encompassing quality, response time, and the effectiveness of rights.

This post-positivist reinterpretation projects the efficiency principle to the centre of the debate on AI adoption by public authorities. Bollotti and Wachowicz (2024) argue that, in the context of neoconstitutionalism, constitutional principles cease to be subsidiary instruments to fill gaps and acquire full normative force. In this framework, maintaining more costly analogue protocols when comprovably superior technological alternatives exist may constitute a misuse of purpose. Unjustified technological inertia thus acquires potential legal consequences — although the characterisation of such duty requires, in each case, the objective demonstration of three elements: verifiable superiority of technological results,

sufficient state capacity to implement and maintain the system, and compatibility with the fundamental rights of those affected.

The concept of the fundamental right to good administration, formulated by Freitas (2007), complements this normative framework: good administration comprises the fundamental right to an efficient, effective, proportionate, transparent, motivated, and impartial Public Administration. In this horizon, AI can act both as an instrument of efficiency and as a vector of participation, by synthesising contributions from public consultations, organising arguments by theme, identifying regulatory gaps, and translating technical language into formats accessible to citizens — democratising public hearing and reducing the costs of participation.

Santos (2025), in a case study on the use of Chat-GPT in a Legislative House in Greater São Paulo, records that civil servants who use generative AI report gains in speed and quality in drafting bill proposals, motions, and information bulletins. The same study documents, however, significant cultural barriers: part of the civil servants interviewed is unfamiliar with the foundations of new public management and the efficiency criteria that should guide public work. This finding reveals that functional training is a precondition for the productive adoption of any technological tool — a perspective equally recognised as a priority on the Brazilian state modernisation agenda (ANESP, 2023).

2.4 Concrete AI applications in Brazilian public administration: patterns and success conditions

Brazil brings together a growing set of AI experiences in the public sector, distributed across two complementary axes: the legal-governmental *back-office* and the citizen-facing *front-office*. The comparative analysis of these cases — developed systematically in section 4.1 — reveals three recurrent patterns that distinguish successful initiatives: delimited functional scope, structured human oversight, and integration with pre-existing control systems.

In the *back-office*, the Sapiens system — and its evolved version, Super Sapiens — of the Attorney General of the Union (AGU) stands out. Integrated with an AI assistant based on GPT-4 technology, the system is capable of triaging and classifying administrative, advisory, and judicial proceedings, supporting the rapid

analysis of documents, and suggesting legally grounded texts. Super Sapiens operates 70% faster than the previous version and was recognised with the Innovare Prize 2020 (BOLLOTTI; WACHOWICZ, 2024). Its success derives from a structural characteristic: all AI-generated documents are submitted to review by solicitors, which maintains human accountability and preserves the justification of administrative acts. It should be noted, however, that the performance data cited comes from the AGU's own institutional sources, and not from independent external evaluation — a limitation that must be considered when reading the results.

In the Judiciary, three tools deserve mention. Victor, developed by the Supreme Court (STF) in partnership with the University of Brasília (UnB), performs the classification of procedural documents and the identification of general repercussion themes. Athos, from the Superior Court of Justice (STJ), optimises the management of qualified precedents. Radar, from the Minas Gerais Court of Justice (TJMG), automates decisions in simple cases under the final supervision of a magistrate (BOLLOTTI; WACHOWICZ, 2024). The common denominator among the three tools is relevant: none of them replaces human judgement on matters that directly affect parties' rights — AI operates in the preparatory and triage phase, not in the final decision.

In external control, the Federal Court of Accounts (TCU) operates the Information Control Laboratory (Labcontas), which runs the robots Alice (Procurement Analysis), Sofia, and Mônica. Alice detects, on the very day of publication, indications of irregularities in public procurement notices and minutes available on ComprasNet. In the field of social control, the Serenata de Amor Operation uses AI to analyse expenses reimbursed by the Parliamentary Activity Exercise Quota (CEAP), allowing any citizen to scrutinise acts of public agents (DESORDI; BONA, 2020). Both cases illustrate that AI can function as an accountability mechanism when applied to structured public data — a condition that simultaneously maximises reliability and restricts the scope of application.

In the *front-office*, Gov.br has established itself as the central platform for digital public service delivery, with more than 4,500 federal digital services available. State-level initiatives such as Gov.Pi Cidadão replicate the federal model at the subnational level (ARAÚJO et al., 2025). *Chatbots* and virtual assistants have

been deployed in health, public safety, and citizen service, reducing wait times and broadening access to information. Contrasting with these initiatives, the COMPAS case in the United States illustrates the systemic risks when AI is applied to high-consequence decisions affecting individual rights without adequate human oversight and without algorithmic transparency (ARAÚJO; ZULLO; TORRES, 2020). The distinction between the Brazilian cases and the COMPAS case is not merely technological: it reflects substantially different normative, institutional, and socio-political contexts, which requires caution in direct argumentative transposition.

