

The State That Cannot See Its Own Tools: Algorithmic Opacity and the Limits of Transparency Law in Brazilian Public Administration

Mayara Cristhina Rodriguero dos Santos¹, Alexandre Pacheco da Silva², Marina Feferbaum³

¹ Licensed Lawyer, Brazil

² Professor of Digital Law and Intellectual Property, FGV São Paulo Law School (Escola de Direito de São Paulo da Fundação Getulio Vargas), Brazil

³ Professor, FGV São Paulo Law School (Escola de Direito de São Paulo da Fundação Getulio Vargas) and Coordinator of the Center for Education and Research in Innovation (CEPI), Brazil

Correspondence: Alexandre Pacheco da Silva, Professor of Digital Law and Intellectual Property, FGV São Paulo Law School (Escola de Direito de São Paulo da Fundação Getulio Vargas), Brazil.

Received: March 23, 2026

Accepted: May 4, 2026

Available online: May 11, 2026

doi:10.11114/ijsss.v14i2.8814

URL: <https://doi.org/10.11114/ijsss.v14i2.8814>

Abstract

This article examines the transparency and accountability challenges raised by the use of AI systems in Brazilian public administration, using the Smart Sampa facial recognition program in the city of São Paulo as a central case study. Drawing on the work of Lessig (1999) and Crootof & Ard (2020) and integrating the multi-level transparency framework developed in public administration scholarship (Giest & Grimmelikhuijsen, 2020; Meijer & Grimmelikhuijsen, 2020), the article diagnoses Smart Sampa's opacity at three distinct levels: algorithmic, organizational, and institutional. It then examines six regulatory frameworks through the same analytical lens, identifying structural gaps common to the EU AI Act, the French Digital Republic Act, the Indian AI Governance Guidelines, and the three most relevant Brazilian legislative proposals. Integrating empirical evidence from citizen preference research (Feng & Chandra, 2025) and experimental studies on institutional trust (Grimmelikhuijsen & Meijer, 2022; Wenzelburger et al., 2024), the article argues that passive transparency mechanisms, whether algorithm registers, explanation rights, or formal access to information channels, cannot in themselves generate the accountability needed to govern high-risk AI systems in Brazilian public security contexts. The article proposes an institutional framework requiring pre-deployment review by a specialized algorithmic gatekeeper within the Judiciary, mandatory multi-level disclosure obligations, administrative appeals for affected individuals, and formal override of trade secrecy claims against gatekeeper access requests. The proposals do not require new primary legislation but the amplification of roles and powers already available in the Brazilian legal system.

Keywords: artificial intelligence, public administration, law enforcement, algorithmic transparency, trade secrecy, facial recognition, Smart Sampa, accountability

1. Introduction

In November 2023, the city of São Paulo launched the Smart Sampa program, announcing the integration of approximately 40,000 surveillance cameras across the city equipped with AI-powered facial recognition capabilities (Prefeitura de São Paulo, 2023). The program was presented to the public as a public security efficiency tool: faster suspect identification, broader spatial coverage, and a technology-driven response to persistent urban crime. What the public announcement did not provide, and what subsequent access to information requests have largely failed to produce, was any substantive information about how the system works, which algorithm underpins it, what its documented false positive and false negative rates are, how it performs across different demographic groups, how the resulting outputs are used by police officers, or which institutional body is empowered to monitor its operation.

Smart Sampa is not an anomaly, it is the paradigmatic expression of a broader pattern documented across Brazilian law enforcement: AI systems acquired without public bidding, deployed without proactive disclosure, operated without accessible channels for contestation, and justified by reference to public security secrecy and trade secrecy arguments that have been consistently used to shield algorithmic governance from scrutiny (Arroyo & Gaspar, 2021). In 2019, two men were wrongly arrested following a false positive generated by a facial recognition system during the World Cup in

Rio de Janeiro. In 2023, a Black man spent 26 days in pretrial detention in Bahia because of an automated decision system error (Alencar, 2023). In neither case was the affected individual given access to information about the system that contributed to their detention.

The opacity that characterizes Smart Sampa and comparable Brazilian programs is not primarily a technical phenomenon. It is a legal and institutional one. Brazilian law already establishes a robust formal architecture for transparency in public administration: the Access to Information Law (*Lei de Acesso à Informação* (LAI)) requires proactive and reactive disclosure of public information; the 1988 Federal Constitution enshrines publicity as a constitutive principle of public administration in its art. 37; the Brazilian Civil Framework for the Internet (*Marco Civil da Internet* (MCI)) establishes transparency principles for digital governance; and the Brazilian General Data Protection Law (*Lei Geral de Proteção de Dados* (LGPD)), despite its public security exemption, creates a framework for automated decision transparency in other contexts. What these instruments lack is the specification and enforcement architecture necessary to give transparency obligations practical effect when applied to algorithmically mediated public security activities.

This article asks how that gap can be closed. The research questions are: how do the transparency challenges raised by AI use in law enforcement, specifically in the Smart Sampa case, illuminate the legal uncertainties posed by AI implementation across Brazilian public administration, and what institutional and regulatory responses are adequate to address them? The analysis proceeds through three connected moves: a multi-level diagnosis of Smart Sampa's opacity; a comparative examination of existing and proposed regulatory frameworks against that diagnostic; and a set of concrete institutional proposals designed to address the gaps the comparative analysis reveals.

The theoretical framework draws on Lessig's (1999) argument that code is not a natural phenomenon but a political one, capable of being regulated by law in the service of public values. This argument is extended through Crootof & Ard's (2020) techlaw typology, which identifies application uncertainty, normative uncertainty, and institutional uncertainty as the three structural conditions that generate legal opacity around new technologies. The multi-level transparency framework developed in public administration scholarship (Giest & Grimmelikhuijsen, 2020; Meijer & Grimmelikhuijsen, 2020) provides the analytical scaffold for diagnosing algorithmic opacity across three dimensions: the specific system (micro), its organizational embedding (meso), and the institutional rules that govern it (macro). The proposals are grounded in Kaminsky's (2020) algorithmic accountability framework, combined with empirical evidence from citizen preference and institutional trust research.

The article proceeds as follows. Section 2 presents the Smart Sampa case and diagnoses its transparency deficits using the multi-level framework. Section 3 analyzes the transparency dilemma in Brazilian public administration, focusing on the structural obstacles created by public security secrecy and trade secrecy regimes. Section 4 examines the theoretical framework governing the relationship between law, code, and regulation. Section 5 analyzes six regulatory frameworks through the multi-level lens. Section 6 presents institutional proposals. Section 7 concludes.

2. Smart Sampa and the Anatomy of Algorithmic Opacity

The Smart Sampa represents the most extensive deployment of AI-based surveillance infrastructure in Brazilian public security to date. The program, contracted between the city of São Paulo and the private technology company Sentinel X, contemplates the deployment of 40,000 cameras across the municipal territory, linked to a centralized platform equipped with facial recognition capabilities, crowd monitoring, and vehicle tracking (Prefeitura de São Paulo, 2023). In October 2025, the program was expanded to integrate private surveillance totems from companies not listed nor participants of any public bidding process (Helder, 2025). Its stated objective is to enhance public security through real-time identification of individuals sought by law enforcement, missing persons, and the automated detection of suspicious behavior patterns in public spaces.

