ALGORITHMIC GOVERNANCE AND THE PUBLIC INTEREST: ETHICAL FOUNDATIONS FOR DECISION-MAKING IN THE PUBLIC SECTOR

GOVERNANÇA ALGORÍTMICA E INTERESSE PÚBLICO: BASES ÉTICAS PARA DECISÕES NO SETOR PÚBLICO

GOBERNANZA ALGORÍTMICA E INTERÉS PÚBLICO: BASES ÉTICAS PARA LA TOMA DE DECISIONES EN EL SECTOR PÚBLICO

https://doi.org/10.56238/sevened2026.008-006

Ailton Ferreira Cavalcante[1]

ABSTRACT

Ongoing technological transformations introduce new layers of decision-making power within the State, particularly through data-driven and algorithmic systems. While these tools can enhance analytical capacity and administrative efficiency, they may also reinforce inequalities, embed invisible biases, and weaken public transparency. This article examines algorithmic governance from an ethical-normative perspective grounded in the principle of the public interest. It engages classical moral philosophy, republican traditions, and contemporary debates on democratic governance, integrating them with the notion of "institutional conscience" and with the GIS Cycle (Governance, Innovation, and Sustainability). Methodologically, this is a theoretical-conceptual study based on a narrative review and critical analysis of legal frameworks and international references. The argument advanced is that the legitimacy of algorithm-mediated decisions depends on four concurrent criteria: universalizability, dignity, transparency, and accountability. From these principles, the article proposes guidelines for public policies involving data and artificial intelligence, including auditability, explainability, bias mitigation, protection of vulnerable groups, and reversibility of automated outcomes. It concludes that, in the digital age, the primacy of the public interest requires institutions to cultivate stable ethical self-control mechanisms capable of aligning technological innovation, social justice, and sustainability. More than a technical matter, algorithmic governance emerges as a profoundly moral and democratic challenge.

Keywords: Governance. Algorithms. Public Ethics. Public Interest. Public Policy. Sustainability.

RESUMO

As transformações tecnológicas em curso introduzem novas camadas de poder decisório no Estado, especialmente por meio de sistemas baseados em dados e algoritmos. Embora ampliem a capacidade analítica e a eficiência administrativa, tais instrumentos também podem reforçar assimetrias, produzir vieses invisíveis e reduzir a transparência pública. Este artigo discute a governança algorítmica a partir de um enfoque ético-normativo orientado

[1] Doctoral student in Public Administration. Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP).
E-mail: prof.ailton@hotmail.com Orcid: https://orcid.org/0009-0009-2386-1245
Lattes: http://lattes.cnpq.br/0389434047626666

pelo princípio do interesse público. Retomam-se contribuições da filosofia moral, da tradição republicana e dos debates contemporâneos sobre governança democrática, articulando-as com a noção de “consciência institucional” e com o Ciclo GIS (Governança, Inovação e Sustentabilidade). Neste estudo optamos por uma abordagem teórico-conceitual, baseado em revisão narrativa e análise crítica de marcos normativos e referenciais internacionais. Argumenta-se que a legitimidade das decisões mediadas por algoritmos depende da observância simultânea de quatro critérios: universalizabilidade, dignidade, transparência e responsabilização. A partir desses princípios, propõem-se diretrizes para políticas públicas que envolvem dados e inteligência artificial, incluindo auditabilidade, explicabilidade, mitigação de vieses, proteção de grupos vulneráveis e reversibilidade das decisões automatizadas. Conclui-se que, na era digital, a centralidade do interesse público exige que as instituições desenvolvam mecanismos estáveis de autocontrole ético, capazes de alinhar inovação tecnológica, justiça social e sustentabilidade. Mais do que uma questão técnica, a governança algorítmica revela-se um problema de natureza moral e democrática.

Palavras-chave: Governança. Algoritmos. Ética Pública. Interesse Público. Políticas Públicas. Sustentabilidade.

RESUMEN

Las transformaciones tecnológicas actuales introducen nuevas capas de poder decisorio en el Estado, especialmente mediante sistemas basados en datos y algoritmos. Si bien pueden aumentar la capacidad analítica y la eficiencia administrativa, también corren el riesgo de reproducir desigualdades, generar sesgos invisibles y disminuir la transparencia pública. Este artículo analiza la gobernanza algorítmica desde un enfoque ético-normativo sustentado en el principio del interés público. Se articulan aportes de la filosofía moral, de la tradición republicana y de los debates contemporáneos sobre gobernanza democrática, integrándolos con la noción de “conciencia institucional” y con el Ciclo GIS (Gobernanza, Innovación y Sostenibilidad). Metodológicamente, se trata de un estudio teórico-conceptual, basado en revisión narrativa y análisis crítico de marcos normativos y referencias internacionales. Se sostiene que la legitimidad de las decisiones mediadas por algoritmos depende de cuatro criterios concurrentes: universalizabilidad, dignidad, transparencia y rendición de cuentas. A partir de ellos, se proponen directrices para políticas públicas que utilizan datos e inteligencia artificial, incluyendo auditabilidad, explicabilidad, mitigación de sesgos, protección de grupos vulnerables y reversibilidad de decisiones automatizadas. Se concluye que, en la era digital, la centralidad del interés público exige que las instituciones desarrollen mecanismos estables de autocontrol ético capaces de alinear innovación tecnológica, justicia social y sostenibilidad.

Palabras clave: Gobernanza. Algoritmos. Ética Pública. Interés Público. Políticas Públicas. Sostenibilidad.

1 INTRODUCTION

The expansion of data-based technologies, machine learning, and automated decision support systems has profoundly changed the architecture of public power. What was once focused on human deliberation and traditional bureaucratic mediation has increasingly been filtered through statistical models, predictive systems, and digital platforms capable of automatically classifying, prioritizing, and allocating resources. This transformation introduces clear gains in efficiency and analytical capacity; However, it reposes classic questions of political philosophy and administrative law in new forms: who decides? Based on what criteria? And with what impacts on justice and democracy? (O'NEIL, 2016; EU, 2021; OECD, 2019).

Recent literature on algorithmic governance indicates that such systems are not neutral: they carry assumptions, modeling choices, incomplete databases, and statistical biases that can amplify historical inequalities and produce asymmetric distributional effects, often invisible to ordinary citizens (PASQUALE, 2015; NOBLE, 2018). At the same time, the growing technical opacity, derived from the complexity of artificial intelligence models, stresses fundamental values of the democratic rule of law, such as transparency, social controllability, and motivation of administrative decisions (FLOOD; MENÉNDEZ, 2022; CASS, 2020).

It is clear that the discussion cannot be reduced to an exclusively technical problem. Rather, it is an ethical-normative issue, which involves the status of the public interest, the legitimacy of decisions mediated by algorithms, and the need for institutional mechanisms of self-control capable of preventing abuses, reducing discrimination, and aligning technological innovation with fundamental rights (RAWLS, 2008; HABERMAS, 1997; UNESCO, 2021). The growing centrality of data highlights the tension between efficiency and justice, revealing that the mere automation of processes does not guarantee greater rationality, and can, on the contrary, institutionalize unfair practices under the appearance of neutrality.

From a philosophical point of view, there are instruments to illuminate these dilemmas. The Kantian notion of inner moral law, for example, suggests that the legitimacy of actions depends on their capacity for universalization and unconditional respect for human dignity (KANT, 2003). Transposed to the level of institutions, this perspective requires that public policies and algorithmic systems be evaluated according to clear normative criteria, which include public justification, non-discrimination, and protection of the vulnerable. Similarly, contemporary debates on democratic governance emphasize the importance of *accountability*, participation, and transparency as conditions for the legitimate exercise of power (PETERS; PIERRE, 2016; RHODES, 2017).

At the same time, the international literature has been consolidating guiding frameworks for the responsible use of artificial intelligence in the public sector, highlighting principles such as explainability, auditability, proportionality, risk assessment, and reversibility of automated decisions (OECD, 2019; EU, 2021; WORLD BANK, 2020). Although relevant, such guidelines often remain disconnected from a deeper ethical reflection, resulting in programmatic documents with little practical operationalization.

In this scenario, this article proposes an integrated ethical-normative framework for algorithmic governance, based on the principle of public interest and articulated with the GIS Cycle[2], Governance, Innovation and Sustainability. It is argued that the legitimacy of decisions mediated by algorithms depends on the simultaneous observance of four structuring criteria, namely: *universalizability*, *dignity*, *transparency,* and *accountability*. Based on these foundations, guidelines are developed for the design, implementation, and monitoring of data-based public policies, with special attention to the protection of vulnerable groups and intergenerational impacts.

In terms of methodology, we opted for a theoretical-conceptual study, supported by a narrative review of the specialized literature and a critical analysis of national and international normative frameworks. The objective is to offer an interpretative matrix that helps public managers, policymakers, and researchers to evaluate and guide the use of algorithmic technologies in the State, contributing to a governance model capable of reconciling innovation, social justice, and democratic sustainability.

In the end, it is argued that the consolidation of an institutional awareness, materialized in rules, processes and permanent ethical safeguards, is an indispensable condition for the digital transformation of the public sector not only to produce faster systems, but, above all, fairer, more transparent institutions aligned with the public interest.