2.5 International regulatory landscape: the European AI Act as reference and its transplantation limits

The debate on algorithmic governance in the public sector gained concrete normative contours with the approval of the European Union Artificial Intelligence Regulation (AI Act, 2024), the first comprehensive regulatory framework dedicated to AI in the world. The AI Act classifies AI systems by risk level — unacceptable, high, limited, and minimal — and imposes proportional obligations for each category. AI systems used in decisions on access to essential public services, risk assessment in security, and administration of justice are classified as high risk, subject to mandatory requirements of transparency, technical documentation, human oversight, and registration in a public database.

For the Brazilian context, the relevance of the AI Act is twofold. First, it serves as a reference for the construction of a national regulatory framework — a discussion that remains at an incipient stage in Brazil, with fragmented legislative initiatives and no approved legislation. Second, it directly affects technology companies that supply solutions to the Brazilian public sector, since many of them operate in the European market and will need to adapt their products to the regulation's requirements.

The comparison between the AI Act and the Brazilian legal order reveals both convergences and gaps. The efficiency principle (art. 37, CF/88), the duty to justify administrative acts (Law 9,784/1999), and the LGPD (Law 13,709/2018) form a normative basis that already includes requirements analogous to those of the AI Act in terms of transparency, accountability, and data protection. The critical gap

lies in the absence of a mandatory algorithmic impact assessment before the deployment of AI systems in public decision-making functions.

A relevant methodological caveat must be added: the importation of supranational regulatory frameworks to specific national contexts faces obstacles that go beyond formal normative compatibility. In the Brazilian case, the difference is not merely the absence of specific regulation, but also limited enforcement capacity, insufficient institutional coordination between oversight bodies, a lack of specialised techno-bureaucratic resources, and heterogeneity of capacity across federal entities. Having a rule and having the capacity to enforce the rule are analytically distinct dimensions. The governance model proposed in this study seeks to address this gap by incorporating, in the fifth axis, training and infrastructure requirements as preconditions for the deployment of AI systems.

2.6 Risks and limits of AI use by the state

The literature converges on a central understanding: the benefits of AI in public administration are only fully realised when accompanied by robust institutional safeguards. Araújo, Zullo and Torres (2020) identify four main risk categories: algorithmic opacity, discriminatory biases, democratic deficit, and threats to privacy and personal data protection.

Algorithmic opacity — frequently referred to as the *black box* problem — refers to the inability to explain, in comprehensible terms, the process by which an algorithm reached a given result. This characteristic directly conflicts with the principles of publicity and justification of administrative acts, enshrined in arts. 2 and 50 of Law 9,784/1999. In a Democratic State under the Rule of Law, the absence of transparency regarding the reasons for decisions compromises both judicial review and the democratic legitimacy of public decisions — and it is precisely this incompatibility that the European AI Act seeks to resolve by requiring explainability as a mandatory requirement for high-risk systems.

Discriminatory biases constitute another first-order risk. Since algorithms are trained on historical databases, automated decision-making tends to reproduce and amplify discrimination based on race, gender, income, and territory. The State of Wisconsin v. Loomis case is paradigmatic: the COMPAS tool was used in a criminal

sentence, and the defence argued a violation of due process given the impossibility of accessing the algorithmic reasons for the conviction. The state Supreme Court acknowledged limitations of the system, including the possibility of racial bias (ARAÚJO; ZULLO; TORRES, 2020). In a country with Brazil's structural inequalities, this risk is amplified and must guide any decision to deploy AI in high-impact functions affecting rights.

Digital exclusion constitutes a central paradox: while automation improves the experience of digitally included users, it may erect insurmountable barriers for the most vulnerable groups. In Brazil, tens of millions of citizens still lack regular internet access or basic digital skills. Toledo and Mendonça (2023) add that the unpreparedness of civil servants and the absence of specific regulation undermine the desired transformation. Although the LGPD establishes a relevant normative framework, its practical application still registers regulatory gaps and insufficient supervisory capacity, reinforcing the urgency of complementary mechanisms of algorithmic governance.

3. Methodology

This research adopts a qualitative approach of an exploratory and descriptive nature. It is an analytical narrative review — rather than a systematic review —, a modality which, as Mendes-da-Silva (2019) notes, enables broad and interpretive discussions on a specific topic from a critical synthesis of prior production. This choice implies recognising inherent reproducibility limits of the method: unlike systematic reviews, analytical narrative reviews do not require prior registration protocols nor guarantee sample exhaustiveness. Its advantage, however, is to permit the transversal integration of contributions from distinct areas — Administrative Law, Public Administration, Computer Science, and Political Science — which proved indispensable given the multidimensional complexity of the object.