The program raises the full spectrum of legal uncertainties identified by Crootof & Ard (2020) in their typology of techlaw. Application uncertainty arises because the existing Brazilian regulatory framework, including the LGPD, LAI, and MCI, was not designed with AI-mediated real-time biometric surveillance in mind, leaving fundamental questions about which and how rules apply unanswered. Normative uncertainty arises because the program's deployment increases the opacity of public administration actions in contexts where fundamental rights are directly implicated, producing what Crootof & Ard (2020, p. 369) name underinclusiveness: a legal condition in which existing norms are insufficient to address the harms generated by a new technology. Institutional uncertainty arises because no specialized body exists with the technical capacity and legal mandate to oversee the program's operation, assess its compliance with constitutional and legal requirements, or provide remedies to individuals affected by its outputs.

The research on AI accuracy in facial recognition contexts is unequivocal as these systems generate false positives at rates that fall disproportionately on people of color, on women, and on younger individuals (Klare et al., 2012; Buolamwini & Gebri, 2018). In the specific conditions of Brazilian public security, where structural racism is well

documented and where police violence against Black communities is a matter of historical record, the deployment of facial recognition technology without disclosed performance metrics, without demographic accuracy disaggregation, and without independent oversight is not merely a technical risk, it is a constitutional one. As Leese (2022, p. 504) argues, discrimination in data-driven profiling is often not a system error but the functional logic of the program itself, reflecting the patterns embedded in the historical data on which it was trained.

The Smart Sampa case is assessed below using the three-level transparency framework derived from Giest & Grimmelikhuijsen (2020) and Meijer & Grimmelikhuijsen (2020). Each level identifies a distinct dimension of the program's opacity and the specific rights implications that opacity generates.

Table 1. Smart Sampa assessed against the multi-level transparency framework

Level	Definition	Assessment of Smart Sampa	Rights implications
Micro (Algorithmic)	Information about how the specific AI system functions: model architecture, training data, false positive and negative rates, output logic, explainability mechanisms.	There is no public information on algorithm type or its decision logic. Its source code remains undisclosed, and no explanation mechanism is available to affected individuals.	Affected individuals cannot challenge the automated decisions, nor oversight bodies can assess its accuracy. Potential bias against marginalized groups remains undetected and uncorrected (Citron, 2008; Buolamwini & Gebru, 2018). High error rates presented in the Transparency Report (e.g. 64% due to outdated judicial data) prove that human checks are not sufficient to prevent mistakes in face of an underlying social-technical structure that it is flawed by itself.
Meso (Organizational)	Information about the organizational embedding of the AI: procurement process, internal governance policies, expertise involved, operator training, monitoring and evaluation practices.	Smart Sampa's Transparency Report (2025) presents the Standard Operational Process for law enforcement; however, it is reactive: the agent only validates the algorithmic output after a suspect has already been flagged, which can often lead to automation bias. No additional information or explanation is available.	Due process obligations are unenforceable without clear documentation of how algorithmic outputs become police actions: middle-man opacity shields providers from accountability (Arroyo & Gaspar, 2021).
Macro (Institutional)	Information about the institutional rules governing AI use: applicable legislation, external oversight mechanisms, regulatory compliance procedures.	Absence of a specialized gatekeeper or dedicated external oversight body. LAI requests produce generic or null responses and LGPD public security exemption (art. 4, III) is invoked to prevent access to information. Adjustments made in the algorithm (e.g. Smart Sampa program raising technical parameters to 92%) are done without independent audit.	Formal legal rights exist but have no practical effect: the absence of an active overseer means compliance with constitutional and legal obligations cannot be verified or enforced (Grimmelikhuijsen & Meijer, 2022). Transparency is limited to post-reporting, restraining possibility of accountability and error prevention before algorithmic deployment.

Description: This table evaluates the Smart Sampa program's transparency deficits at the algorithmic, organizational and institutional levels. It identifies specific gaps, such as undisclosed source code and reactive validation, and maps these to their direct impact on the fundamental rights of citizens.

As Table 1 illustrates, Smart Sampa's opacity operates simultaneously at all three levels. This simultaneity is not coincidental, it reflects the structural dynamics that Arroyo & Gaspar (2021) documented across Latin American law enforcement AI deployments: governments acquire systems from private providers under contractual arrangements that preclude public disclosure, invoke public security secrecy and trade secrecy arguments to deny access requests, and

operate in a regulatory vacuum that provides no institutional actor with the mandate to pierce these defenses. Understanding why this structure persists, and what legal resources already exist to challenge it, requires examining the transparency dilemma at its root.

3. The Transparency Dilemma: Public Security Secrecy and Trade Secrecy

In liberal democratic constitutions, transparency is not merely a policy value, it is a condition of legitimacy. In Brazil, the right to access information is enshrined in art. 5, inc. XIV of the 1988 Federal Constitution, and the principle of publicity is established as a mandatory obligation of all public administration in art. 37. The LAI gives these constitutional principles operational form, establishing both active transparency obligations (proactive disclosure of information of general public interest) and passive transparency channels (citizen-initiated access requests). The symbolic success of the LAI, with a federal response rate of 99.66% as of April 2024, tends to obscure its substantive limitations: public authorities routinely comply formally with information requests while providing information so generic or incomplete as to be analytically meaningless (Arroyo & Gaspar, 2021). Response is not disclosure.

The Smart Sampa program illustrates how public administration treats transparency as a formal exercise instead of a substantive safeguard: details about how the system works and how is deployed are published in biannual reports that lack the depth needed for independent audit over the algorithm's accuracy. For example, the October 2025 report revealed that 82 individuals were wrongly detained and conducted to police districts after being flagged by the system. The justification for these false positives was the outdated data in the Brazilian national warrant bank, with no additional explanation allowing further audit over these errors.

The abovementioned incident also proves the failability of the "human-in-the-loop" argument. Smart Sampa's reports contain operational guidelines mandating how law enforcement agents must interact with the system and apply its outputs. Police is required to "manually validate" alerts before taking any type of action. However, agents seem to be affected by automation bias, that is, the tendency to treat algorithm's output as an indisputable fact rather than a statistical probability. These incidents and risks could be significantly reduced by an independent, ongoing oversight as proposed in this paper.

In the specific context of AI in law enforcement, two legal arguments have become the primary instruments of substantive opacity. The first is public security secrecy. The argument proceeds as follows: disclosing detailed information about AI systems used for surveillance and suspect identification would enable bad actors to evade detection, compromise ongoing operations, and reduce the effectiveness of law enforcement. Zarsky (2013, p. 1553-1554) calls this the 'game the system' argument, and it is not entirely without merit: algorithmic transparency in adversarial contexts can produce incentives for avoidance behavior. However, as Zarsky (2013) himself observes, the argument has routinely been applied far beyond the narrow circumstances where it is defensible, functioning as a blanket justification for opacity across the entire spectrum of AI deployment rather than a proportionate and case-specific limitation on disclosure. Secrecy claims are often broader than required, and secrecy about the nature of government powers, how they are exercised, and why they are justified diminishes democratic accountability in ways that are not offset by any genuine operational security benefit (Moses & Koker, 2017, p. 10).

The second instrument of opacity is trade secrecy. Many AI systems used by Brazilian public administration are acquired from private technology providers under contractual arrangements that treat the system's source code, parameters, training data, and validation results as confidential commercial information protected under Brazil's Industrial Property Law (Law 9.279/1996). The LGPD reinforces this structure in its art. 20, para. 1, which requires that access to automated decision system information observe specifically established rights that include trade secrecy protection, while its art. 4, inc. III excludes public security applications from the law's scope entirely. When public authorities respond to access requests by invoking trade secrecy, they are frequently not making this argument on their own account as they are reproducing arguments made to them by their technology providers, who retain contractual control over information about the systems the public authority has deployed.

