noemicajournal.org

e-ISSN: 3086-7037

ARTICLE

State capacities for digital transformation: a theoretical model of responsible artificial intelligence governance in local governments

Capacidades estatais para a transformação digital: um modelo teórico de governança responsável da inteligência artificial nos governos locais

Capacidades estatales para la transformación digital: un modelo teórico de gobernanza responsable de la inteligencia artificial en los gobiernos locales

Robson Barbosa[1]

[1] Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP), São Paulo, SP, Brasil. ORCID: https://orcid.org/0009-0009-5896-0323

SUBMITTED	ACCEPTED	PUBLISHED	DOI
20/03/2025	30/10/2025	05/12/2025	in process

ABSTRACT

Local governments face increasing pressure to adopt artificial intelligence in policymaking and service delivery, but they often operate under constraints of personnel, infrastructure, data, coordination, and institutional control. This article develops a theoretical model of state capacities aimed at the responsible governance of artificial intelligence in local governments. The approach is theoretical and integrative, drawing together the literatures on state capacity, digital government, data governance, public value, and trustworthy artificial intelligence. The result is a set of six interdependent capacities: strategic, regulatory, technical, informational, organizational, and participatory. The model shows that weaknesses in any of these dimensions tend to produce vendor dependency, fragmented automation, opaque decisions, and widening inequalities. The central conclusion is that digital maturity is not measured by the number of tools adopted, but by the public sector's ability to select, oversee, evaluate, contest, and, when necessary, discontinue systems, while balancing efficiency, rights, transparency, accountability, and public value.

Keywords: state capacities, local government, digital transformation, artificial intelligence, public governance

ABSTRACT

Local governments face increasing pressure to use artificial intelligence in policymaking and service delivery, although they frequently experience limitations in personnel, infrastructure, data, coordination, and institutional control. This article develops a theoretical model of state capacities for responsible artificial intelligence governance in local governments. It adopts a theoretical-integrative approach combining state capacity, digital government, data governance, public value, and trustworthy artificial intelligence. Six interdependent capacities are proposed: strategic, regulatory, technical, informational, organizational, and participatory. The model indicates that weaknesses in any dimension may generate vendor dependency, fragmented automation, opaque decisions, and increased inequalities. Digital maturity should not be equated with the number of tools adopted, but with the public sector's capacity to select, oversee, assess, contest, and, when necessary, discontinue systems while balancing efficiency, rights, transparency, accountability, and public value.

Keywords: state capacity, local government, digital transformation, artificial intelligence, public governance

RESUMEN

Los gobiernos locales enfrentan una presión creciente para utilizar inteligencia artificial en la formulación de políticas y la prestación de servicios, aunque con frecuencia presentan limitaciones de personal, infraestructura, datos, coordinación y control institucional. Este artículo desarrolla un modelo teórico de capacidades estatales para la gobernanza responsable de la inteligencia artificial en los gobiernos locales. Se adopta un enfoque teórico-integrador que articula capacidades estatales, gobierno digital, gobernanza de datos, valor público e inteligencia artificial confiable. Como resultado, se proponen seis capacidades interdependientes: estratégica, regulatoria, técnica, informacional, organizacional y participativa. El modelo indica que las debilidades en cualquiera de estas dimensiones pueden producir dependencia de proveedores, automatización fragmentada, decisiones opacas y ampliación de desigualdades. Se concluye que la madurez digital no corresponde al número de herramientas adoptadas, sino a la capacidad pública para seleccionar, supervisar, evaluar, impugnar y, cuando sea necesario, descontinuar sistemas, equilibrando eficiencia, derechos, transparencia, responsabilidad y valor público.

Palabras clave: capacidades estatales, gobierno local, transformación digital, inteligencia artificial, gobernanza pública

INTRODUCTION

The incorporation of artificial intelligence-based systems into public administration has changed the means by which governments diagnose problems, organize services, and interact with citizens. Tools for classification, prediction, screening, and automated service delivery have come to be embedded in processes that, until recently, depended almost exclusively on the judgment of civil servants. At the local level, these technologies already appear in health, education, social assistance, mobility, inspection, taxation, and urban planning, often arriving before any specific municipal regulation governing their use.