Bibliographic data collection was conducted in the Web of Science, Scopus, and Google Scholar databases, following this procedure: (a) initial search by combining descriptors in Portuguese and English — *inteligência artificial AND administração pública*; *artificial intelligence AND public administration*; *algorithmic governance AND public sector*; *IA AND eficiência administrativa*; *AI AND citizen*

participation —, applied to the title, abstract, and keyword fields; (b) application of temporal filters (2017–2025), with the exception of classical reference works from earlier periods; (c) reading of the title and abstract of all returned results; (d) selection for full reading of studies that met the inclusion criteria; (e) backward reference tracking of documents cited in selected studies. The process resulted in a bibliographic corpus of 28 references, comprising peer-reviewed journal articles, legal doctrine books, and reference institutional documents. The normative documents — CF/88, CA 19/1998, CA 85/2015, Law 9,784/1999, Law 13,709/2018, Decree 10,332/2020, and the European AI Act (2024) — were included as an additional documentary corpus, used in the analysis as case evaluation criteria, and not merely as theoretical contextualisation.

The inclusion criteria for studies were: (i) directly addressing the relationship between AI and public management; (ii) published in a peer-reviewed scientific journal or a recognised reference institutional source; (iii) presenting empirical results or legal-normative analyses with explicit grounding. Excluded were opinion pieces without scientific support, commercial sources with evident conflicts of interest, and studies whose main focus was technology in private contexts without interface with the public sector.

Data were analysed through thematic content analysis. The four analytical axes — administrative efficiency; citizen participation; risks and limits; and parameters for responsible adoption — were constructed inductively from iterative reading of the corpus: the first two axes derive from the research question and the stated objectives; the third emerged from the convergence of risks identified in the algorithmic governance literature; and the fourth resulted from the synthesis between the empirical findings of the cases and the normative requirements identified in the national legal order and the European AI Act.

Case selection followed the criterion of deliberate typological representativeness: (a) cases of operational support in legal back-office, with low risk of directly affecting rights — Sapiens/AGU, Victor/STF, and Athos/STJ; (b) cases of external control and social control, with an accountability function — Alice/TCU and the Serenata de Amor Operation; (c) a citizen service case, with risk of digital exclusion — Gov.br; and (d) an international limit case, with high rights

impact and insufficient human oversight — COMPAS/USA. This typology enabled the systematic comparison developed in Table 2 (section 4.1), structured around six fixed variables: AI function, task type, risk level according to the AI Act classification, degree of human oversight, and potential impact on rights. Cases were described based on specialised secondary literature and institutional documentation. It must be acknowledged, as a limitation, that performance data for Brazilian systems comes predominantly from the implementing agencies themselves, without validation by independent empirical audit — which recommends caution when reading assertions about documented efficiency.

4. Results and Discussion

4.1 Adoption patterns and success conditions: comparative case analysis

The comparative analysis of the studied cases, summarised in Table 2, identifies three structural conditions that distinguish successful AI experiences in the Brazilian public sector from those presenting risks or unsatisfactory results. The matrix was built using fixed variables applied to each case, so as to reveal analytical patterns rather than merely descriptive narratives.

Table 2 – Comparative matrix of AI cases in the public sector

| Case | AI function | Task type | Risk level (AI Act) | Human oversight | Potential impact on rights |
|------------------------------|-------------------------------|---|-------------------------------|---|---|
| Sapiens/ Super Sapiens (AGU) | Legal operational support | Triage, classification, and drafting of legal documents | Limited (decision support) | Mandatory review by solicitors before any legal act | Indirect: improves quality of acts, but does not decide on rights |
| Victor (STF) | Procedural management support | Classification of documents and identification of general repercussion issues | Limited (preparatory support) | Law clerks retain responsibility for final classification | Indirect: accelerates triage, does not decide on merits |
| Athos (STJ) | Precedent management | Analysis of decisions for application of binding precedents | Limited (technical support) | Justices validate all algorithmic suggestions | Indirect: guides decisions, but does not replace them |
| Alice/TCU | External control and audit | Detection of irregularities in public procurement | Limited (alert) | Auditors validate leads before any notification | Indirect: enables legality control, without binding decisions |
| Serenata de Amor | Social control | Analysis of parliamentary expenses (CEAP) | Minimal (transparency) | Citizen as end user; automated decision | Positive: expands accountability without restricting rights |

| Case | AI function | Task type | Risk level (AI Act) | Human oversight | Potential impact on rights |
|---------------------------|---------------------------|---|----------------------------|---|---|
| Gov.br | Digital service delivery | Delivery of standardised services and information | Minimal to limited | User chooses channel; human review in complex cases | Indirect: broadens access, risk of digital exclusion |
| COMPAS (USA) — limit case | Judicial decision support | Criminal recidivism risk assessment | High (decision on liberty) | No structured human oversight; opaque algorithm | Direct and severe: informs criminal sentences with documented racial bias |

Source: Prepared by the authors based on Araújo, Zullo and Torres (2020), Bollotti and Wachowicz (2024), Desordi and Bona (2020), and AI Act (2024).