The legal problem created by this structure is not subtle: the refusal to produce the code of automated decision-making systems allows agencies to enforce laws that no one can see or monitor, which is the very antithesis of open government (Citron, 2008, p. 1293). Wexler (2018) developed this argument in the context of American criminal proceedings, demonstrating through a series of appellate cases that trade secrecy has been used to shield AI systems from defendants' discovery rights even in capital cases, and arguing that this represents a categorical subordination of liberty interests to commercial ones. In Brazil, as in the cases Wexler examined, courts have not yet developed a consistent doctrine governing the conflict between trade secrecy and the transparency obligations of public law. The result is structural uncertainty that systematically favors opacity, because the party with the strongest interest in nondisclosure, the private technology provider, is also the party with the most complete information about what is being concealed.

Understanding this structure points to the inadequacy of purely legislative solutions since governments are large

purchasers of AI systems and have the leverage to impose transparency requirements as procurement conditions (Desai & Kroll, 2017). The challenge is that exercising this leverage requires institutional capacity, an entity with both the technical expertise to specify meaningful transparency requirements and the legal authority to enforce them. Identifying what regulatory frameworks have done, or failed to do, to build such capacity is the task of the following section.

4. Theoretical Framework: Code, Law, and the Structuring of Techlaw

The relationship between law and code is not one of external regulation applied from above. Lessig's foundational argument in 'The Law of the Horse: What Cyberlaw Might Teach' (1999) is that code is itself a form of regulation, one that operates through the architecture of systems rather than the formal enactment of legal norms. Code writers legislate: they determine what inputs are accepted, what outputs are generated, what can be logged, what can be audited, and what remains opaque by design. The asymmetry between code's regulatory power and law's regulatory capacity in technologically complex domains is not a permanent feature of the landscape, it is a consequence of design choices that law can, in principle, reverse.

For the purposes of this article, Lessig's argument has three specific implications. First, the opacity of AI systems deployed in Brazilian public security is not a natural or inevitable consequence of technical complexity, it is the product of design and contractual choices that could have been made differently and that can be regulated otherwise. Second, the interaction between law and code must account for the other constraints Lessig identifies, particularly market forces: the commercial logic of AI development systematically incentivizes opacity, because the proprietary value of AI systems depends on the difficulty of replication, and replication difficulty is maintained through nondisclosure. Third, the implication for regulatory design is that law should aim to change the code, not merely to impose post-hoc sanctions for its misuse.

The Smart Sampa program provides a concrete example of Lessig's theory in action. While Lessig argues that law should regulate code to reflect and protect public values, Smart Sampa shows the opposite: how Brazilian public administration currently undermines the law, allowing the code to legislate through technical and legal opacity. As mentioned before, Smart Sampa system's does more than recognizing faces, it also performs crowd monitoring and detects suspicious behavior patterns. As we lack access to the algorithm and its weights, the meaning behind these terms remains a secrecy. The code is legislating by defining what counts as suspicious behavior without a public debate or oversight, which is a dangerous shift in face of the history of police oppression and racism in Brazil.

Although Smart Sampa is marketed as a neutral security tool, it is not free from political or social bias. When the program reported that 82 individuals were wrongly detained and conducted to police districts after a system alert, officials blamed the error on an outdated database, framing it as a purely "technical" mistake. This explanation ignores how the system's design and the manual police checks (as required by their operational guidelines) can reinforce social prejudices. For example, would the system flag and agents act over a white man standing near a car for a long time as "suspicious" in the same way that it might happen with a black man?

Crootof & Ard's (2020) techlaw framework provides a granular analytical vocabulary for the specific legal uncertainties generated by the dynamic describe above. Application uncertainty arises when it is unclear how existing law applies to a new technology; normative uncertainty arises when existing law produces outcomes that conflict with established values; institutional uncertainty arises when it is unclear which legal actors have the authority and competence to govern the technology. The Smart Sampa program raises all three forms of uncertainty simultaneously, as documented in Section 2. Crootof & Ard (2020, p. 384) recommend a precautionary approach as the appropriate regulatory stance in such conditions: a pro-regulation disposition that favors protective intervention when the potential for rights deprivation is high and the capacity for ex-post correction is limited.

The multi-level transparency framework (Giest & Grimmelikhuijsen, 2020; Meijer & Grimmelikhuijsen, 2020) operationalizes this precautionary orientation by specifying the three levels at which transparency obligations must operate to be effective. The framework distinguishes between information about the algorithm itself (micro), information about the organizational structures surrounding its use (meso), and information about the institutional rules governing both (macro). The framework's central empirical contribution is the demonstration that transparency at any single level, absent the others, fails to produce meaningful accountability. A legal framework that requires disclosure of algorithm documentation but not of organizational practices produces information without context. A framework that establishes institutional oversight without access to technical information produces oversight without substance. Effective algorithmic governance requires all three levels to be addressed simultaneously.

The accountability framework developed by Kaminsky (2020) complements this structural analysis by specifying what meaningful accountability requires at the level of individual rights. Kaminsky identifies three concerns that algorithmic accountability must address: an instrumental concern with preventing biased outcomes; a dignitary concern with protecting individuals from reduction to data profiles; and a justificatory concern with ensuring that AI-driven decisions

can be examined, challenged, and revised. The justificatory concern is the most legally demanding: it requires not only that information be technically accessible but that it be presented in forms that allow affected individuals, their advocates, and oversight institutions to assess the logic and validity of the decisions being made on its basis.

5. The Regulatory Landscape: Transparency Frameworks and Their Limits

The comparative analysis of regulatory frameworks in this section applies the multi-level transparency framework developed in Section 4 to six regulatory instruments: the EU AI Act, the French Digital Republic Act, the Indian AI Governance Guidelines, PL 2338/2023, PL 4612/2019, and PL 846/2026. The analysis asks, for each framework, which levels of the transparency architecture it addresses and which it leaves inadequately covered.

5.1 The EU AI Act: Institutional Architecture and Structural Limits

The EU AI Act, approved in February 2024, is the most comprehensive AI regulatory framework yet enacted. Its risk-based architecture subjects high-risk AI systems, including facial recognition technologies deployed in publicly accessible spaces and AI systems used in criminal investigation and prosecution, to obligations covering risk management, data governance, technical documentation, automatic logging, and transparency to deployers. Chapter 4 establishes notifying authorities and competent bodies empowered to audit high-risk AI systems and verify conformity with the Act's requirements. Article 23 requires providers of high-risk systems to furnish competent authorities with all documentation necessary to demonstrate conformity upon reasoned request.

From the perspective of the multi-level framework, the Act's most significant contribution is macro-level: it creates an institutional architecture for algorithmic governance that goes beyond passive disclosure requirements. The requirement of pre-market conformity assessment, combined with the audit powers of notifying bodies, represents a structural shift toward proactive rather than reactive oversight. Nevertheless, three limitations are significant for the Brazilian context.

First, the Act's transparency requirements operate primarily between regulated entities and notifying bodies, not between those entities and the public or affected individuals. Second, source code and model parameters are specifically protected from disclosure in audit contexts under Article 70, and the audit is reactive by design. As Simsek (2021) argues, high-risk AI systems used in law enforcement will typically have been in operational use for extended periods before a complaint triggers investigation, by which time substantial harm may have been generated without remedy. Third, the Act does not resolve the conflict between trade secrecy and defendants' discovery rights in criminal proceedings, a conflict that Wexler (2018) identified as structurally irresolvable within an intellectual property framework that privileges commercial interests over due process claims.