2 THEORETICAL FRAMEWORK

Every society is born from the tense encounter between freedom and limit. As human beings made mistakes, learned and disputed meanings, they understood that living together requires more than good intentions: it requires measure, word, pact. It was in this process

[2] The GIS Cycle (Governance, Innovation and Sustainability) is an analytical model proposed by Ailton Ferreira Cavalcante, within the scope of his doctorate in Public Administration (IDP), which understands state action as a continuous process of formulation, execution, monitoring and feedback of public policies, articulating institutional governance, responsible innovation and sustainability as criteria for creating public value. Although developed in a systematic way in previous studies, the GIS Cycle can be understood, in synthetic terms, as a spiral process of planning, execution, monitoring and institutional learning, in which governance, innovation and sustainability operate as inseparable dimensions of the same movement. It is not just a managerial arrangement, but a normative architecture that guides public decision-making based on criteria of democratic legitimacy, intergenerational responsibility, and continuous creation of public value.

that the garden of innocence gave way to the city, a space where power needs to be contained by rules and where each decision starts to carry, simultaneously, promise and risk.

It is in this horizon that the contemporary debate on algorithms and public decisions is inserted. Digital technologies, far from being just instruments, have become part of the very architecture that organizes collective life, designing flows, defining priorities, and distributing opportunities. They can favor cooperation and justice, but from another angle, they can also erect invisible walls. The central question is not strictly technical: it is to know what kind of world we are helping to shape by delegating increasing portions of our decision-making power to machines.

Understanding this challenge requires returning to traditions that have reflected on truth and appearance, order and freedom, creation and responsibility, and from there reinterpreting the role of public interest and democratic governance in the digital age. The theoretical itinerary that follows follows this path: it starts from human moral ambivalence, dialogues with classical philosophy, revisits contemporary constitutionalism and integrates current debates on governance and algorithms. In the end, these contributions converge in the GIS Cycle[3], which will serve as an analytical matrix for subsequent sections, indicating that the digital transformation of the State involves not only new tools, but a new way of interrogating power and responding for it.

2.1 MORAL CONSCIENCE AND HUMAN AMBIVALENCE

Philosophical literature recognizes that the human condition is crossed by a constitutive ambivalence: the capacity for cooperation, compassion and the creation of collective goods coexists with impulses of domination, appropriation and moral indifference. This tension, already intuited in ancient ethical traditions, finds a paradigmatic formulation in Hobbes[4], for whom the absence of shared norms and legitimate authority places individuals

[3] The GIS (Governance, Innovation and Sustainability) Cycle proposes that decisions about the use of algorithms in the public sector be treated as processes of justification and control, and not just as technical decisions. In the Governance phase, the public problem is precisely defined, ethical and legal risks are assessed, alternatives are registered, and it is verified whether the solution can be publicly justified (universalizability and dignity). In the Innovation phase, the development of the system must be interdisciplinary, documented and tested, with prior algorithmic impact assessment, transparent metrics, meaningful human oversight and challenge channels. Finally, the Sustainability phase requires continuous monitoring of errors and distributive effects, periodic reviews, audit trails, real possibility of correction or discontinuation, and public record of lessons learned. The Cycle operates in an iterative manner: problems detected in later stages return for prior adjustments, ensuring that the technology remains subordinate to the public interest, reviewable and responsible.

[4] In Hobbes, the image of *homo homini lupus* does not describe an essential evil of the human being, but the condition of vulnerability and reciprocal distrust that emerges when common norms and legitimate authority are lacking. In the so-called "state of nature", each individual, seeking self-preservation, becomes a potential threat to others. The social contract and the institution of a common political power thus emerge as rational mechanisms to contain the predatory dimension of human relations and stabilize expectations, converting fear into public security (HOBBES, 2003).

in a state of potential war, synthesized in the expression homo homini lupus (HOBBES, 2003). The recourse to the social contract and political authority would therefore not be explained only by organizational efficiency, but by the need to contain the predatory dimension of human relations and produce reciprocal security.

Modernity introduces a decisive shift: the problem of evil and violence cannot be solved exclusively by external means. Kant[5] shows that morality does not derive primarily from the fear of punishment or the calculation of advantages, but from the discovery of a universalizable practical law, inscribed in reason itself (KANT, 2003). Right action is that which can be conceived as a rule valid for all and which respects humanity, in itself and in others, as an end. In this way, conscience comes to be understood as an inner court, prior to and superior to coercive control, a space in which the subject is confronted with moral obligation even in the absence of external surveillance.

This movement has profound implications for the State and for public policies. If individuals do not become righteous just because they are watched, institutions do not become legitimate just because they are legal. Legitimacy, as contemporary theories of justice argue, depends on the ability to subject collective decisions to standards of public justification, in which any person, considered as free and equal, could recognize acceptable reasons (RAWLS, 2008; HABERMAS, 1997). It is a shift from the exclusive domain of authority to the domain of public reason.

However, human ambivalence reappears within the institutions themselves. Even structures created to protect rights can become instruments of capture, excessive bureaucratization or symbolic exclusion. Hence the relevance of thinking of the State as the bearer of a kind of institutional conscience, materialized in rules, processes and organizational cultures that induce ethical behavior and prevent the naturalization of abuse.

In the digital age, this debate gains additional density. Algorithmic systems, by automating decision criteria, seem to shift the moral problem from the individual to the code. However, far from eliminating human ambivalence, they re-inscribe it in new formats: modeling choices, data curation, definition of weights and thresholds, design of interfaces, and usage policies (PASQUALE, 2015; O'NEIL, 2016). The "court of conscience" does not disappear, it moves to the spaces where technical parameters are defined that will later affect concrete lives.

[5] In Kant, morality does not derive from desirable consequences, but from the ability of the agent to submit his actions to a law that can be universally valid. The so-called categorical imperative demands that each person always be treated as an end in himself, never only as a means, the foundation of the principle of human dignity. Moral autonomy, in this sense, implies responsibility: to act is to legislate for oneself and for all, which makes the public justification of norms a condition of legitimacy (KANT, 2003).

Thus, understanding algorithmic governance requires recognizing that there is no purely technological shortcut to justice. Technical rationality needs to be anchored in ethical rationality, otherwise automation will only accelerate and reinforce existing inequalities. In these terms, human ambivalence is not an obstacle to be eliminated by technology, but an anthropological truth that imposes the need for reflective institutions, capable of controlling the power they exercise, even when mediated by algorithms.

2.2 TRUTH, APPEARANCE, AND THE MYTH OF THE CAVE

The Myth of the Cave, presented by Plato in Book VII of the Republic, remains one of the most powerful images for understanding the relationship between knowledge, power, and emancipation. In the narrative, prisoners chained since childhood see only shadows projected on the cave wall and, because they have never experienced any other form of perception, they take them as reality itself (Plato, 2000). The process of liberation, marked by pain, vertigo and resistance, symbolizes the passage from the immediate sensible world to intelligibility and, ultimately, to the Good.

The decisive element of the allegory lies in the fact that the shadows are not the result of chance, but result from an organization of the scenery: there is a fire, objects manipulated by others and a wall that functions as a projection surface. In contemporary terms, the myth points to the existence of regimes of visibility, structures that determine what can be seen, understood, and considered true in a society. There, ignorance is not the absence of data, but imprisonment in a field of appearances produced and mediated.

This reading has become particularly fruitful in the digital context. Recent literature shows how platforms, recommendation systems, and predictive models start to filter and prioritize information, creating highly personalized but non-transparent informational environments (ZUBOFF, 2019; KITCHIN, 2017). *Government dashboards* , risk maps, vulnerability indexes, and automated classification tools produce representations of the world that, due to their technical appearance, tend to acquire a status of unquestionable objectivity. However, such representations depend on methodological decisions, cuts, weights, proxies, which are rarely discussed publicly (PASQUALE, 2015).

The ethical-political risk lies not only in the possibility of error, but in the naturalization of a technically sophisticated shadow regime. By confusing representation with reality, space is opened for the replacement of public debate by opaque technicalities. O'Neil's (2016) critique of the so-called "weapons of mathematical destruction" shows that algorithmic models, when applied to security, credit, education, or social assistance policies, can consolidate trajectories of exclusion precisely because they feed back on historical data

marked by inequalities, which suggests that getting out of the cave also requires critically reviewing the data that feeds it.

On the other hand, the myth also points to an ethical dimension of responsibility: the one who comes out of the cave and contemplates the sunlight, a metaphor for the Good, faces the difficult task of returning to share knowledge and contribute to the liberation of others.

However, Plato reminds us that the passage from darkness to light is not immediate: the eyes of the apprentice, not accustomed to brightness, can be dazzled before seeing clearly. Clarification involves a gradual process, which requires time, mediation and pedagogical care.

In this key, politics is not a mere administration of resources, but a permanent effort to clarify, aware that excessively abrupt transparency can produce misunderstanding and resistance.

In democratic governance, this translates into practices of substantive and gradual transparency: making understandable not only the results, but also the processes, criteria, and uncertainties that structure public decisions (FLORIDI et al., 2018).

Finally, the Myth of the Cave illuminates the contemporary debate on algorithmic governance by revealing that the central issue is not only to produce more information, but to ensure conditions of intelligibility and public contestation. Without this, technically efficient systems can trap societies in new, more refined, but equally coercive forms of obscurity.

Thus, the construction of fair public policies requires institutions capable of deliberately promoting movements to "get out of the cave": critical education about data, open models, citizen participation, and auditing mechanisms that restore to the public space the debate about what counts as administrative truth. Such movements, however, demand accompaniment and formative guidance, as the transition between darkness and light is rarely immediate; Without adequate support, transparency itself can overshadow rather than clarify.

2.3 NUMBER, HARMONY, AND THE PYTHAGOREAN SCHOOL

The Pythagorean School[6] occupies a unique place in the history of thought for articulating, in an unprecedented way, mathematics, metaphysics and ethical formation. More

[6] The Pythagorean tradition conceived of number as a constitutive principle of reality. More than a tool of calculation, arithmetic expressed a cosmic order that should also guide moral and political life. Harmony, proportion and measure were not only mathematical categories, but criteria of justice and social balance. This view anticipated the idea that forms of quantification can shape the world, and therefore need to be thought of ethically.

than a circle of research in numbers, it was a philosophical community in which the study of the world order was intrinsically linked to the moral improvement of its members (BURKERT, 1972). For the Pythagoreans, number was not a mere tool of calculation; it constituted the very intelligible structure of reality, manifesting itself in musical harmony, in the rhythms of nature, and in the proportions of the human body.