The first structural condition identified is delimited functional scope. All documented success cases apply AI to tasks with clear limits: triage, classification, pattern detection, and delivery of standardised information. None of them delegate the final decision on rights to the machine. This pattern contrasts directly with the COMPAS case, in which AI was used to inform criminal sentences — a high-impact, factually complex decision with direct implications for individual liberty —, resulting in judicial challenge on grounds of racial bias (ARAÚJO; ZULLO; TORRES, 2020). It should be emphasised that the COMPAS case operates in a normative and institutional context that differs substantially from the Brazilian one; its inclusion as a limit case serves to demarcate the boundary between acceptable and problematic AI uses, and not to establish functional equivalence with national cases.

The second structural condition is formalised human oversight as a mandatory protocol. In Sapiens/AGU, all AI-generated texts are reviewed by solicitors before any legal act. In Victor/STF, algorithmic results feed the work of law clerks, who retain final responsibility for classification. In TCU's robots, auditors validate leads before any notification. This pattern is consistent with art. 50 of Law 9,784/1999, which requires express justification in acts that deny, limit, or affect rights — justification which, by its nature, requires identifiable and accountable human authorship.

The third condition is integration with pre-existing control processes. Successful systems do not replace established workflows but expand them: Alice/TCU enhances the capacity to audit procurement for which legal procedures and defined institutional responsible parties already exist; Victor/STF organises processes that civil servants already managed. This integration ensures that pre-

existing legal accountability obligations also apply to algorithmically mediated outputs, avoiding the creation of accountability vacuums.

From a normative perspective, the results show that the Brazilian legal order already contains the foundations for responsible algorithmic governance — albeit in fragmented form. Art. 37, caput, CF/88 imposes the duty of efficiency; art. 5, LV, guarantees due process even in administrative proceedings; art. 2 of Law 9,784/1999 requires justification, proportionality, and legal certainty; and the LGPD establishes limits on the automated processing of personal data. What is lacking, unlike the European model, is a mandatory mechanism for algorithmic impact assessment prior to deployment — a gap that Decree 10,332/2020 does not fill, as it is limited to the digitalisation of services without addressing algorithmic governance.

4.2 AI and administrative efficiency: documented gains and capture risks

The analysed cases confirm that AI presents concrete potential to expand administrative efficiency in triage, classification, automation, and decision-support tasks. Super Sapiens operates 70% faster than the previous AGU system, according to data released by the institution itself. Victor, from the STF, reduces process triage time from dozens of minutes to a few seconds. The TCU's robots Alice and Mônica sweep public procurement databases with efficiency unattainable by conventional human control (ARAÚJO; ZULLO; TORRES, 2020). These results suggest compatibility with the constitutional efficiency imperative and, according to Bollotti and Wachowicz (2024), may indicate that, in contexts such as those described, AI adoption tends to be not only legitimate but legally justifiable — provided the three structural conditions identified are met.

These gains, however, carry a specific risk that the literature tends to underestimate: efficiency capture. AI projects deployed without rigorous technical assessment and without public performance metrics can generate technological dependencies that, in the future, translate into excessive costs and integration difficulties — a phenomenon documented by Araújo et al. (2025) in the context of the federal fragmentation of Brazilian digital government. The cultural resistance of civil servants — documented by Santos (2025) and Toledo and Mendonça (2023)

— constitutes a real and recurrent barrier: even when the tool is available and technically adequate, its adoption remains marginal in the absence of training and usage protocols. Furthermore, the asymmetry between federal entities means that efficiency gains verified in high-capacity federal bodies do not automatically replicate in smaller municipalities.

The evaluation of public AI systems must include measurable operational criteria, both *ex ante* and *ex post*. At the *ex ante* level, the following are recommended: (a) algorithmic impact assessment published before deployment, identifying risks by population groups; (b) pilot testing with limited functional scope; and (c) verification of minimum required state capacity — data infrastructure, civil servant training, and oversight mechanisms. At the *ex post* level, the relevant indicators include: reduction in mean processing time; contestation rate for AI-mediated decisions; demographic impact disparity; auditability of the decision-making process; and user satisfaction, both internal (civil servants) and external (citizens). The incorporation of these criteria into the proposed governance model is what distinguishes a principled list from a genuinely operational instrument.

4.3 AI and citizen participation: typology, evidence, and risks of algorithmic mediation

The analysis of AI-mediated citizen participation requires, as a preliminary step, a conceptual distinction that the public debate frequently neglects. Digitalisation of service delivery, informational transparency, AI-supported social control, and algorithmically mediated public deliberation are analytically distinct phenomena with different democratic implications. Conflating them obscures both the possibilities and the risks of each modality.