5.2 The French Digital Republic Act: Public Law Primacy

The French Digital Republic Act of 2016 addresses the trade secrecy problem through a different mechanism. By establishing the common information domain, the Act treats information produced or used by public authorities and their contracted private service providers as public information subject to the transparency obligations of administrative law. Private entities that provide technology services to public administration assume the transparency obligations of the public law regime. Article 8 explicitly overrides intellectual property protections, including trade secrecy, with respect to information falling within the domain of public interest.

The doctrinal logic of the French Act is directly applicable to the Smart Sampa situation. In Brazilian terms, the same argument can be constructed from existing constitutional and administrative law: entities that benefit from public contracts and provide services that affect citizens' fundamental rights are not merely commercial actors but participants in the public function, subject to the constitutional principles of art. 37. The French Act gives this argument a specific legislative form that has no Brazilian equivalent, but the normative foundation it rests on is available within the Brazilian constitutional order.

The French Act's limitation is institutional rather than doctrinal. It expands the category of public information but does not create the technical capacity to audit it, the institutional mandate to investigate discrepancies, or the specialized expertise to interpret what is disclosed. The empirical literature on algorithm registers in analogous transparency-forward legal contexts suggests that formal disclosure requirements without active institutional oversight tend to produce what the literature has termed a meaningful box-ticking exercise: registers that record algorithmic systems without enabling genuine scrutiny of them (Cath & Jansen, 2022; Floridi, 2020).

5.3 India's AI Governance Guidelines: Accountability and Transparency

In 2025, India published its AI Governance Guidelines aiming to foster innovation while ensuring safety, trust and inclusion. Unlike the EU AI Act's focus on new, horizontal legislation, India prioritizes adaptation. This approach leverages updating existing laws (such as 2000 IT Act, 2023 Digital Personal Data Protection Act, and 2021 IT Rules with its amendment on 2026) to address AI-related harms. A central pillar of this strategy is the National AI Incidents

Database, a public platform designed to record and disclose operational failures reported by developers and private companies. By mandating an “accountability trail”, the Guidelines focus on documenting incidents and real-world mitigation rather than relying solely on a risk-classification framework. Through a system of grievance redressing mechanisms, it enables India’s public administration to use complaint data as a tool to identify systemic biases and regulatory gaps.

This framework moves beyond treating transparency as a static duty to disclose information. Instead, it pushes for an active institutional role to resolve operational harms and risks as they arise. Compared to the European and French models discussed above, India’s approach helps bridge the gap between formal transparency and substantive evidence of how AI impacts individual rights. Such orientation is new and would require a data-to-mitigation change of view for Brazilian public administration, at the same time would assure timely protection over fundamental rights.

The Indian model has limitations. First, it relies on the capacity of Indian public bodies to process the grievances and act in an efficient and timely manner, which can create bureaucracy bottlenecks and overburden the public machine. Second, it does not address its applicability to public security and law enforcement. Finally, it does not predict how it will respond and deal with trade secrecy arguments.

5.4 Brazilian Legislative Proposals: Convergence and Persistent Gaps

PL 2338/2023, Brazil’s most advanced AI governance proposal, imports key elements of the EU AI Act framework: a list of prohibited and high-risk AI applications, documentation and impact assessment obligations, and individual rights to explanation and contestation (art. 8). Its article 5, inc. III establishes the right to contest automated decisions producing legal effects, without qualification based on the domain of application. This is a significant advance over both the LGPD’s art. 20, which permits trade secrecy as a limit on access rights, and the LGPD’s art. 4, which excludes public security from the law’s scope. However, like the EU AI Act it models, PL 2338/2023 protects trade secrets in both the impact assessment stage (art. 26) and the public AI database (art. 43), reproducing the structural tension the framework was ostensibly designed to resolve.

PL 4612/2019, focused on facial recognition and biometric AI systems, offers a partial departure from this pattern. Its art. 6, para. 1 states explicitly that commercial and industrial secrets cannot serve as justification for violations of rights, ethical standards, or the norms established by the law. This is the only instance across all analyzed proposals in which trade secrecy is formally subordinated to public law transparency obligations rather than protected alongside them. Art. 4 establishes the ANPD as the competent oversight authority for biometric AI, and art. 11 provides for audit rights with access to logs and infrastructure documentation. The proposal is limited in scope and institutional depth, but the normative principle it establishes is foundational to any adequate regulatory response.

PL 846/2026 establishes, for the first time within the current Brazilian law proposals, the possibility of overview and human review of automated decisions (art. 2, inc. IV; art. 3, single para., inc. I and III). Its art. 5 goes beyond and requires that all AI systems hired and implemented by public administration be open source, which converges with the French Digital Republic Act framework for algorithmic transparency. However, the PL leaves many open questions: it does not define who would be responsible for algorithmic audit; what would be the frequency of such audits; where the open source codes would be available and how it would handle potential circumvention attempts from bad actors; how it would respond to trade secrecy complaints from private technological providers; nor who would be responsible for enforcing liabilities and penalties in case the law is disregarded. Even though the proposal is still incipient and requires civil society and academia debate over its terms, it is a first step towards a bigger algorithmic transparency framework.

5.5 What Citizens Actually Want: Accountability Preferences and Institutional Design

Regulatory design for AI governance has frequently treated citizen preferences as a secondary consideration, subordinating them to technical and commercial feasibility assessments. The empirical evidence suggests this is both normatively and instrumentally mistaken: citizens have clear and consistent preferences about accountability mechanisms, and those preferences diverge systematically from the passive disclosure frameworks that dominate existing regulatory instruments.

A discrete choice experiment conducted with 2,080 Chinese citizens by Feng & Chandra (2025) found that all four categories of accountability mechanisms tested, procurement conditions, transparency, auditing, and appeal mechanisms significantly increased citizen preference for AI-based government systems over the baseline. The most striking finding concerns transparency type: ecosystem transparency, defined as information about the development, management, operations, and organizational processes surrounding an AI system, increased preference by 157%, compared to 65% for technical transparency alone. The preference gap was statistically significant (log-odds difference of 0.447, $p < 0.001$). Citizens are not primarily interested in source code or model architecture. They want to understand the institutional and organizational context in which algorithmic decisions are made and who is responsible for overseeing

them. This finding maps almost precisely onto the meso and macro levels of the multi-level transparency framework: ecosystem transparency is organizational and institutional transparency by another name.

Equally significant is the finding on appeal mechanisms: administrative appeals increased preference by 432% compared to 156% for judicial appeals, with a log-odds difference of approximately 0.73 (OR approximately 2.08, $p < 0.001$) (Feng & Chandra, 2025). In a legal system with chronic judicial overload, citizen preference for administrative redress channels reflects a rational assessment of accessibility and timeliness, not a rejection of legal protection.

A parallel body of experimental research conducted in the Netherlands tested two dimensions of institutional transparency on citizen trust in predictive policing: legislation transparency (disclosure of applicable legal frameworks) and monitoring transparency (disclosure of active external oversight) (Grimmelikhuijsen & Meijer, 2022; Wenzelburger et al., 2024). Legislation transparency produced no significant effect on any trust dimension. Monitoring transparency, however, produced significant increases in perceived competence, benevolence, and integrity. This result establishes an important design principle: what generates citizen trust is not the existence of legal rules but the visible, active enforcement of those rules by an independent institutional actor.