The famous discovery of the numerical relations that underlie musical intervals became a symbol of this vision: that which seems qualitative, beautiful, harmonious, pleasant, reveals itself, at a deeper level, quantitatively structured. This *insight* had lasting consequences for Western science, fueling the conviction that the cosmos can be described by mathematical laws and that to know such laws is, in a sense, to participate in the rationality of the world itself (HADOT, 2002).

However, the Pythagorean heritage includes something often overlooked: mathematics was not conceived as an autonomous activity, detached from life. Access to the order of the cosmos required discipline, inner silence, temperance, and principled community life. In other words, there was no science without ethics. The idea of "harmony" referred both to the ratio between numbers and to the balanced integration of the human soul with the greater order of the universe.

Transposed to the contemporary debate, this perspective offers a critical lens on the growing quantification of government. Performance indicators, scoring systems, risk metrics, and prioritization algorithms seem to materialize the ancient Pythagorean aspiration to rule by number.

Contemporary literature has rightly reminded us that quantification is far from neutral: it always presupposes choices about what goes into account, how it is measured, and for what purpose (KITCHIN, 2017). A vulnerability index, for example, can serve to expand social protection, but it can also be instrumentalized to justify selective cuts. It all depends on the way it is constructed, interpreted and put at the service of public decisions.

This ambiguity shows that numerical rationality can serve both justice and domination. In contexts of structural inequality, apparently objective systems can consolidate hierarchies, especially when statistical criteria are confused with moral judgments.

The Pythagorean School reminds us, in this sense, that the mathematical order needs to be subordinated to an ethical horizon: harmony is not only formal balance, but adequacy to the purposes of the good and the just life.

For algorithmic governance, the lesson is straightforward. Mathematical models that guide decisions about credit, security, social benefits, or urban policies should be evaluated not only for their accuracy, but for their distributive impact, discriminatory potential, and

compatibility with fundamental rights. The "beauty" of a technical solution, its algorithmic elegance, is no substitute for the political question of who wins and who loses with its application.

In conclusion, recovering the Pythagorean heritage means reaffirming that number and ethics are not separate. In a scenario in which data and algorithms progressively shape the action of the State, it becomes essential to reconnect mathematical precision with the demand for justice, preventing the fascination with metrics and models from becoming a new form of practical irrationality. Only in this way will quantification be able to fulfill its emancipatory vocation, contributing to more rational, but also more humane, policies.

2.4 DEMIURGE, FALSE ORDER, AND MODELING POWER

The figure of the Demiurge[7], presented by Plato in the dialogue Timaeus, constitutes a sophisticated metaphor for thinking about the exercise of power. Unlike an arbitrary deity, the Demiurge is described as a cosmic craftsman who, when contemplating the world of Forms, seeks to order chaotic matter according to measure, proportion and Good (Plato, 2001). It does not create from nothing, but gives form to what exists, guided by a principle of rationality that links creation and responsibility.

This image, although belonging to the ancient philosophical imaginary, offers a fruitful key to interpreting the role of contemporary institutions. Governing, whether through laws, policies, or technologies, is, to a large extent, shaping the common world: defining categories, setting priorities, organizing flows, and assigning meanings. The public power, in this sense, assumes a demiurgic function by transforming social chaos into normative, economic and symbolic order. The decisive question, then, is not only how to make this process efficient, but at the service of what values the created order is structured.

The Gnostic tradition radicalizes this reflection by introducing the figure of the false demiurge, a creative instance that institutes a degraded cosmos, marked by domination and ignorance (JONAS, 2006; PAGELS, 1995). In this reading, the order of the world is not necessarily good; It can be oppressive, yet extremely organized. The Gnostic critique works as a warning against the risk of confusing order with justice, reminding us that highly structured systems can serve to perpetuate inequalities when guided by particularistic interests.

[7] In the Platonic tradition, the Demiurge is not an absolute creator god, but an artificer who organizes chaotic matter according to models of order and beauty. In certain Gnostic readings, however, this figure gains ambiguous traits: the power that organizes can also imprison, producing a "false order" that gives the appearance of necessity to what is, in fact, a human decision. The metaphor helps to think about institutions and technologies that, in shaping the social world, require constant ethical vigilance.

Transposed to the current scenario, this ambivalence is particularly pertinent in the face of the advance of digital infrastructures and algorithmic systems. Such technologies increasingly participate in the very constitution of the social: they classify individuals, distribute visibility, mediate access to services and modulate opportunities. Its performance is not merely instrumental; they produce reality by organizing the modes of interaction and decision. In this sense, algorithms symbolically assume a demiurgic function, they become artisans of a social world mediated by data.

The critical literature has shown that this modeling power is ambivalent. On the one hand, it enables greater coordination, predictability, and analytical capacity to face complex problems. On the other hand, it can result in new forms of silent domination, in which structural decisions are shifted to technical environments that are not very transparent, making public contestation difficult (ZUBOFF, 2019; PASQUALE, 2015). The risk is the constitution of a false technocratic order: an arrangement that presents itself as inevitable and scientifically neutral, when, in fact, it crystallizes preferences and asymmetries of power.

Recognizing this demiurgic dimension of the digital state does not imply demonizing technology, but assuming it as a moral and political issue. If governing is modeling, then every act of institutional or algorithmic design requires public justification and accountability mechanisms. In the absence of these brakes, the creative capacity of power can slide into subtle forms of authoritarianism, no longer imposed by force, but by decision-making architectures that limit alternatives before they can even be discussed.

In this interpretative horizon, recent studies have recovered the metaphor of the Demiurge to think about the responsibility of those who design technosocial systems, suggesting that every institutional architecture is, above all, an ethical decision about the type of order that is intended to be established (CAVALCANTE, 2025). Such a reading reinforces that the central issue is not simply to create efficient systems, but to ensure that efficiency does not become an instrument of injustice.

In this way, the reflection on the Demiurge and its gnostic counterpart contributes to making explicit the normative core of algorithmic governance: the problem is not to create order, but to ensure that the created order is compatible with human dignity, distributive justice, and democratic integrity. It is at this point that ethics ceases to be a theoretical adornment and becomes a condition for the possibility of institutional legitimacy itself in the digital society.

2.5 PUBLIC INTEREST AND SUPREMACY IN A CONSTITUTIONAL KEY

The concept of public interest is one of the pillars of Administrative Law, functioning as a guiding criterion and limit for state action. In traditional readings, the so-called "supremacy of the public interest" was often interpreted as generic authorization for the Administration to restrict individual rights whenever it alleged reasons of collective convenience.

This vision, although functional for the construction of the intervening State, proved to be insufficient in the face of the requirements of contemporary constitutionalism, centered on human dignity and the protection of fundamental rights (DI PIETRO, 2012; JUSTEN FILHO, 2014).

Democratic constitutionalism introduces a decisive shift: the public interest is not a power above the Constitution, but a principle that must be interpreted in the light of the constitutional text and its structuring values. Supremacy thus becomes an argumentative burden: every restriction of rights needs to be rationally justified, proportional, and guided by legitimate purposes (BINENBOJM, 2017). In other words, it is not enough to invoke the public interest, it is necessary to demonstrate it.

This transformation brings Administrative Law closer to broader ethical-political debates. Rawls' theory of justice, for example, proposes that institutional arrangements are legitimate when they could be accepted by free and equal citizens under conditions of impartiality (RAWLS, 2008). Habermas' discourse ethics, on the other hand, emphasizes public justification: norms are only valid when they can find assent in communicative processes free of coercion (Habermas, 1997). Applied to the administrative field, such perspectives suggest that the public interest should be sought through transparent, motivated and socially controllable decisions.

On the practical level, this reinterpretation implies recognizing that the State operates in constant tension between efficiency and the guarantee of rights. Efficiency, an important value in complex societies, cannot serve as a pretext for the compression of individual guarantees, nor for the adoption of technocratic solutions immune to democratic scrutiny. The qualified public interest requires that administrative decisions balance results and procedures, avoiding both paralyzing formalism and opaque decisionism.

The emergence of algorithmic systems puts this debate back on new bases. By transferring a significant part of the decision-making capacity to statistical models, there is a risk of naturalizing exclusion criteria that have not been publicly debated. The literature shows that risk forecasting, benefit distribution, or policy prioritization tools can reproduce existing biases, disproportionately impacting vulnerable groups (EUBANKS, 2018; NOBLE, 2018). In

these cases, the abstract invocation of the public interest, under the argument of efficiency gains, would cover up violations of material equality and substantive due process.

It is important to note that the interpretation of the public interest as an ethical instance that conditions the use of technologies and decision-making models has gained prominence, preventing efficiency from being converted into an absolute criterion. Recent studies show that the uncritical incorporation of technocratic solutions can produce subtle forms of exclusion, especially when automated decisions start to operate as if they were neutral and inevitable (CAVALCANTE, 2025).

For this reason, the supremacy of the public interest, in a constitutional key, must be read as a principle of containment and responsibility. It obliges the manager to demonstrate: (a) the legitimate purpose of the decision, (b) the adequacy of the means used, (c) the need for less onerous alternatives, and (d) proportionality in the strict sense, weighing impacts on fundamental rights. Such requirements become even more relevant when automated systems participate in the decision, requiring impact assessments, minimal explainability, and significant human review.