Digitalisation of service delivery refers to the migration of public services to digital channels, with or without AI. Gov.br and similar initiatives belong to this category: they broaden access and reduce transaction costs, but do not necessarily deepen citizens' voice in public decisions. Informational transparency consists of making government data available in a comprehensible and accessible form, and can be enhanced by AI through automatic synthesis, translation into plain language, and thematic organisation of large volumes of information. AI-supported social

control goes further: it allows citizens themselves to use algorithms to scrutinise acts of public agents. The Serenata de Amor Operation is the most established case in this modality, having analysed parliamentary expenses and identified irregularities with high accuracy (DESORDI; BONA, 2020). Algorithmically mediated public deliberation, finally, is the most ambitious and least developed modality: it refers to the use of AI to process contributions in public consultations, synthesise arguments, identify demand patterns, and present structured diagnostics to managers — converting participation from a symbolic formality into a real input for policy design (SCHIEFLER; CRISTÓVAM; PEIXOTO, 2020).

In the Brazilian context, experiences of algorithmically mediated public deliberation remain at an incipient stage. Public consultations conducted by federal bodies frequently receive volumes of contributions that make timely manual analysis impracticable — which renders algorithmic synthesis technically feasible and administratively relevant. The digital participatory budgeting agenda in states and municipalities presents similar potential: thematic classification algorithms could organise proposals, identify convergences, and present results transparently to decision-makers.

The risks of algorithmic mediation of citizen participation are, however, substantial and require specific attention. The first is that of prioritisation and invisibilisation: classification algorithms tend to amplify contributions from groups already over-represented digitally — urban, more educated, with internet access — and to undervalue the voices of minority, peripheral, or technologically less fluent groups. The second is the asymmetry of power between algorithm producers and affected citizens: the criteria for synthesis, prioritisation, and classification are defined by those who hold the technology, not by the participants. The third is deliberative opacity: if citizens do not know how their contributions were processed, filtered, and weighted, AI-mediated participation may function as symbolic legitimisation of decisions already made, rather than as a genuine channel of influence.

For AI to expand — rather than merely simulate — citizen participation, certain conditions are non-negotiable: intentional design oriented towards inclusion, with accessible interfaces for varied user profiles; multiple participation channels,

both digital and in-person; public disclosure of the algorithmic criteria used to process contributions; and mechanisms for citizens themselves to verify the synthesis results. Without these guarantees, AI-mediated participation tends to reproduce pre-existing access hierarchies, amplifying already-represented groups at the expense of historically marginalised populations.

4.4 Democratic and ethical risks: opacity, bias, and exclusion

The systematic analysis of cases and literature confirms that the risks of AI in public management are real, documented, and demand structured normative and institutional response. Algorithmic opacity is incompatible with the requirement to justify administrative acts under art. 50 of Law 9,784/1999. A public decision that affects citizens' rights without being able to explain its reasoning in intelligible terms violates the Democratic State under the Rule of Law, regardless of its technical efficiency — and it is precisely for this reason that the European AI Act makes explainability a mandatory requirement, not a recommendation.

Discriminatory bias has direct relevance for the Brazilian context. In a country marked by deep racial, regional, and socioeconomic inequalities, AI systems trained on historical data tend to reproduce and amplify these asymmetries. Algorithms for granting social benefits, triaging health services, or assessing risk in public safety can operate as instruments of institutional discrimination — formally neutral, but materially exclusionary. Araújo, Zullo and Torres (2020) warn that police actions and judicial decisions supported by algorithmically discriminatory processes attack human dignity, the right to equality, and individual liberty.

Digital exclusion constitutes a central paradox: while automation improves the experience of digitally included users, it may erect insurmountable barriers for the most vulnerable groups. The accelerated adoption of AI without parallel digital inclusion policies is not only ethically questionable — it is constitutionally problematic, as it violates the principle of equal access to public services.

4.5 Algorithmic governance guidance model: foundations, tensions, and applicability

The five-axis guidance model proposed in this study does not result from a bibliographic compilation of pre-existing principles, but emerges from the joint

analysis of three sources: (a) the patterns identified in the systematic comparison of cases, which revealed the three structural success conditions; (b) the normative requirements of the Brazilian legal order; and (c) the parameters of the European AI Act adapted to the national institutional context. Its distinction from previous proposals lies in the explicit articulation between each axis and a measurable evaluation indicator — which transforms the model from a principled list into an operational management instrument.