Table 2. Comparative analysis of regulatory frameworks: levels addressed and structural gaps

Framework	Micro-level	Meso-level	Macro-level	Trade secrecy override
EU AI Act (2024)	Partial: disclosure to notifying bodies; source code protected under art. 70.	Limited: documentation obligations on providers; organizational context not disclosed publicly.	Yes: institutional oversight architecture; notifying bodies; conformity assessment.	No: art. 70 explicitly protects trade secrets during audits.
French Digital Republic Act (2016)	Indirect: public information doctrine covers all data used by public services.	Partial: organizational integration as contracted service treated as public information.	Limited: no dedicated technical oversight body created.	Yes: overrides IP and trade secrecy for public service providers (art. 8).
India AI Governance Guidelines (2025)	Yes: the “Understandable by Design” sutra requires comprehensible and accessible disclosures.	Yes: mandates grievance redressing mechanisms and public administration integration in different levels, as well as reporting of incidents.	Yes: the National AI Incidents Database and the AI Governance Group (AIGG).	Partial: accountability based on real-world harms; incident reporting is public.
PL 2338/2023 (Brazil)	Yes: individual explanation rights and documentation obligations (art. 8).	Limited: impact assessments required but no organizational monitoring body.	Limited: no dedicated oversight body or gatekeeper established.	No: trade secrecy protected in impact assessments (art. 26) and database (art. 43).
PL 4612/2019 (Brazil)	Partial: documentation and audit access (art. 11).	Yes: ANPD as a competent oversight authority for biometric AI.	Partial: limited to biometric AI; no general law enforcement mandate.	Partial: art. 6 para. 1 formally subordinates secrecy to rights violations.
PL 846/2026 (Brazil)	Yes: individual explainability rights, challenge of automated decisions and audit (art. 3).	Limited: no authority defined both for monitoring and enforcement of penalties.	Limited: law enforcement focused but no oversight body or gatekeeper defined.	Partial: art. 5 mentions the preferability for open source programs but does not mandate it nor defines monitoring and accountability structures.

Description: This table compares international and Brazilian AI regulations and frameworks. It highlights that most frameworks lack a mechanism to override trade secrecy when high-risk systems violate public rights.

Table 2 reveals a consistent pattern presenting how each framework addresses some levels of the transparency architecture while leaving others inadequately covered. No single framework simultaneously achieves micro-level technical disclosure, meso-level organizational transparency, macro-level active institutional oversight, and formal override of trade secrecy in rights-violation contexts. The Smart Sampa program occupies the intersection of all four gaps: a high-impact system where none of the existing or proposed frameworks would generate the accountability the multi-level framework identifies as necessary.

6. Proposals: An Institutional Framework for Algorithmic Accountability

The analysis developed across the preceding sections converges on a single organizing claim: the accountability deficit of programs like Smart Sampa is not primarily a product of legislative absence, but of institutional absence: the absence of an actor with the technical expertise, legal mandate, and operational resources to give existing transparency obligations practical effect in the context of algorithmically mediated public security. As Citron (2008) observed, the law has limited tools for handling technological black boxes but extensive experience handling humans, processes, and organizations. The proposals here work through the latter to reach the former.

The proposals are structured around Kaminsky's (2020) five-question framework: who, what, how, when, and why. They do not require new primary legislation but the extension and amplification of institutional roles and powers already available within the Brazilian legal system.

6.1 Who: The Algorithmic Gatekeeper

The central institutional proposal is the creation of a specialized body within the Judiciary, designated as an algorithmic gatekeeper, with the mandate and technical capacity to serve as an intermediary between the opacity of law enforcement AI systems and the transparency obligations of public law. The rationale for institutional intermediation is both empirical and normative. Passive transparency tools directed at the general public, including algorithm registers, explanation requirements, and access to information channels, consistently fail to generate meaningful accountability in AI governance contexts because citizens lack the technical literacy to assess disclosed information and because disclosed information is routinely incomplete (Cath & Jansen, 2022). The experimental evidence reviewed in Section 5.4 reinforces this conclusion: it is active external monitoring, not legislative disclosure, that generates citizen trust (Grimmelikhuijsen & Meijer, 2022; Wenzelburger et al., 2024).

The gatekeeper should be structured analogously to the notifying bodies established under Chapter 4 of the EU AI Act, with adaptations reflecting the Brazilian legal context. Its core characteristics should include: multidisciplinary staffing combining technical AI expertise with administrative and criminal law competence; a formal prohibition on commercial consultancy activities; strict confidentiality obligations and post-service exclusivity requirements; advanced cybersecurity infrastructure; and a mandate extending to both pre-deployment review and ongoing operational monitoring. The location within the Judiciary is motivated by the courts' established competence in handling sensitive information, adjudicating fundamental rights disputes, and enforcing disclosure obligations against public authorities and private contractors (Gomes et al., 2018). It also enables the gatekeeper's pre-deployment review function to be linked to the judicial authorization requirement proposed in PL 2338/2023 for biometric surveillance in public spaces.

Preventing institutional capture requires specific safeguards against the revolving door phenomenon. As Kaminsky (2020) emphasizes, algorithmic governance bodies are particularly vulnerable to capture when staff members move between regulatory and industry roles involving the same technology systems. The proposed safeguards include a minimum four-year post-service hiring freeze with respect to AI companies in the same sector; mandatory conflict of interest disclosure throughout employment; secure internal reporting channels; and stakeholder participation mechanisms ensuring external oversight of gatekeeper decisions.

6.2 What: Multi-Level Information Access

The information framework proposed here distinguishes between two levels of access subject and two orders of transparency object. The distinction follows from the multi-level framework's insight that different actors require different kinds of information to discharge different accountability functions.

The gatekeeper requires access to first-order information: the system's architecture, training data sources and quality assessment, false positive and negative rates disaggregated by race, gender, and age cohort, confidence thresholds, output documentation, and records of human decision points in both development and deployment. This information is commercially sensitive and should be protected from public disclosure under a confidentiality regime modeled on EU AI Act Article 70, adapted to ensure that protection covers the information, not the fact of non-compliance with applicable standards. The principle from PL 4612/2019, art. 6, para. 1, is foundational: commercial secrets cannot justify violations of rights, ethical standards, or applicable legal norms.

The gatekeeper also requires access to second-order information: organizational governance structures surrounding the system, including the full text of the procurement contract, internal operational policies, operator training and qualification records, audit logs linking algorithmic outputs to police actions, and the monitoring and evaluation procedures actually in place. As Ananny & Crawford (2018, p. 974) argue, transparency of the sociotechnical system as a whole, rather than merely the algorithm, is the condition for accountability that means something.

The general public should have access to high-level information about the purpose, scope, and performance of AI systems deployed in public security contexts, in compliance with LAI's art. 7 active transparency obligations. This

information should be proactively disclosed prior to deployment, presented in plain language, and updated at regular intervals. The minimum public disclosure package should include: system purpose and functional description; geographic and temporal deployment scope; categories of personal data processed; a plain-language summary of the pre-deployment impact assessment; and identification of the institutional body responsible for ongoing oversight.