Thus, the public interest ceases to be a rhetorical concept and begins to operate as an institutional moral law, which links power to the obligation to justify oneself. Far from paralyzing administrative action, this paradigm gives it ethical and democratic density, preventing the transformation of efficiency into an absolute value and ensuring that technological innovation remains subordinated to constitutional purposes.

Ultimately, understanding the supremacy of the public interest from a constitutional perspective means recognizing that the legitimacy of state action, whether analogical or algorithmic, depends on its ability to produce socially just results without giving up transparency, control, and unconditional respect for human dignity.

2.6 DEMOCRATIC GOVERNANCE AND WAVES OF ACCOUNTABILITY

The contemporary debate on governance is based on the recognition that the State is no longer the only instance capable of coordinating collective action. Social complexity, the fragmentation of interests, and the interdependence between public, private, and community actors have produced polycentric decision-making arrangements, in which multiple organizations share responsibilities (PETERS; PIERRE, 2016; RHODES, 2017). In this scenario, governing means less "commanding" and more orchestrating, creating mechanisms of cooperation, coordination and reciprocal control.

The literature identifies, in this process, successive "waves" of accountability. The first, centered on legal-formal control, sought to ensure compliance with standards and

procedures. The second, associated with managerial reforms, emphasized performance, goals and results. A third, more recent wave, replaces the values of transparency, integrity, participation, and social control at the center, articulating legal, managerial, and democratic dimensions in the same framework (BOVENS; SCHILLEMANS; HART, 2008).

This displacement does not eliminate tensions. On the contrary, it reveals that accountability is not a univocal concept, but a field of dispute between different logics of control. While managerial approaches tend to privilege metrics and indicators, democratic perspectives insist that legitimacy requires public justification, deliberation, and the possibility of contestation. Good governance results from the balance between these dimensions, avoiding both procedural formalism and productivism devoid of civic sense.

Digital transformation introduces new challenges to this picture. Information systems, platforms, and algorithms began to structure their own decision-making processes, changing who decides, how they decide, and based on what evidence.

In many cases, critical decisions are mediated by outsourced technical infrastructures, whose operating logic escapes the traditional scrutiny of public institutions. There is then the risk of a shift of power to invisible architectures, in which the criteria for classification, recommendation, and prioritization remain opaque to citizens and even to the managers themselves (PASQUALE, 2015; ZUBOFF, 2019).

This phenomenon requires broadening the concept of *accountability* beyond traditional accountability. There is increasing talk of algorithmic *accountability*, which involves transparency about the data used, model documentation, minimal explainability of results, independent risk assessment, and the possibility of meaningful human review (FLORIDI et al., 2018; WORLD BANK, 2020). It is a matter of recognizing that, when public decisions are mediated by code, the code also becomes part of the law, and must be subject to equivalent controls.

At the same time, democratic governance implies substantive participation. The opening of data and the availability of interfaces are not enough if citizens and social organizations are not integrated into deliberative processes that influence the design and monitoring of public technologies.

In this sense, co-design initiatives, public hearings on automated systems, digital ethics councils, and participatory impact evaluations represent promising ways to repoliticize technical decisions and prevent innovation from becoming a zone of normative exceptionality.

It is important to highlight that such mechanisms do not replace the individual responsibility of managers, but reconfigure it. The public leader is now responsible not only for tangible results, but for institutional architecture choices, including technological ones, that

shape behaviors and opportunities. As already suggested by theories of justice and the public interest, responsibility means being able to give reasons for decisions made and demonstrate how they align with constitutional principles and the protection of vulnerable groups.

In summary, democratic governance in the digital age demands a multidimensional accountability model, which combines: (a) legality and institutional control, (b) evaluation of performance and results, and (c) informed social participation. Without this synthesis, there is a risk of shifting the center of decisions to barely visible technical spheres, weakening public trust and bringing digital management closer to subtle forms of decisionism. With it, on the contrary, it becomes possible to align innovation, efficiency and democracy, paving the way for a governance capable of learning from its own mistakes and correcting directions, an essential condition for any public project guided by the collective interest.

2.7 ALGORITHMIC GOVERNANCE: INTERNATIONAL RISKS AND PRINCIPLES

The advancement of systems based on massive data, machine learning, and decision-making automation has ushered in a new field of ethical and legal concerns often referred to as algorithmic governance. It is about analyzing how algorithms, understood not only as mathematical codes, but as socio-technical arrangements, come to perform functions traditionally associated with public authorities: allocating resources, classifying citizens, estimating risks, and guiding policies (KITCHIN, 2017; PASQUALE, 2015). The growing centrality of these systems requires a normative framework that goes beyond engineering and dialogues with fundamental democratic values.

The specialized literature identifies three main sets of risks. The first concerns opacity. Complex, often proprietary, models make it difficult to understand why certain decisions were made, which weakens the right to explanation and hinders judicial and social control (PASQUALE, 2015).

The second is linked to the reproduction of biases. As algorithms learn from historical data, they tend to replicate and even amplify existing inequalities, disproportionately impacting minorities and vulnerable groups (O'NEIL, 2016; NOBLE, 2018; EUBANKS, 2018).

The third involves systemic effects: the simultaneous use of multiple models in areas such as health, education, security, and assistance can generate unexpected feedbacks, reconfiguring social opportunities without prior public debate.

In response to these challenges, international organizations have been formulating converging principles to guide the responsible use of artificial intelligence and automated systems. The OECD Guidelines uphold values such as inclusive growth, human well-being, transparency, technical robustness and accountability (OECD, 2019). The European Union

proposes a risk approach that combines fundamental rights, impact assessment and the requirement for meaningful human oversight (EU, 2021). UNESCO, for its part, emphasizes justice, cultural diversity, environmental protection, and cooperative global governance (UNESCO, 2021).

Despite important advances, these milestones usually face difficulties in operationalization. Many documents remain at the abstract level, lacking clear instruments to translate principles into everyday administrative practices. From this perspective, methodologies such as algorithmic impact assessments, public model registries, independent audits, and explainability protocols have been highlighted, aimed at making automated decisions and their effects traceable (FLORIDI et al., 2018; WORLD BANK, 2020). Still, the effective implementation of these tools depends on institutional capacities, technical resources, and ethical-oriented organizational culture.

Another sensitive point refers to the outsourcing of decision-making infrastructure. Governments often hire private monitoring, analysis, and prediction solutions, creating informational asymmetry between government and suppliers. This technological dependence can compromise decision-making sovereignty and make it difficult to carry out audits, especially when contracts include commercial secrecy clauses. Hence the importance of transparency by design clauses, open standards, and requiring access to data and models for control and independent research purposes.

In addition, algorithmic governance involves a pedagogical dimension. For principles of justice, non-discrimination and proportionality to be effective, it is necessary to invest in the ethical and technical training of managers, developers and evaluators. Without critical awareness, automated systems can be adopted only because they promise efficiency, without adequately evaluating their social costs. Here, the convergence between moral philosophy, law, and data science becomes essential.

In light of the above, algorithmic governance demands a multi-layered normative architecture: international principles, national legislation, sectoral regulation, responsible public procurement, and everyday organizational practices. Its objective is not to prevent innovation, but to condition it to respect for fundamental rights and the public interest. In this sense, the incorporation of mechanisms of transparency, risk assessment and citizen participation represents a necessary, although not sufficient, condition for decision-making automation in the public sector to go hand in hand with democracy, social justice and ethical responsibility.

2.8 THE GIS CYCLE AS AN INTEGRATING MATRIX

From the previous discussions, it becomes evident that the challenges posed by the digital transformation of the public sector cannot be met by piecemeal approaches. Issues of technique, law, moral philosophy, administration and citizen participation appear interconnected, requiring a conceptual framework capable of integrating multiple dimensions without reducing complexity. From this perspective, the GIS Cycle emerges as an analytical and normative proposal that articulates democratic values, technical rationality and intergenerational responsibility.

The first axis, Governance, refers to the construction of transparent, participatory institutional arrangements with robust accountability mechanisms. It is about recognizing that public decisions mediated by algorithms are not limited to technical choices, but constitute exercises of power that must remain open to criticism and social control. Governance, in this key, is not mere formal regulation: it is a permanent practice of public justification, in which managers explain reasons, listen to counterpoints and review directions when necessary.

The second axis, Innovation, is not to be confused with the uncritical adoption of technologies. On the contrary, it supposes the creation of solutions that expand rights, reduce inequalities and strengthen institutional capacities, always subordinated to the public interest. This implies incorporating, from the design of systems, principles such as bias prevention, auditability, transparent documentation, and the possibility of human review. Ethically-oriented innovation thus translates into processes of continuous institutional learning, in which technologies are evaluated not only for their efficiency, but for their distributive and democratic effects.

The third axis, Sustainability, broadens the time horizon of decisions. The discussion about algorithms cannot be restricted to the immediate impact on performance indicators; it must consider intergenerational, social, environmental and institutional consequences. Systems that concentrate power, erode public trust, or reinforce inequalities can produce short-term gains at the cost of lasting damage to social cohesion and the very legitimacy of the state. Sustainability, in this sense, operates as a long-term safeguard against technological solutions that seem ingenious, but weaken the democratic pact.

The strength of the GIS Cycle lies in their functioning as interdependent dimensions. There is no legitimate innovation without transparent governance; there is no sustainability without responsible innovation; and there is no effective governance without consideration of future consequences. This circularity dialogues with philosophical traditions that see knowledge as a process of self-transformation, getting out of the "caves" of ignorance, harmonizing numbers and ethics, guiding creative power through universal values. In

contemporary language, it means building reflective institutions, capable of learning from their own mistakes and submitting to clear ethical criteria.