Table 1 – Algorithmic governance guidance model for Brazilian public administration

| Axis | Core directive | Legal basis | Evaluation indicator |
|---|--|--|---|
| 1. Legality and proportionality | Adopt AI only when the problem is clearly identified, the solution is appropriate, and benefits outweigh costs and risks | Art. 37, caput, CF/88; Art. 2, Law 9,784/1999; AI Act, art. 9 | Published algorithmic impact assessment prior to deployment |
| 2. Transparency and explainability | Disclose criteria, training data, error rates, and contestation mechanisms; ensure intelligible explanations for citizens | Art. 5, XXXIII, CF/88; Art. 50, Law 9,784/1999; AI Act, art. 13 | Explainability index audited annually |
| 3. Human oversight and accountability | Subject all decisions affecting rights to human review; maintain audit trail; accountability rests with the public manager | Art. 5, LV, CF/88; Art. 37, §6, CF/88; AI Act, art. 14 | Percentage of decisions reviewed by humans and mean response time to contestations |
| 4. Non-discrimination and data protection | Periodic bias audits; full compliance with LGPD; impact assessment on vulnerable groups before large-scale use | Art. 5, caput, CF/88; LGPD, arts. 20 and 44; AI Act, arts. 10 and 15 | Demographic disparity rate by group (semi-annual report) |
| 5. Training and inclusion | Continuous functional training for civil servants; assisted service for citizens without digital access; accessible interfaces for all user profiles | Decree 10,332/2020; AI Act, art. 4 (AI Literacy); Recital 20 | Percentage of trained civil servants and citizens served through alternative non-digital channels |

Source: Prepared by the authors based on Araújo, Zullo and Torres (2020), Bollotti and Wachowicz (2024), LGPD (2018), and AI Act (2024).

The first axis — legality and proportionality — requires that AI be adopted only when there is a clearly identified problem, an algorithmic solution appropriate to its nature, and measurable benefits that outweigh the costs and risks involved. The second axis — transparency and explainability — demands that the criteria used, training data, documented error rates, and contestation mechanisms be disclosed and intelligible to citizens. Explainability is not an optional requirement in a Democratic State under the Rule of Law — it is a validity requirement of administrative acts.

The third axis — human oversight and accountability — stipulates that every decision affecting rights be subject to human review, that the decision-making process be auditable, and that responsibility clearly fall on the public manager. The fourth axis — non-discrimination and data protection — mandates periodic bias audits, full compliance with the LGPD, and impact assessment on vulnerable groups as a condition for any large-scale use. The fifth axis — training and inclusion — establishes that continuous functional training for civil servants, assisted service for citizens without digital access, and inclusive interfaces are non-negotiable conditions.

It is important to acknowledge that the five axes are not automatically convergent: tensions may arise between them in concrete situations. The requirement of algorithmic transparency (axis 2) may conflict with data confidentiality obligations or the protection of strategic information (axis 4). Mandatory human oversight (axis 3) may reduce the efficiency gains that motivated the system's adoption (axis 1). The mass training of civil servants (axis 5) requires resources that may be scarce in contexts of budgetary constraint, creating pressure on the legality of deployment (axis 1). The recognition of these tensions is a condition for the responsible application of the model: in each concrete case, the public manager must carry out an explicit and grounded balancing exercise, documenting the criteria used to balance the conflicting axes. The model does not eliminate administrative discretion — it structures and renders it auditable.

5. Conclusion

This study started from the following research question: under what conditions can the adoption of AI solutions expand administrative efficiency and strengthen citizen participation in Brazilian public administration without violating fundamental rights, deepening structural inequalities, or generating a democratic deficit? The results allow a precise and empirically grounded answer: AI adoption tends to be constitutionally legitimate and administratively beneficial when three structural conditions are met — delimited functional scope, formalised human oversight, and integration with pre-existing control processes. When these

conditions are absent, the risks of algorithmic opacity, discriminatory bias, and digital exclusion outweigh the potential benefits.

The documented evidence confirms that AI is already an operational reality in Brazilian public administration. Cases such as Super Sapiens at the AGU, Victor at the STF, Alice at the TCU, and Gov.br demonstrate gains in time, quality, and uniformity in service delivery, compatible with the imperative of art. 37, caput, CF/88. The comparative analysis revealed that successful initiatives share a common normative denominator: in all of them, AI operates as a support instrument, and the ultimate responsibility for the administrative act remains with an identifiable public official — which preserves the justification of acts and the possibility of judicial review. It is important to emphasise, however, that part of the performance evidence comes from the institutional sources of the implementing agencies themselves, and not from independent empirical evaluation — which requires caution in the generalisation of findings.

In contrast, the COMPAS case and the analysis of Brazilian normative gaps reveal the limits of the current model. The absence of a mandatory algorithmic impact assessment — a central requirement of the European AI Act of 2024 — constitutes the main gap in the Brazilian legal order. The dialogue with the European regulatory framework shows that Brazil possesses sufficient normative foundations — art. 37, Law 9,784/1999, and the LGPD —, but lacks a specific mechanism to articulate them into a mandatory system of algorithmic governance. The importation of this regulatory model, however, cannot disregard the enforcement limits, federal heterogeneity, and unequal state capacity that characterise the Brazilian context.