6.3 How: Extending the LAI and Overriding Trade Secrecy

Two complementary legal extensions of the existing framework are proposed. First, the active transparency obligations of LAI art. 8 should be formally extended, through a binding regulatory interpretation by the Brazilian *Controladoria Geral da União* or other competent authority, to require proactive disclosure of information about AI systems used in the provision of public services. The French Digital Republic Act's doctrinal model is instructive as it requires that information produced or used by public authorities and their contracted partners in the delivery of public services belongs to the domain of public law transparency, regardless of the private origin of the underlying technology. This interpretation does not require new legislation, but instead it requires a recognition that AI systems deployed in the exercise of public functions are, in the relevant sense, public information.

Second, trade secrecy should be formally precluded as a legal defense against gatekeeper access requests. The precedent from PL 4612/2019 and the jurisprudential argument developed by Wexler (2018) support the categorical principle that the public administration's interest in verifying the accuracy, legality, and constitutional compliance of AI systems it deploys against citizens is superior to the commercial interests of private technology providers. This does not require public disclosure of trade secrets; it requires their disclosure to a trusted institutional intermediary operating under a robust confidentiality regime that protects commercial value while enabling public accountability for deployment. The LGPD sanctions framework (art. 52), as detailed in ANPD Resolution CD/ANPD 4/2023, provides a model for the administrative and criminal liability structure that should govern gatekeeper operations.

6.4 When: Pre-Deployment Review and Ongoing Monitoring

The comparative regulatory analysis in Section 5 identifies a consistent limitation of reactive oversight models: by the time harm is documented and an investigation initiated, affected individuals have typically been subjected to algorithmic decisions without remedy for extended periods. Simsek (2021) characterizes this as the fundamental design flaw of the EU AI Act's audit mechanism. The alternative is a mandatory pre-deployment review requirement, analogous to the conformity assessment procedure under the EU AI Act but structured as a condition of deployment rather than a reactive audit trigger.

Pre-deployment review by the gatekeeper should be mandatory for all AI systems intended for use in law enforcement contexts, covering facial recognition, predictive policing, automated suspicion algorithms, and any other system whose output may directly influence decisions affecting individual liberty. Review should assess: compliance with applicable legal requirements; adequacy of documentation, impact assessment, and data governance; proportionality of deployment relative to identified rights impacts; and demographic performance disaggregation. Deployment should be conditional on gatekeeper approval. The financial and administrative costs of this requirement should be borne by the deploying public authority, creating an incentive for selective and justified acquisition decisions.

Ongoing monitoring should be structured to verify that operational behavior conforms to the information provided in the pre-deployment review. Regular audit cycles, with publicly accessible summaries of findings, identified deficiencies, corrective measures required, and compliance status, should be a core gatekeeper function. Following the Indian model, a grievance redressing system could also be included to complement the information collected through ongoing audits. Public reporting of monitoring findings and a public incident repository are mechanisms through which the monitoring transparency effect documented by Grimmelikhuijsen & Meijer (2022) and Feng & Chandra (2025) is realized: citizens trust algorithmic governance when they can observe that an independent body is actively checking it.

6.5 Why: Legitimacy, Due Process, and the Accountability Imperative

The efficiency arguments that accompany AI deployment in public security are not without substance. Algorithms offer genuine advantages in processing speed, predictive consistency, and operational scale that manual surveillance cannot match (Kitchin, 2021). Acknowledging these advantages is not, however, an argument for opacity. It is, if anything, an argument for more rigorous accountability: systems whose outputs significantly influence police actions affecting individual liberty require more scrutiny, not less, precisely because their scale and speed multiply the potential rights impact of any systematic error.

The documented accuracy deficits of facial recognition systems are not marginal. They fall disproportionately on individuals who are already most vulnerable to law enforcement error: people of color, women, younger individuals (Klare et al., 2012; Buolamwini & Gebru, 2018). In Brazil, where police violence against Black communities is a structural and well-documented phenomenon, deploying AI systems with undisclosed and unaudited demographic

accuracy disparities is not a neutral technical choice. As Kaminsky (2020, p. 1539) argues, models are opinions embedded in mathematics: they reflect the assumptions and values of those who designed them and the historical data on which they were trained. The legitimacy of those opinions in the context of public administration depends on their subjection to the same scrutiny that art. 37 of the Brazilian Constitution requires of all exercises of public power.

The citizen preference evidence reviewed in Section 5.4 reinforces the institutional design implications of this argument. The strong preference for administrative appeals over judicial remedies (OR approximately 2.08, Feng & Chandra, 2025) suggests that accessibility and speed are not secondary features of accountability mechanisms but conditions of their effectiveness. A right to challenge that requires filing a lawsuit to exercise is not, in any practical sense, available to individuals who have just been detained or identified as suspects. An administrative appeals mechanism before the gatekeeper, with mandatory response within defined time limits, is the institutionally adequate form for this right to take.

Table 3. Regulatory recommendations for the Smart Sampa program and analogous law enforcement AI systems

Level	Recommendation	Legal basis	Expected effect
Micro	Mandatory disclosure to the gatekeeper of algorithm architecture, training data sources, false positive and negative rates disaggregated by race, gender, and age, confidence thresholds, and output decision logic.	PL 4612/2019 arts. 6 and 11; PL 2338/2023 art. 8; PL 846/2026, art. 2; gatekeeper mandate.	Enables technical bias assessment; creates evidential basis for individual contestation; makes demographic accuracy disparities visible to institutional oversight.
Micro	Individual notification when subject to facial recognition; plain-language explanation of algorithmic output and subsequent human decision, with reference to contestation mechanism.	LAI art. 7; PL 2338/2023 arts. 5 and 8; CNJ Directive 332/2020.	Restores individual agency; generates audit trail linking algorithmic outputs to police actions; addresses Citron's (2008) due process concern.
Meso	Mandatory full publication of procurement contracts including technical specifications and performance obligations imposed on the private provider.	Law 14.133/2021; LAI art. 7 inc. V; CF art. 37.	Eliminates 'middle-man opacity' documented by Arroyo & Gaspar (2021); enables civil society and parliamentary scrutiny of contractual accountability mechanisms.
Meso	Creation and proactive publication of an algorithm register including: purpose, operational scope, data sources, human oversight protocols, evaluation methodology, and update history.	LAI art. 8; CF art. 37; PL 2338/2023 art. 43 extended to law enforcement; PL 846/2026 art. 5.	Enables ongoing organizational scrutiny; creates internal disciplining effect on algorithmic governance practices (Floridi, 2020); generates audit reference point for gatekeeper.
Macro	Mandatory pre-deployment review by the algorithmic gatekeeper, with deployment conditional on gatekeeper approval after impact and rights assessment.	Modeled on EU AI Act art. 23; PL 4612/2019 art. 4; operationalized via CNJ resolution or Judiciary regulation.	Converts oversight from reactive to proactive; prevents rights deprivations that cannot be retroactively remedied; aligns with precautionary approach (Crotoft & Ard, 2020).
Macro	Trade secrecy formally precluded as defense against gatekeeper access; confidentiality regime protecting commercial information from public disclosure while enabling institutional oversight.	PL 4612/2019 art. 6 para. 1; Wexler (2018) principle applied to Brazilian public law; EU AI Act art. 70 adapted; PL 846/2026 art. 5.	Removes the primary legal obstacle to AI oversight in Brazilian law enforcement; operationalizes the constitutional principle that commercial interests yield to fundamental rights.
Macro	Accessible administrative appeal mechanism before the gatekeeper for individuals whose rights were affected by an AI-mediated decision, with mandatory response within defined time limits.	PL 2338/2023 art. 5 inc. III; CF art. 5 inc. XXXIV; Feng & Chandra (2025); PL 846/2026 art. 3, single para., III; India AI Guidelines (Grievance Redressal Mechanisms).	Provides the accessible, specialized, and timely redress channel strongly preferred over judicial appeals by citizens (OR approximately 2.08); generates operational feedback for system correction.