From a methodological point of view, the GIS Cycle operates as a matrix for diagnosing policies, guiding the design of systems, and evaluating their implementation. It allows us to formulate structuring questions: (a) who governs and how is he accountable? (b) what innovation is being produced and at the service of whom? (c) What impacts, present and future, are being considered and for whom? Answering these questions helps to transform abstract principles of algorithmic governance into concrete decision-making procedures, bringing ethics, law, and public administration closer together.

Finally, the GIS Cycle does not intend to offer a definitive formula, but a hermeneutic and practical scheme to align technology, justice and democracy. By collecting contributions from classical philosophy, contemporary political theory, and international guidelines for responsible AI, he indicates that the legitimacy of decisions mediated by algorithms depends on the simultaneous observance of four structuring criteria, namely, *universalizability*, *dignity*, *transparency,* and *accountability*, which will guide the proposals presented in the following sections.

In this way, the theoretical framework developed here prepares the ground for normative and applied analysis, demonstrating that the digital transformation of the State is not only a technical challenge, but a historical opportunity to reconfigure the exercise of public power in the light of robust ethical principles.

3 METHODOLOGY

The present study adopts a theoretical-conceptual, qualitative and ethical-normative design, aimed at the construction of an analytical framework capable of interpreting the challenges of algorithmic governance and proposing guidelines compatible with the public interest. It is an analytical-argumentative essay that articulates philosophical tradition, contemporary constitutionalism and recent literature on governance and artificial intelligence, seeking to bring administrative theory and practice closer together (RAWLS, 2008; HABERMAS, 1997; OECD, 2019).

The empirical basis consists of a critical narrative review, appropriate for fields in consolidation and that require interpretive integration between multiple sources (BAKER, 2016). Reference works were selected in: (a) moral and political philosophy; (b) administrative law and public interest; (c) democratic governance and *accountability*; (d) international guidelines on AI and algorithms in the public sector; and (e) critical studies on biases, opacity and social impacts of automated systems (PASQUALE, 2015; O'NEIL, 2016;

EU, 2021; UNESCO, 2021). The inclusion of sources observed three criteria: thematic relevance, academic/institutional recognition, and capacity for dialogue with the Brazilian context.

Also, from the point of view of source selection, the review focused on texts that explicitly articulate technology, public ethics, and collective interest, combining classical contributions from political and moral philosophy with recent normative documents on artificial intelligence and data governance (RAWLS, 2008; HABERMAS, 1997; OECD, 2019; EU, 2021; UNESCO, 2021; WORLD BANK, 2020). The time frame privileged the production of the last decade, without excluding previous works considered structuring for the debate, as well as international reports and guidelines with a direct impact on the design of public policies. Studies of a strictly technical-engineering orientation were deliberately left out, whose focus is limited to the computational performance of the models, without a substantive interface with problems of democratic legitimacy, protection of rights and public interest.

The analysis took place in two complementary stages. First, a conceptual reconstruction of core categories was carried out, namely *truth*, *order*, *public interest*, *responsibility and governance,* examining their evolution and internal tensions. Then, the normative translation of these categories was carried out into evaluative criteria applicable to decisions mediated by algorithms, organized in the Cycle as an interpretative matrix. This procedure allows transforming theoretical foundations into operational principles, preserving ethical density and institutional viability.

Finally, it is recognized that the results result from the internal coherence of the argument and the robustness of the sources mobilized. The study is not a substitute for empirical investigations; rather, it offers a normative framework that can guide future impact assessments, case studies, and comparative analyses on the use of algorithms in the public sector. Finally, it is important to note that no systematic empirical studies have been conducted, which constitutes a future research agenda.

4 RESULTS AND DISCUSSION

Operationally, the GIS Cycle is structured as a spiral movement in four articulated moments: planning, execution, monitoring and feedback. In planning, public problems, criteria of justice and ethical risks are defined; in execution, technologies are tested with safeguards; in monitoring, distributive effects and impacts on rights are evaluated; Finally, in feedback, decisions are reviewed, corrected or interrupted, recording institutional lessons for the following cycles. This dynamic avoids irreversible decisions and transforms technological innovation into a continuous deliberative process.

The paths taken in the previous sections have shown that technology, far from being a mere neutral tool, participates in the very architecture of public power. Algorithms classify, prioritize, and distribute, and in doing so, shape concrete experiences of citizenship. If the theoretical framework has made it possible to understand the philosophical, legal and institutional bases of this phenomenon, the task that is now imposed is another: to translate principles into criteria capable of guiding real decisions.

This section takes up this deliberately difficult passage. Instead of celebrating promises or dramatizing risks, it seeks to organize a normative framework that dialogues with the complexity of the digital world, without losing sight of the public interest. It is based on the assumption that algorithmic systems should not be judged only by their efficiency, but by their compatibility with democratic values and the requirement of justification that underpins the rule of law.

In this way, the results and the discussion presented here convert the theoretical itinerary into four structuring criteria, namely, *universalizability, dignity, transparency and accountability,* integrated by the GIS Cycle. From them, guidelines for the life cycle of algorithmic systems are formulated and institutional tensions and safeguards necessary for technological innovation not to become opaque power are analyzed. More than closed answers, what is offered is a method of judgment: a way of asking, in a systematic and reflective way, what it means to govern with algorithms in a democracy.

An example helps to visualize the GIS Cycle in operation. Suppose the adoption of an algorithm for screening social benefits. In planning, the Administration defines the problem (fraud and delays), establishes fairness criteria, explainability requirements and safeguards for vulnerable groups. In execution, the system is implemented on a pilot basis, with human supervision and dispute channels. In the monitoring, patterns of undue exclusion are identified in certain territories and it is verified that the performance metric favored statistical efficiency to the detriment of equity. In the feedback phase, the parameters are reviewed, the model is independently audited, and the decision rules are rewritten, recording the lessons learned for the next cycle. The result is not just a "more accurate" algorithm, but a public policy that learns about itself and reinforces its democratic legitimacy.

The GIS Cycle is a conceptual and practical arrangement intended to guide contemporary state action, by articulating three dimensions that, historically, tend to appear dissociated: the way of deciding (governance), the ability to transform processes and services (innovation) and the normative horizon that gives meaning to public choices (sustainability). This articulation is not limited to the search for administrative efficiency; it organizes

government action in such a way as to produce legitimate, enduring, and socially recognizable public value.

Represented as an evolutionary spiral, the GIS Cycle expresses a continuous process that combines four interdependent movements: (i) planning guided by participation and evidence; (ii) implementation with institutional cooperation and ethical safeguards; (iii) monitoring and evaluation of results and impacts; and (iv) feedback, in which decisions are reviewed and learning is institutionalized. The ascending character of the spiral symbolizes the progressive improvement of government capacities, with sustainability operating as an ethical and temporal reference that projects decisions beyond the immediate present.

Figure 1

The GIS Cycle as an Evolutionary Spiral of Public Governance

Source: Prepared by the authors (2025).

The figure represents the continuous and ascending dynamics of the public management process structured in the GIS Cycle, articulating: (i) Democratic Strategic Planning, (ii) Collaborative Implementation, (iii) Monitoring and Evaluation, and (iv) Feedback and Learning. The rise of the spiral symbolizes the progressive advance of institutional capacity towards sustainability as an ethical and intertemporal horizon.

4.1 FROM THE THEORETICAL FRAMEWORK TO THE NORMATIVE CRITERIA

Although the analysis is predominantly theoretical, it is possible to visualize its practical developments in recent experiences. Facial recognition systems adopted in security policies, for example, increased false positive rates among racialized populations, evidencing structural biases in training data. In another direction, algorithms used to screen social benefits, when opaque, generated undue exclusions and reduced the capacity for administrative contestation. There are also cases of risk prediction in public health that, by prioritizing statistical efficiency, neglected territorial inequalities. These examples show that automated decisions are not mere technical resources: they reorganize opportunities, distribute burdens, and require robust legitimacy criteria.

The theoretical path developed in this study shows that algorithmic governance is not only a technological advance, but a re-actualization of the classic problem of power: who decides, based on what reasons, for the benefit of whom and under what limits? This question has crossed political philosophy since its origins and reappears mediated, now, by digital infrastructures capable of organizing sensitive dimensions of collective life.

Human moral ambivalence, present since Hobbes and Kant, reveals that social coexistence oscillates between cooperation and conflict, requiring norms that contain predatory impulses without eliminating freedom (HOBBES, 2003; KANT, 2003). In the algorithmic context, this tension intensifies: tools designed for efficiency and precision can reinforce inequalities, produce opaque decisions, and consolidate power asymmetries (PASQUALE, 2015; O'NEIL, 2016; GREEN, 2021). The Platonic image of the cave remains current: societies run the risk of confusing outputs from automated systems with objective truths, forgetting that they result from methodological choices, models and implicit values (Plato, 2000; ZUBOFF, 2019; CRAIG; BROWN, 2022).

In a convergent way, the Pythagorean tradition recalls that number is not neutral: it expresses a conception of order that can harmonize or dominate, according to the ethical horizon that guides it (BURKERT, 1972; HADOT, 2002). Indicators and algorithms not only describe reality, they contribute to constructing it, defining priorities, acceptable risks and forms of social visibility (LESSIG, 1999; BOVENS; ZOURIDIS, 2002). The metaphor of the Demiurge, critically reinterpreted by Gnosticism, deepens this diagnosis: every architecture of order can generate both justice and "false order", highly organized, but excluding (JONAS, 2006; PAGELS, 1995). In the digital public sphere, this means recognizing that algorithms play a creative role and, therefore, demand explicit moral responsibility (CAVALCANTE, 2025; CRAWFORD, 2021).