The five-axis guidance model proposed offers an operational roadmap for managers and policy-makers. Its specific contribution lies in the articulation between legal basis and measurable evaluation indicator, as well as the explicit recognition of tensions between axes — which distinguishes it from generic principled proposals and makes it an instrument of structured, auditable, and accountable discretion. Constitutional efficiency is not measured by the speed of processes, but by the quality, transparency, and inclusiveness of decisions.

The limitations of this study must be explicitly acknowledged. First, the bibliographic and documentary nature of the research implies dependence on secondary and institutional sources, without direct empirical validation of the performance, oversight, and decision quality claims regarding the described systems. The efficiency data cited comes predominantly from the implementing agencies themselves, which creates a risk of institutional confirmation bias. Second, the studied cases are concentrated in federal and large-scale bodies with high state capacity; results are not automatically extrapolable to subnational governments of lesser capacity. Third, the temporal scope (2017–2025) may not capture recent developments in AI regulation in Brazil. Fourth, case selection, although deliberately typological, is not exhaustive: there are state and municipal experiences still underinvestigated in the literature. Fifth, the study does not include primary data on citizens' perceptions, user satisfaction, or the distributive impact of the analysed systems — a gap that the future research agenda must prioritise.

As a future research agenda, the following are suggested: (i) case studies on AI adoption in subnational governments, with emphasis on municipalities of lesser institutional capacity; (ii) investigations with primary data on citizens' perceptions and trust in algorithmically mediated decisions; (iii) comparative empirical performance evaluations of public AI systems, conducted by independent auditors; (iv) analysis of Brazilian AI legislative proposals under consideration and their compatibility with the AI Act; and (v) studies on the distributive impact of public service digitalisation on vulnerable groups. In the twenty-first century, good administration begins by asking the right questions — and only then chooses the algorithms.

References

ANESP – ASSOCIAÇÃO NACIONAL DOS ESPECIALISTAS EM POLÍTICAS PÚBLICAS E GESTÃO GOVERNAMENTAL. Inteligência artificial e gestão pública: mundo em transformação. Brasília: Anesp, 2023.

ARAÚJO, Alan Bruno da Silva; BRITO, Bruna Ohana Silva; GUIMARÃES, Victor Silva; OLIVEIRA, Francisco Mesquita de; CASTRO, Maurício Mendes Boavista de. Transformação digital no setor público brasileiro: uma abordagem teórica sobre o impacto inovador e transformador de sua adoção. *Revista Multidisciplinar do Nordeste Mineiro*, v. 11, 2025. DOI: 10.61164/rmm.v11i1.4037.

ARAÚJO, Valter Shuenquener de; ZULLO, Bruno Almeida; TORRES, Maurílio. Big data, algoritmos e inteligência artificial na Administração Pública: reflexões para a sua utilização em um ambiente democrático. A&C – Revista de Direito Administrativo & Constitucional, Belo Horizonte, v. 20, n. 80, p. 241-261, abr./jun. 2020. DOI: 10.21056/aec.v20i80.1219.

BOLLOTTI, Joelson Júnior; WACHOWICZ, Marcos. A aplicação da inteligência artificial pela administração pública diante do princípio da eficiência. Revista da AGU, Brasília-DF, v. 23, n. 4, dez. 2024. DOI: 10.25109/2525-328X.v.23.n.4.2024.3429.

BRASIL. [Constituição (1988)]. Constituição da República Federativa do Brasil de 1988. Brasília, DF: Presidência da República, 1988.

BRASIL. Decreto n.º 10.332, de 28 de abril de 2020. Institui a Estratégia de Governo Digital para o período de 2020 a 2022. Diário Oficial da União, Brasília, DF, 29 abr. 2020.

BRASIL. Emenda Constitucional n.º 19, de 4 de junho de 1998. Modifica o regime e dispõe sobre princípios e normas da Administração Pública. Diário Oficial da União, Brasília, DF, 5 jun. 1998.

BRASIL. Emenda Constitucional n.º 85, de 26 de fevereiro de 2015. Atualiza o tratamento das atividades de ciência, tecnologia e inovação na Constituição Federal. Diário Oficial da União, Brasília, DF, 27 fev. 2015.

BRASIL. Lei n.º 9.784, de 29 de janeiro de 1999. Regula o processo administrativo no âmbito da Administração Pública Federal. Diário Oficial da União, Brasília, DF, 1 fev. 1999.

BRASIL. Lei n.º 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Diário Oficial da União, Brasília, DF, 15 ago. 2018.

CAMPILLO, Laura Pérez; HONORATO, Cláudio Belmonte de Athayde. Tecnologia e aplicativos no combate à COVID-19 e na proteção de dados pessoais na Espanha e no Brasil. Revista General de Derecho Administrativo, n. 56, 2021.