Macro	Mandatory annual public reporting by the gatekeeper of monitoring findings, identified deficiencies, corrective measures required, and compliance status.	CF art. 37; LAI art. 8; Grimmelikhuijsen & Meijer (2022); Wenzelburger et al. (2024); India AI Guidelines (National AI Incidents Database).	Generates the active monitoring signal empirically identified as the primary driver of citizen trust; makes institutional oversight visible to the public and civil society.
-------	---	---	--

Description: This table presents seven institutional proposals. Each measure is linked to existing Brazilian legal principles and international examples, and describes the expected improvement in accountability and public trust.

The seven recommendations in Table 3 constitute an integrated framework, not a menu of options. Their functional interdependence is the key design feature: micro-level interventions are meaningless without an institutional actor empowered to enforce them; macro-level oversight is empty without access to micro and meso-level information; and the accountability signal that generates citizen trust requires the entire architecture to function simultaneously. As Giest & Grimmelikhuijsen (2020) demonstrate, organizations are adept at producing formal compliance with transparency requirements while maintaining substantive opacity. The proposals here attempt to address this structural tendency by making the gatekeeper's active oversight function publicly visible, creating the external accountability signal that Feng & Chandra (2025) and Grimmelikhuijsen & Meijer (2022) identify as the mechanism through which institutional legitimacy is built.

7. Conclusion

This article began with a question about transparency and ended with a proposal about institutions. That trajectory reflects the central analytical finding of the research: the transparency deficit of Smart Sampa and comparable Brazilian law enforcement AI programs is not primarily a problem of missing legal rights but of missing institutional actors with the capacity and mandate to give existing rights practical effect. Brazilian law already establishes the formal architecture for public administration transparency. What it lacks, in the specific context of algorithmically mediated law enforcement, is the specialized institutional presence that can pierce the dual opacity produced by public security secrecy and trade secrecy arguments that public authorities and their technology providers have deployed with consistent success against citizen access requests.

The multi-level transparency framework applied throughout this article demonstrates that this institutional absence has consequences at every level of algorithmic governance. At the micro level, affected individuals have no access to information about the systems that have produced decisions about them, and no practical mechanism for challenging those decisions. At the meso level, the organizational embedding of AI systems in police operations is invisible to civil society, academia, and legislative oversight. At the macro level, the absence of an active external monitor means that the formal legal obligations of public administration transparency exist without the enforcement pressure that would give them operational effect. The experimental evidence reviewed in Section 5.4 suggests that citizens understand this structure intuitively: they trust algorithmic governance not when they are told about the laws governing it but when they can observe an independent institution actively checking compliance with those laws.

The proposals advanced in Section 6 are designed to be operationally realistic. They do not require the enactment of new primary legislation but the amplification of institutional roles available within the existing Brazilian legal system. The proposal for a judicial gatekeeper builds on the existing competence of the Judiciary in handling sensitive information and adjudicating fundamental rights disputes. The proposal to extend LAI's active transparency obligations to AI systems builds on interpretive precedents already available in administrative law. The proposal to override trade secrecy against gatekeeper access requests builds on the principle already embedded in PL 4612/2019 and the jurisprudential arguments developed in comparative context by Wexler (2018). The proposal for administrative appeals builds on the constitutional right of petition and the administrative redress infrastructure established by the LGPD's ANPD model.

It is important to conclude with an acknowledgment of what transparency cannot do. Zarsky (2013, p. 1568-1569) correctly observes that transparency regimes, however robust, cannot eliminate the structural biases embedded in the historical data on which AI systems are trained, nor the institutional biases that shape how algorithmic outputs are translated into police decisions. A facial recognition system that produces more false positives for Black Brazilians reflects, at the level of its training data, the historical patterns of a criminal justice system shaped by structural racism. Making that disparity visible through the transparency framework proposed here is a necessary condition for addressing it, not a sufficient one. The root causes require political responses and public deliberation of a kind that transparency facilitates but does not substitute for.

Just as code can shape law, as Lessig (1999) demonstrated, law can and must shape code, particularly when the code in

question is making decisions about the liberty, dignity, and fundamental rights of citizens in a constitutional democracy. The Smart Sampa program is an example of code running ahead of the law. The proposals in this article are designed to close that gap.

Author Information

¹ Mayara Cristhina Rodriquero dos Santos is a Brazilian licensed lawyer with professional experience focused on Digital Law and Regulatory Compliance. She holds a Master's and a Bachelor's degree in Law from FGV São Paulo Law School (Escola de Direito de São Paulo da Fundação Getulio Vargas). She was granted a recognition award from FGV São Paulo Law School as one of the best Master's dissertations of 2024, as well as was a recipient of the FGV Presidency Merit-Socioeconomic Grant during her bachelors. She completed specialized training in Digital Law at the Geneva Digital Law School at the University of Geneva.

² Alexandre Pacheco da Silva is a Professor of Digital Law and Intellectual Property at FGV São Paulo Law School (Escola de Direito de São Paulo da Fundação Getulio Vargas), where he teaches in both the undergraduate and graduate programs. He holds a Ph.D. in Science and Technology Policy from the Institute of Geosciences at the University of Campinas (UNICAMP), as well as a Master's degree and a Bachelor's degree in Law from FGV São Paulo Law School. He also serves as Coordinator of the Center for Education and Research in Innovation (CEPI) at FGV São Paulo Law School.

³ Marina Feferbaum is a Professor at FGV São Paulo Law School (Escola de Direito de São Paulo da Fundação Getulio Vargas) and Coordinator of the Center for Education and Research in Innovation (CEPI). Her work focuses on law and technology, AI governance, digital rights, and innovation in legal education. She leads interdisciplinary research and international collaborations on public interest technology, digital governance, and the impacts of emerging technologies on legal institutions and society.

Acknowledgments

The authors wish to thank Professor Raquel Fortes Gato for her thoughtful engagement with this research, including her insightful comments and the recommendations she generously offered throughout the process. The authors are likewise grateful to Professor Guilherme Forma Klafke for sharing relevant documents on the Smart Sampa program and for his valuable observations on the research.

Author contributions

This article originates from research conducted by Mayara Cristhina Rodriquero dos Santos as part of her Master's degree program at FGV São Paulo Law School. Data collection and bibliographic review were carried out by Mayara Cristhina Rodriquero dos Santos, while the adaptation of the dissertation and the specific analysis of the Smart Sampa program were developed by Professors Alexandre Pacheco da Silva and Marina Feferbaum. Manuscript drafting was the responsibility of Professors Alexandre Pacheco da Silva and Marina Feferbaum, and the manuscript was revised by all three authors through a series of joint working sessions. All authors have read and approved the final version of the text and consider that they have contributed equally to it.

Funding

Not applicable.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Not applicable.