Contemporary constitutionalism replaces the public interest as a criterion of legitimacy, not as unlimited authorization, but as a duty of justification guided by human dignity and fundamental rights (DI PIETRO, 2012; BANDEIRA DE MELLO, 2015; BINENBOJM, 2017). Consequently, automated decisions need to be evaluated both by their results and by their procedures, preserving transparency, participation and social control (HABERMAS, 1997; PETERS; PIERRE, 2016; HOOD; DIXON, 2015). In Brazil, the debate is also connected to the protection of personal data and the idea of informational self-determination, reinforcing the demand for public justification (DONEDA, 2020; MENDES, 2021).

The convergence of these traditions made it possible to derive, in this work, a set of structuring normative criteria, organized in the GIS Cycle. It is argued that the legitimacy of algorithmically mediated decisions depends on the simultaneous observance of four dimensions described above. Such criteria result: (a) from the requirement that public rules can be accepted as reasonable by any citizen (KANT, 2003; RAWLS, 2008); (b) the prohibition of treating people only as means (KANT, 2003); (c) the need to make intelligible the processes that shape public truth (Plato, 2000; FLORIDI et al., 2018; EU, 2021); and (d) the democratic obligation to answer for decisions that structure collective life (HABERMAS, 1997; WORLD BANK, 2020; OECD, 2019).

The main contribution of this stage is, therefore, the passage from the interpretative to the normative plane. From the theoretical framework, principles capable of acting as ethical and legal filters for automated decisions in the State are identified. The following subsections examine each criterion and show how they are articulated in the GIS Cycle, becoming guidelines for the life cycle of algorithmic systems.

4.2 FOUR CRITERIA FOR ALGORITHMIC GOVERNANCE ORIENTED TO THE PUBLIC INTEREST

From the reconstructed theoretical framework, it becomes possible to derive a set of normative criteria capable of guiding the use of algorithms in public administration. These criteria are not external to technology; on the contrary, they emerge from the recognition that computer systems come to compose the very architecture of decision-making power, reorganizing priorities, filtering information and distributing opportunities and restrictions among citizens (LESSIG, 1999; PASQUALE, 2015). It should be noted that the idea of public interest only maintains its meaning when associated with ethical and legal parameters that condition the exercise of technical power, preventing it from becoming a mere instrumental rationality.

Based on this horizon, the study identifies the four structuring criteria already articulated within the scope of the GIS Cycle. They work as successive and complementary filters. Each one illuminates specific dimensions of the algorithmic problem, but only their joint observance allows us to speak of governance compatible with the Democratic Rule of Law.

The first criterion, *universalizability*, derives from the Kantian tradition and the contemporary theory of justice, according to which legitimate norms are those that could be accepted by any subject under equal conditions (KANT, 2003; RAWLS, 2008). Transferred to the algorithmic sphere, the principle requires that the rules embedded in computational models can be justified publicly, without arbitrary discrimination and without dependence on secret criteria. This means assessing whether automated policies are consistent with constitutional values and whether their effects do not fall disproportionately on vulnerable groups, a problem widely documented in the recent literature on data biases and algorithmic discrimination (O'NEIL, 2016; BAROCAS; SELBST, 2016; GREEN, 2021). Thus, universalizability establishes a substantive limit: decisions that could not be accepted as reasonable by all those affected lack legitimacy, even if they are technically efficient.

The second criterion, *dignity*, expands this reflection by rejecting any form of reduction of the individual to a mere object of classification, prediction or management. Dignity, the foundation of contemporary constitutionalism, requires that people not be treated only as means to administrative ends (KANT, 2003; BINENBOJM, 2017). In practice, this implies questioning models that fully replace human judgment, that deny possibilities of contestation, or that produce permanent and stigmatizing labeling, for example, risk scores, surveillance profiles, and automated screening mechanisms in social and public security policies (CITRON; PASQUALE, 2014; CRAWFORD, 2021). The criterion of dignity reminds us that, even when data allow for more accurate predictions, there are ethical boundaries that the state cannot cross without compromising its own moral legitimacy.

The third criterion, *transparency*, seeks to respond to the risk that modern societies will live in a "new cave", confusing algorithmic outputs with self-evident truths. Transparency means making visible the purposes, premises and limits of automated systems, allowing citizens, control bodies and researchers to understand their basic logic and submit it to criticism (Plato, 2000; FLORIDI et al., 2018; OECD, 2019). It is not just about disclosing source codes, which are often inaccessible or protected by trade secrets, but about producing explainability and institutional documentation, including through algorithmic impact assessments, public model registries, and clear guidelines for risk communication (EU, 2021; WORLD BANK, 2020). Without transparency, decisions become unappealable and unquestionable, shifting power to technocratic spheres of difficult democratic control.

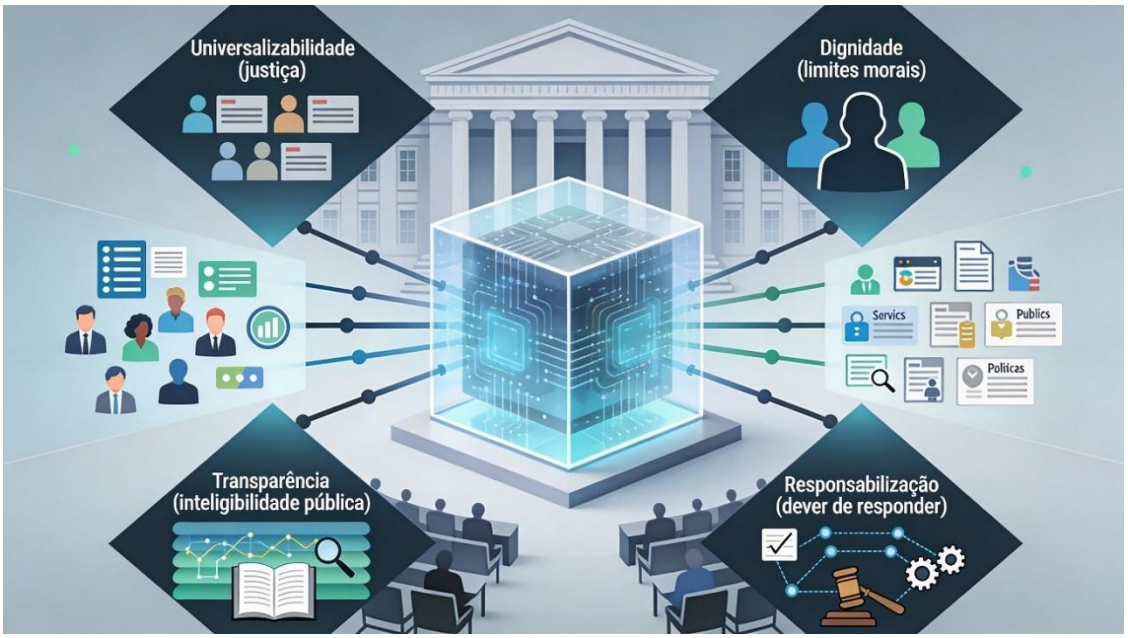
Finally, the criterion of *accountability* states that, if algorithms model social reality, it is necessary to clearly identify who is responsible for their consequences. The public governance literature warns that decision-making regimes based on automatic rules tend to spread authorship, making it difficult to assign blame, correct errors, and repair damage (BOVENS; ZOURIDIS, 2002; HOOD; DIXON, 2015). Accountability requires the definition of specific duties of diligence for managers, developers, and suppliers; it requires audit trails, review procedures and correction mechanisms when injustices are detected (HABERMAS, 1997; PETERS; PIERRE, 2016). By articulating Governance and Sustainability, this criterion prevents innovation from becoming an act without an author, a "faceless power" that no one can question.

The central contribution of this section is to show that these four criteria are not occasional lists of good practices, but contemporary translations of philosophical and constitutional requirements. They operate as mediations between abstract values and concrete decisions, allowing algorithms to be evaluated not only for their performance, but for their compatibility with justice and democracy. When incorporated into the GIS Cycle, they become stable references for the design, implementation, and monitoring of algorithmic systems, favoring institutions that are more reflective and willing to learn from their own limits.

In summary, the results indicate that the legitimacy of algorithmic governance depends on the simultaneous observance of universalizability (justice), dignity (moral limits), transparency (public intelligibility), and accountability (duty to respond). The following subsections explore how these criteria unfold throughout the life cycle of systems, from planning to discontinuity, and how they can be transformed into operational guidelines for public policies oriented to the collective interest.

Figure 2

Four criteria for Algorithmic Governance



Source: prepared by the author

To reduce the degree of indeterminacy, each criterion can be associated with operational indicators. Universalizability can be tested through impact analyses on vulnerable groups and non-discrimination audits. Dignity requires verification of meaningful human oversight, the right to object, and limits on permanent labeling. Transparency can be measured by the existence of public documentation, understandable explanations, and records of decisions. Accountability, on the other hand, requires audit trails, clear identification of decision-makers, and reparation procedures. These indicators do not exhaust the problem, but they offer verifiable parameters for institutional control.

4.2.1 Universalizability: from the moral imperative to public policy

The criterion of universalizability is based on the idea that legitimate norms are those that could be accepted by any citizen, considered equal in dignity and autonomy. In the Kantian tradition, a rule is only morally valid when it can be erected into universal law, without contradiction and without privilege (KANT, 2003). Contemporary justice theory reinforces this intuition by arguing that principles of government need to be justified from an impartial position, in which agents ignore contingent advantages and particularistic interests (RAWLS, 2008).

Transposed to the algorithmic sphere, this principle imposes that automated decisions can be publicly justified to those who will be affected. It is not enough for models to be efficient or technically sophisticated; It is necessary to verify whether the criteria embedded in the

computational routines, variables used, weights assigned, and performance metrics stand the test of reasonableness when exposed to democratic criticism. This point is particularly relevant in the face of evidence that algorithms can reproduce historical inequalities, amplifying biases present in the data and disproportionately affecting vulnerable groups (BAROCAS; SELBST, 2016; O'NEIL, 2016; GREEN, 2021).