CEPAL; CAF. As oportunidades da digitalização na América Latina frente ao COVID-19. Santiago: CEPAL, 2020.

CONEGLIAN, Caio Saraiva; SEGUNDO, José Eduardo Santarem. Inteligência artificial e ferramentas da web semântica aplicadas à recuperação da informação: um modelo conceitual com foco na linguagem natural. Informação & Informação, Londrina, v. 27, n. 1, p. 625-651, 2022. DOI: 10.5433/1981-8920.2022v27n1p625.

DA SILVA, W. F.; SILVA, F. S.; RABÊLO, O. da Silva. Tendências no uso de inteligência artificial e sua influência na requalificação da força de trabalho no setor público. Cadernos de Prospecção, Salvador, v. 14, n. 3, p. 808-824, 2021. DOI: 10.9771/cp.v14i3.36727.

DESORDI, Danubia; BONA, Carla Della. A inteligência artificial e a eficiência na Administração Pública. Revista de Direito, Viçosa, v. 12, n. 2, p. 1-22, 2020. DOI: 10.32361/202012029112.

DINIZ, Eduardo; BARBOSA, Alexandre; JUNQUEIRA, Alvaro; PRADO, Otávio. O governo eletrônico no Brasil: perspectiva histórica a partir de um modelo estruturado de análise. Revista de Administração Pública, Rio de Janeiro, v. 43, n. 1, p. 23-48, 2009.

FARINIUK, Tássia Menezes Débora. Smart cities e pandemia: tecnologias digitais na gestão pública de cidades brasileiras. Revista de Administração Pública, Rio de Janeiro, v. 54, n. 4, p. 860-873, ago. 2020.

FREITAS, Juarez. Discricionariedade administrativa e o direito fundamental à boa Administração Pública. São Paulo: Malheiros, 2007.

HACHEM, Daniel Wunder; GABARDO, Emerson. El principio constitucional de eficiencia administrativa: contenido normativo y consecuencias jurídicas de su violación. *Cuestiones Constitucionales: Revista Mexicana de Derecho Constitucional*, Ciudad de México, n. 39, p. 3-38, jul./dez. 2018.

KAUFMAN, Dora; SANTAELLA, Lucia. O papel dos algoritmos de inteligência artificial nas redes sociais. *Revista FAMECOS: Mídia, Cultura e Tecnologia*, Porto Alegre, v. 27, n. 1, p. e34074, 2020. DOI: 10.15448/1980-3729.2020.1.34074.

McCARTHY, John. *What is Artificial Intelligence?* Stanford, CA: Stanford University, 2007. Available at: <https://www-formal.stanford.edu/jmc/whatisai.pdf>.

MENDES-DA-SILVA, Wesley. Contribuições e limitações de revisões narrativas e revisões sistemáticas na área de negócios. *Revista de Administração Contemporânea*, Maringá, v. 23, n. 2, p. 1-11, mar. 2019.

MINISTÉRIO DE CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES (MCTIC). *Estratégia Brasileira para a Transformação Digital (E-Digital)*. Brasília: MCTIC, 2018.

OCDE. *Recomendação do Conselho sobre Inteligência Artificial*. Paris: OECD Publishing, 2023. Available at: <https://oecd.ai/en/ai-principles>.

SANTOS, André Roberto Ramos dos. Inteligência artificial na administração pública: estudo de caso sobre o Chat-GPT como ferramenta de desenvolvimento no poder legislativo. *Práticas de Administração Pública*, Santa Maria, v. 9, e89377, 2025. DOI: 10.5902/2526629289377.

SCHIEFLER, Eduardo André Carvalho; CRISTÓVAM, José Sérgio da Silva; PEIXOTO, Fabiano Hartmann. A inteligência artificial aplicada à criação de uma central de jurisprudência administrativa. *Revista do Direito*, Santa Cruz do Sul, n. 60, p. 18-34, 2020. DOI: 10.17058/rdunisc.v3i50.14981.

SPANÓ, Eduardo; SILVA, Viviane. Inovação do serviço público no Brasil pós-covid-19: plataformas digitais nos níveis da administração pública. *Revista Tecnologia e Sociedade*, Curitiba, v. 18, n. 52, p. 324-348, jul./set. 2022.

TOLEDO, Adriana Teixeira de; MENDONÇA, Milton. A aplicação da inteligência artificial na busca de eficiência pela Administração Pública. *Revista do Serviço Público*, Brasília, v. 74, n. 2, p. 410-438, abr./jun. 2023. DOI: 10.21874/rsp.v74i2.6829.

UNIÃO EUROPEIA. Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial). *Jornal Oficial da União Europeia*, L 1689, 12 jul. 2024.