Ethics approval

The Publication Ethics Committee of the Redfame Publishing.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

References

- Alencar, I. (2023). Com mais de mil prisões na BA, sistema de reconhecimento facial é criticado por 'racismo algorítmico'. *Portal G1*, Bahia.
- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973-989. <https://doi.org/10.1177/1461444816676645>
- Arroyo, V., & Gaspar, P. (2021). *Tecnologia de vigilância na América Latina*. Access Now.
- BRASIL. (1988). *Constituição Federal de 1988*. Brasília: Senado Federal.
- BRASIL. (1996). *Lei nº 9.279, de 14 de maio de 1996*. Regula direitos e obrigações relativos à propriedade industrial.
- BRASIL. (2011). *Lei nº 12.527, de 18 de novembro de 2011*. Lei de Acesso à Informação (LAI).
- BRASIL. (2014). *Lei nº 12.965, de 23 de abril de 2014*. Marco Civil da Internet.
- BRASIL. (2018). *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD).
- BRASIL. (2019). *Projeto de Lei nº 4612/2019*. Dispõe sobre o desenvolvimento, aplicação e uso de tecnologias de reconhecimento facial e emocional.
- BRASIL. (2020). *Resolução CNJ nº 332/2020*. Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário.
- BRASIL. (2021). *Lei nº 14.133, de 1º de abril de 2021*. Lei de Licitações e Contratos Administrativos.
- BRASIL. (2023). *Projeto de Lei nº 2338/2023*. Dispõe sobre o uso da Inteligência Artificial.
- BRASIL. (2026). *Projeto de Lei nº 846/2026*. Institui a Lei de Governança de Inteligência Artificial no Setor Público.
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on Fairness, Accountability and Transparency*. PMLR, pp. 77-91.
- Cath, C., & Jansen, F. (2022). AI-registers: Realising the EU AIA requirements for algorithmic transparency through open government data. *Federal Communications Law Journal*.
- Citron, D. (2008). Technological due process. *Washington University Law Review*, 85, 1249.
- Crootof, R., & Ard, B. J. (2020). Structuring techlaw. *Harvard Journal of Law & Technology*, 34, 347. <https://doi.org/10.2139/ssrn.3664124>
- Degeling, M., & Berendt, B. (2018). What is wrong about Robocops as consultants? A technology-centric critique of predictive policing. *AI & Society*, 33, 347-356. <https://doi.org/10.1007/s00146-017-0730-7>
- Desai, D. R., & Kroll, J. A. (2017). Trust but verify: A guide to algorithms and the law. *Harvard Journal of Law & Technology*, 31, 1.
- EUROPEAN COMMISSION. (2024). Regulation (EU) 2024/1689 (Artificial Intelligence Act). *Official Journal of the European Union*.
- Feng, N., & Chandra, Y. (2025). Accountability in government use of AI: Citizen concerns and preferences. *Public Administration*, 0, 1-22. <https://doi.org/10.1111/padm.70030>
- Floridi, L. (2020). The fight for digital sovereignty: What it is and why it matters, especially for the EU. *Philosophy & Technology*, 33(3), 369-378. <https://doi.org/10.1007/s13347-020-00423-6>
- Giest, S., & Grimmelikhuijsen, S. (2020). Practitioner perceptions of algorithmic tools in administrative decision-making. *Policy & Internet*, 12(4), 135-164.
- Gomes, L. et al. (2018). Opacidade algorítmica: desafios para o acesso à informação. *Liinc em Revista*, 14(1).
- GOVERNMENT OF INDIA. (2025). *India AI Governance Guidelines: Enabling Safe and Trusted AI Innovation*.

- Grimmelikhuijsen, S., & Meijer, A. J. (2014). Effects of transparency on the perceived trustworthiness of a government organization. *Journal of Public Administration Research and Theory*, 24(1), 137-157. <https://doi.org/10.1093/jopart/mus048>
- Grimmelikhuijsen, S., & Meijer, A. J. (2022). Legitimacy of algorithmic governance: Algorithms in public sector decision making. *Public Administration*.
- Kaminsky, M. E. (2019). Binary governance: Lessons from the GDPR's approach to algorithmic accountability. *Southern California Law Review*, 92, 1529. <https://doi.org/10.2139/ssrn.3351404>
- Kaminsky, M. E. (2020). Algorithmic accountability. *Notre Dame Law Review*, 98(3), 1-68.
- Kitchin, R. (2021). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences* (2nd ed.). London: Sage.
- Klare, B. F. et al. (2012). Face recognition performance: Role of demographic information. *IEEE Transactions on Information Forensics and Security*, 7(6), 1789-1801. <https://doi.org/10.1109/TIFS.2012.2214212>
- Lapin, A. (2022). Transparency as trust in algorithmic systems of public administration. In *Proceedings of the 22nd Annual International Conference on Digital Government Research*.
- Leese, M. (2014). The new profiling: Algorithms, black boxes, and the failure of anti-discriminatory safeguards in the European Union. *Security Dialogue*, 45(5). <https://doi.org/10.1177/0967010614544204>
- Lessig, L. (1999). The law of the horse: What cyberlaw might teach. *Harvard Law Review*, 113(2), 501-549. <https://doi.org/10.2307/1342331>
- Maranhão, J. S. A., Junquillo, T. A., & Tasso, I. (2023). Transparencia de sistemas de inteligência artificial: uma taxonomia baseada em direitos. *Revista de Direito Administrativo*, 282(2), 151-182.
- Meijer, A. J., & Grimmelikhuijsen, S. (2020). Responsible and accountable algorithmization: How to generate citizen trust in governmental usage of algorithms. *ACM FAccT*. <https://doi.org/10.4324/9780429261404-5>
- Moses, L. B., & Koker, L. De. (2017). Open secret: Balancing operational secrecy and transparency in the collection of location data by law enforcement. *Monash University Law Review*, 43(1).
- Nieuwenhuizen, E. (2025). *Algorithmic transparency in government: A multi-level perspective on transparency of and trust in algorithm use by governments*. Utrecht: Utrecht University.
- Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge: Harvard University Press. <https://doi.org/10.4159/harvard.9780674736061>
- Porumbescu, G. Et al. (2022). Connecting organizational transparency to citizen coproduction: A cross-country comparative study. *Governance*, 35(4), 1243-1262.
- PREFEITURA DE SÃO PAULO. (2023). *Smart Sampa: Programa de videomonitoramento inteligente*. São Paulo.
- PREFEITURA DE SÃO PAULO. (2025). *Relatório de Transparência Smart Sampa*. São Paulo.
- Reisman, D. et al. (2018). *Algorithmic impact assessments: A practical framework for public agency accountability*. AI Now Institute.
- RÉPUBLIQUE NUMÉRIQUE. (2016). *Loi no 2016-1321 du 7 octobre 2016 pour une République numérique*. Journal Officiel de la République Française.
- Simsek, D. (2021). Algorithmic transparency and accountability: A framework for the public sector. *European Public Law*, 27(1), 1-35.
- TRANSPARÊNCIA BRASIL et al. (2020). *Inteligência artificial e segurança pública: mapeamento de usos e riscos no Brasil*.
- Wenzelburger, G. et al. (2024). Trustworthy artificial intelligence and the demand for institutional transparency: Evidence from an experimental vignette study. *Public Administration*.
- Wexler, R. (2018). Life, liberty, and trade secrets: Intellectual property in the criminal justice system. *Stanford Law Review*, 70(5), 1343. <https://doi.org/10.2139/ssrn.2920883>
- Wischmeyer, T., & Rademacher, T. (Eds.). (2020). *Regulating Artificial Intelligence*. Berlin: Springer. <https://doi.org/10.1007/978-3-030-32361-5>
- Zarsky, T. (2013). Transparent predictions. *University of Illinois Law Review*, 4, 1503-1569.