Thus, universalizability acts as a substantive limit in the GIS Cycle, especially in the Governance axis. By requiring that technical rules can be defended as equitable, the capture of public policies by private rationalities or implicit discriminations is prevented. Algorithmic policies that would not pass the scrutiny of universalization, because they select beneficiaries in an opaque way, stigmatize territories or systematically penalize the same groups, lack legitimacy, regardless of the operational gains they present. Universalizability reintroduces, therefore, the republican ideal that no one should be subject to norms that they could not, in principle, co-sign.

4.2.2 Dignity: limits to instrumental rationality

The criterion of dignity deepens the reflection by recalling that people can never be treated as mere means for administrative, economic or security ends. In the classical formulation, dignity derives from the fact that human beings are ends in themselves, endowed with moral autonomy and the right to self-determination (KANT, 2003). In contemporary constitutionalism, this value becomes the foundation of the State and a parameter for controlling public policies (BINENBOJM, 2017; BANDEIRA DE MELLO, 2015).

Applied to the digital environment, the dignity criterion questions practices that reduce individuals to statistical profiles, risk scores or permanently suspect categories. Automated screening systems in social policies, education, or security can produce labels that are difficult to overcome, restricting future opportunities and naturalizing stigmas (CITRON; PASQUALE, 2014; CRAWFORD, 2021). The risk increases when relevant decisions are fully automated, with no real possibility of human review, converting citizens into objects of silent evaluation by machines.

Within the scope of the GIS Cycle, dignity acts as an ethical counterweight to innovation. She recalls that not every technical possibility is morally admissible. Even when algorithms promise significant efficiency gains, there is still a duty to assess whether the logic used violates essential boundaries of the human person, such as privacy, freedom of expression, equal treatment, and the right to challenge decisions that affect it (DONEDA, 2020; MENDES, 2021). Finally, dignity imposes a clear limit: it is not legitimate to govern by turning citizens into mere objects of calculation.

4.2.3 Transparency: Coming Out of the Cave in the Age of Code

Transparency responds to the danger that societies will live in a new form of cave, in which algorithmic projections acquire the appearance of incontestable truth. Since Plato, it has been known that appearances can imprison when not subjected to critical examination (Plato, 2000). In the digital context, opacity derives from technical complexity, trade secrets, and the probabilistic nature of the models, factors that make it difficult to understand why certain decisions were produced (PASQUALE, 2015; FLORIDI et al., 2018).

Transparency, however, does not mean indiscriminately disclosing everything, but making intelligible the purposes, premises and limitations of the systems used. This involves accessible documentation, algorithmic impact assessments, public model registries, the right to an explanation for affected persons, and institutional mechanisms for clear risk communication (OECD, 2019; EU, 2021; WORLD BANK, 2020). Transparency also presupposes real conditions for independent auditing, including by control bodies and researchers.

Within the GIS Cycle, transparency occupies a central position in the Governance axis, as it enables democratic control. Without it, decisions become unappealable, and authority moves to technocratic spheres that are not very visible. By illuminating decision-making processes, transparency prevents algorithms from operating as "black boxes" and allows society to distinguish between fair outcomes and mere shadows cast by statistical models. It is, therefore, a requirement for technological innovation to remain subordinate to the public sphere, and not the other way around.

4.2.4 Accountability: the power to model and the duty to respond

The criterion of accountability closes the set by stating that, if algorithms model the social world, someone must answer for their consequences. The public administration literature shows that automated systems tend to disseminate the authorship of decisions, making it difficult to identify those responsible and correct injustices (BOVENS; ZOURIDIS, 2002; HOOD; DIXON, 2015). Without clear rules, the phenomenon of "faceless power" occurs, in which errors are attributed to the machine and not to the human choices that configured it.

Accountability, in this context, means establishing duties of diligence throughout the entire life cycle of the system: planning, development, contracting, implementation, monitoring, and discontinuity. Each stage must have clearly identified agents, with obligations to document, test, review and correct when damage is detected (HABERMAS, 1997;

PETERS; PIERRE, 2016). This includes the right to reparation and the existence of institutional channels for contesting and reviewing automated decisions.

In the GIS Cycle, accountability integrates Governance and Sustainability: it prevents innovation from being an act without an author and ensures institutional learning over time. By recognizing that every technological architecture incorporates moral and political choices, the criterion obliges managers to justify, record, and publicly answer for them. With this, it transforms the automation of a democratic risk into an opportunity to strengthen the culture of accountability, an indispensable condition for the use of algorithms to remain compatible with a State that is intended, above all, to be republican and democratic.

4.3 GUIDELINES FOR THE LIFE CYCLE OF ALGORITHMIC SYSTEMS IN THE PUBLIC SECTOR

The normative criteria presented, universalizability, dignity, transparency and accountability, only reveal their full potential when translated into concrete guidelines for the life cycle of algorithmic systems. Recent literature shows that many ethical failures do not stem from bad intentions, but from the absence of institutional processes capable of anticipating risks, monitoring effects, and correcting course (WORLD BANK, 2020; OECD, 2019). Thus, more than listing good practices, it is about organizing a sequence of public decisions consistent with the GIS Cycle, in which governance, innovation, and sustainability act in an integrated way.

From the point of view of public management, the proposal developed here does not intend to replace existing norms, but to offer a decision roadmap. In practical terms, this means incorporating algorithmic impact assessments into contracting processes, recording the technical and legal motivations for choices, instituting formal channels for contestation, and providing for periodic reviews of the systems in place. It is, therefore, a matter of translating ethical principles into verifiable administrative routines, strengthening *accountability* without paralyzing innovation.

4.3.1 Planning: defining the problem before the solution

The first moment is to prevent technology from taking precedence over politics. Before hiring tools or developing models, the government must clearly delimit the public problem, its objectives and the values involved. The initial question is not "which algorithm to use?", but whether the use of algorithms is necessary, proportionate and justifiable, compared to less intrusive alternatives (LESSIG, 1999; MENDES, 2021).

At this stage, it is necessary to explain which goods will be prioritized, reduction of inequalities, expansion of access, improvement of targeting and how such goals dialogue with the criteria of universalizability and dignity. When values and purposes remain implicit, space is opened for technical decisions to replace political choices, shifting democratic deliberation into codes and databases (PASQUALE, 2015).

4.3.2 Design and development: incorporating ethics from the beginning

Once the problem is defined, the design of the system must occur in an interdisciplinary way. Technical teams need to dialogue with professionals in law, philosophy, public policy, and institutional control, preventing modeling from being the exclusive domain of data specialists. The literature emphasizes that apparently neutral options, choice of variables, treatment of missing data, performance metrics, carry normative decisions (BAROCAS; SELBST, 2016; GREEN, 2021).

It is therefore recommended to document justifications, carry out prior algorithmic impact assessments and identify, from the outset, possible discriminatory effects, especially on historically vulnerable groups (EU, 2021; OECD, 2019). In cases of contracting with third parties, transparency clauses, access for auditing, and clear rules for data ownership and sharing become essential. The axis of Governance, in this phase, consists of recognizing that each line of code translates a public policy option.

4.3.3 Implementation: human oversight and right to challenge

During implementation, systems should not fully replace human judgment, especially when they produce high-impact decisions on fundamental rights. The literature points out that public trust depends on the possibility of revision and the existence of a way out for unfair decisions (CITRON; PASQUALE, 2014; DONEDA, 2020).

Thus, it is recommended that algorithms operate as decision support, and not as an indisputable final authority. Affected citizens need to have accessible channels to question results and obtain understandable explanations, which at the same time reinforces the criteria of dignity and transparency. From the point of view of institutional sustainability, this stage requires continuous monitoring, with the collection of evidence on systematic errors, biases, and unforeseen effects, transforming implementation into a learning process.

4.3.4 Monitoring, review, and discontinuity: learning from the effects

No algorithmic system should be considered definitive. Use over time produces externalities, alters behaviors, and can generate dependencies that are difficult to reverse.

For this reason, international best practices recommend periodic reviews, with the participation of control bodies, independent experts and, when possible, representatives of civil society (WORLD BANK, 2020; EU, 2021).

These reviews should evaluate not only technical performance, but also distributive impacts, respect for rights, and compatibility with the public interest. When adverse effects prove to be serious, discrimination, unappealable opacity, erosion of fundamental guarantees, the State needs to have clear mechanisms of correction or discontinuity, including adequate reparation for injured persons (HABERMAS, 1997; PETERS; PIERRE, 2016).

This last movement materializes the axis of Sustainability in the GIS Cycle: recognizing that technological decisions leave institutional and intergenerational traces and, therefore, require the ability to review, record, and respond to the paths adopted.

4.3.5 Synthesis: from technique to self-government

The guidelines presented show that governing algorithms means, ultimately, governing the exercise of power itself. When careful planning, ethical design, responsible implementation, and reflective monitoring are articulated, digital transformation ceases to represent a technocratic threat and becomes an opportunity to deepen democracy.

The key is to understand that technological neutrality is a myth. Algorithmic systems incorporate worldviews, moral choices, and distribution criteria. Submitting them to the GIS Cycle, through the four reconstructed normative criteria, means putting politics and ethics back at the center of the decision-making process, preventing technology from becoming a new incontestable "Demiurge" (JONAS, 2006; CRAWFORD, 2021).

Thus, the results suggest that public policies mediated by algorithms only achieve legitimacy when anchored in processes that deliberate, explain and respond, three verbs that synthesize the passage from technological innovation to a true democratic governance of the digital.

4.4 INSTITUTIONAL TENSIONS, RISKS AND SAFEGUARDS

The analysis carried out so far shows that algorithmic governance does not eliminate historical dilemmas of public power; on the contrary, it reconfigures them into new forms. Instead of conflicts between bureaucracy and politics, or between efficiency and legality, tensions emerge between technical complexity and democratic control, between algorithmic personalization and equality, between massive data and privacy, between speed of innovation and institutional prudence (HOOD; DIXON, 2015; OECD, 2019). Acknowledging

these tensions is a decisive step toward formulating safeguards that don't demonize technology but also don't capitulate to its allure.

A first tension concerns the promise of efficiency. Automated systems are often presented as rational answers to problems of resource allocation and policy focus. However, efficiency gains can occur at the expense of rights, for example, when automation increases the undue exclusion of beneficiaries or makes it difficult to challenge administrative decisions (PASQUALE, 2015; CITRON; PASQUALE, 2014). The literature shows that technological solutions tend to hide values and choices under the appearance of neutrality, shifting moral debates into mathematical models (CRAWFORD, 2021; GREEN, 2021). The criterion of universalizability, in this context, works as a prudential brake, requiring that "efficient" policies also be justifiable to those affected.

A second tension involves structural opacity. Complex models, such as neural networks, have high performance, but make it difficult to explain specific results. When associated with trade secrets and restrictive contracts, they produce zones of public unintelligibility, in which not even managers can explain why someone was selected, classified or excluded (FLORIDI et al., 2018; EU, 2021). Added to this is the risk of excessive "epistemic authority": decisions tend to be accepted simply because "the algorithm said so." Transparency, here, ceases to be a mere administrative virtue and becomes a condition for the possibility of democratic control.

There are also tensions linked to technological dependence and the asymmetry of power between public administrations and large private suppliers. Contracting proprietary systems can generate institutional lock-in, making it difficult to audit, review, and migrate to more rights-compatible alternatives (WORLD BANK, 2020). This scenario increases the importance of contractual clauses that guarantee access to data, documentation, and logs, as well as the need for internal technical capabilities capable of critically dialoguing with providers (BOVENS; ZOURIDIS, 2002). Accountability, distributed throughout the life cycle, prevents structural failures from being naturalized as "inevitable machine errors."

Another relevant risk is the normalization of monitoring. The growth of databases and surveillance devices can stimulate preventive policies based on behavioral prediction, with profound impacts on privacy and freedom. Even when motivated by legitimate ends, such practices threaten to turn citizens into permanent objects of observation, eroding social trust (ZUBOFF, 2019; DONEDA, 2020). In these cases, the criterion of dignity reminds us that not every administrative benefit authorizes sacrificing essential spheres of autonomy.

In the face of this set of tensions, institutional safeguards play a decisive role. International experiences point to three minimum layers. The first is normative, involving data

protection laws, responsible AI frameworks, the requirement for impact assessments, and active advertising rules (OECD, 2019; EU, 2021). The second is organizational, with ethics committees, units specialized in data governance, audit protocols, and formal instances of social participation. The third is cultural, dependent on the continuous training of managers and civil servants to understand that technology is always a political decision mediated by technique, and not a substitute for deliberation.

These safeguards, when integrated into the GIS Cycle, make it possible to transform risks into institutional learning opportunities. Governance creates spaces of justification; innovation is submitted to ethical and legal criteria; Sustainability ensures that decisions can be reviewed in light of their intergenerational effects. The tension does not disappear, but becomes governable.

In summary, the results suggest that the central challenge is not to choose between adopting or rejecting algorithms, but to decide how to adopt them, under what limits, and with what guarantees. Automated systems can deepen inequalities, entrench opacities, and erode rights, but they can also expand administrative capacities and improve resource allocation when anchored in clear criteria. The decisive point is that, in the absence of safeguards, technology tends to assume the role of a new "world orderer", without justification. With robust safeguards, it becomes an instrument subordinate to the democratic project.

5 CONCLUSION

The path developed throughout this article started from an apparently simple concern: what actually changes when the State starts to decide with the support of algorithms? The answer proved to be broader than a mere technological update. By articulating classical philosophy, contemporary constitutionalism, and recent debates on governance and artificial intelligence, it has become evident that the central issue is not the "use of tools", but the reconfiguration of the exercise of power in societies increasingly mediated by data and computational models.

Human ambivalence, the Myth of the Cave, the Pythagorean heritage, the metaphor of the Demiurge, and the reconstruction of the supremacy of the public interest in a constitutional key converged to the same diagnosis: every form of order, including the digital order, carries a normative core. Algorithmic systems don't just describe reality; They model it, defining who shows up, who gets monitored, who will have access to policies, and who will stay on the margins. In this context, insisting on the neutrality of technology is equivalent to confusing, again, shadows with reality, only with more sophisticated mathematical tools.

The main contribution of this work was to transform this finding into a structured normative framework, organized in the GIS Cycle. Based on it, four criteria were identified that should guide public decisions mediated by algorithms: universalizability, dignity, transparency, and accountability. These criteria do not constitute an occasional list of good practices, but a contemporary translation of requirements already present in moral philosophy, in the theory of justice and in administrative law. Universalizability reminds us that decisions are only legitimate when they can be justified to all those affected; dignity imposes limits on the reduction of people to mere objects of calculation; transparency prevents the technique from becoming a new form of cave; accountability ensures that the power to shape reality does not dissolve into the anonymity of the machine.

By applying these criteria to the life cycle of algorithmic systems, from planning to discontinuity, the paper showed that there is no responsible governance without structured processes. Defining the problem before the solution, designing models with interdisciplinary participation, ensuring meaningful human oversight, opening channels of contestation, instituting periodic reviews, and providing for the possibility of interruption of systems incompatible with fundamental rights are not optional gestures: they are minimum conditions for technological innovation to remain subordinate to the public interest, and not the other way around. When these safeguards do not exist, automation tends to reinforce inequalities, opacities, and dependencies; When they exist, real possibilities are opened up to expand state capacities and qualify public policies.

It is important to recognize, however, the limits of the proposal. It is a study of a theoretical-conceptual nature, whose strength lies in the internal coherence of the argument and in the pertinence of the sources mobilized. The normative model presented here does not replace empirical investigations, field experiments or comparative case studies. On the contrary, it intends to prepare the ground so that these surveys can be conducted with greater clarity of criteria, avoiding both technofetishism and technopessimism. A promising agenda involves, for example, the application of the GIS Cycle to the analysis of systems already in use in areas such as health, education, social assistance, public security, and fiscal management, as well as the comparison between different institutional arrangements of algorithmic governance at the national and international levels.

Despite these limits, the central argument remains: there is no technical way out of a problem that is, above all, ethical and political. The growing presence of algorithms in the action of the State does not dispense with democratic deliberation; on the contrary, it makes it more urgent. In a scenario in which decisions tend to be justified by the authority of statistical models, reaffirming the need to justify oneself before concrete citizens and not just before

indicators is a gesture of democratic resistance. It is also a way of remembering that the State, even when digital, remains bound by the commitment to treat each person as an end in himself.

Ultimately, governing algorithms is a way of governing oneself as an institution. By proposing that all technological innovation in the public sector be passed through the sieve of universalizability, dignity, transparency, and accountability, this article suggests that the digital society does not need to be a new chapter of heteronomy imposed by the code. On the contrary, it can become an opportunity to deepen the ideal of a power that knows itself to be dangerous, and therefore limits itself; he knows himself to be fallible, so he lets himself be controlled; he knows he is a creator, so he is responsible for the world he helps to build. If this intuition is taken seriously, the digital transformation of the State will no longer be just a project of administrative modernization and can be understood as what, in fact, is at stake: a new stage in the history of the struggle for a fairer, more reflective, and truly democratic public governance.

REFERENCES

Arendt, H. (2014). A condição humana (10ª ed.). Forense Universitária.

Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. European Law Journal, 13(4), 447–468.

Brasil. (2018). Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Presidência da República.

Cavalcante, A. F. (2025). O Demiurgo e a construção do real algorítmico: Ética, poder e responsabilidade no Estado digital. Revista Aracê, 7(12), 1–28.

Cavalcante, A. F., & Lima, J. S. (2025). Entre o planejamento e a democracia: Orçamento, participação popular e o Ciclo GIS. Revista PPC – Políticas Públicas e Cidades (no prelo).

Crawford, K. (2021). Atlas of AI: Power, politics, and the planetary costs of artificial intelligence. Yale University Press.

European Commission. (2021). Proposal for a Regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act).

Green, B. (2022). The smart enough city: Putting technology in its place to reclaim our urban future. MIT Press.

Habermas, J. (1997). Direito e democracia: Entre facticidade e validade. Tempo Brasileiro.

Hobbes, T. (2003). Leviatã. Martins Fontes.

Jonas, H. (2006). O princípio responsabilidade. PUC-Rio; Loyola.

Kant, I. (2003). Fundamentação da metafísica dos costumes. Martins Fontes.

Kitchin, R. (2017). The data revolution: Big data, open data, data infrastructures and their consequences. SAGE.

Moore, M. (1995). Creating public value: Strategic management in government. Harvard University Press.

Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. NYU Press.

Organisation for Economic Co-operation and Development. (2019). OECD Principles on Artificial Intelligence. OECD.

O'Neil, C. (2016). Weapons of math destruction. Crown.

Osborne, S. (2006). The new public governance? Public Management Review, 8(3), 377–387.

Pasquale, F. (2015). The black box society. Harvard University Press.

Platão. (2006). A República. Martins Fontes.

Rawls, J. (2008). Uma teoria da justiça. Martins Fontes.

United Nations Educational, Scientific and Cultural Organization. (2022). Recommendation on the Ethics of Artificial Intelligence. UNESCO.

Zuboff, S. (2019). The age of surveillance capitalism. PublicAffairs.