Municipal governments occupy a peculiar position in this process. They are responsible for the services closest to people's daily lives and maintain direct contact with the population, yet they generally have less fiscal, technical, and informational capacity than state and federal governments. This asymmetry produces a tension between the pressure for modernization, fueled both by social demands and by the commercial supply of artificial intelligence solutions, and the real capacity to select, procure, oversee, and evaluate complex systems. In many cases, the adoption of a tool originates from the isolated initiative of a single department, a vendor's offer, or an emergency need, without any municipal strategy capable of guiding future decisions.

The digitalization of administrative processes is not a neutral phenomenon. Automated systems reorganize authority, alter workflows, and influence the distribution of benefits and constraints among citizens. When an algorithm prioritizes service appointments, identifies vulnerable families, calculates default risk, or selects areas for inspection, it comes to embed criteria that must be understandable and open to challenge. The technical opacity of these systems can turn choices that are essentially administrative, and therefore political, into decisions that appear objective merely because they arise from a calculation, thereby reducing the visibility of who, ultimately, answers for them.

Given this scenario, this article is guided by the following question: what state capacities are necessary for local governments to govern artificial intelligence responsibly and in the public interest? The aim is to develop a theoretical model that articulates strategic, regulatory, technical, informational, organizational, and participatory competencies, and that can serve as a reference for both academic analysis and institutional planning. The central argument is that digital maturity depends less on the quantity of tools deployed and more on the administration's capacity to govern them throughout their entire life cycle, from conception to eventual discontinuation.

The relevance of the topic stems from a risk seldom discussed in the more enthusiastic debates about innovation: the possibility that digital transformation simultaneously expands efficiency, inequality, and dependency. Municipalities with low institutional capacity run the risk of becoming captive consumers of proprietary solutions, unable to audit their own administrative decisions or to retain technical knowledge when a contract ends or a vendor changes its commercial strategy. The model proposed here seeks to offer a framework for analysis, planning, and evaluation for this reality, without assuming that all automation is desirable or that all resistance to automation amounts to backwardness.

The article is organized into seven sections, in addition to this introduction. The second section presents the theoretical framework, bringing together the literature on state capacity, digital government, data governance, trustworthy artificial intelligence, and bureaucratic discretion. The third section describes the methodological path of the integrative review. The fourth section presents the model of six capacities and discusses their interdependencies. The fifth section details the life cycle of artificial intelligence in local government and the mechanisms for risk assessment, procurement, and accountability. The sixth section discusses applications of the model, maturity indicators, federative cooperation, people management, transparency, financial sustainability, and an agenda of propositions for future empirical research. The seventh section presents the concluding remarks.

THEORETICAL FRAMEWORK

The theoretical foundation of this article draws together five bodies of literature that rarely appear integrated within a single model: historical sociology and political science on state capacity, public administration studies on digital government and data governance, the ethics and governance literature on artificial intelligence, and studies on bureaucratic discretion and street-level bureaucracy. This combination makes it possible to treat digital transformation as a political and organizational process, rather than merely as a technical innovation to be adopted or rejected.

State capacities: from the classical tradition to the digital agenda

The concept of state capacity describes the set of attributes that enable a government to formulate objectives, mobilize resources, and implement policies effectively. The classical literature of historical sociology brought this concept back to the center of the debate on development and governance, emphasizing professional bureaucracy, autonomy from interest groups, internal coordination, and the capacity for social penetration (Skocpol, 1985; Evans, 1995). Evans (1995) coined the expression "embedded autonomy" to describe bureaucracies that combine technical competence and political insulation with dense ties to social and economic actors, a combination that allows the state both to formulate coherent policies and to obtain information and cooperation from society.

Fukuyama (2013, 2004) contributed to this tradition by distinguishing state capacity from the scope of state action and from the quality of governance. For the author, a state may have a broad scope, that is, act in many areas, and still have low capacity to execute well what it sets out to do; the strength of the state, in this sense, refers to its ability to plan and execute policies and to enforce laws transparently. This distinction is useful for thinking about local governments that rapidly expand the scope of their digital activities, acquiring artificial intelligence systems across several areas, without the capacity to operate and oversee them growing at the same pace. Centeno, Kohli, Yashar, and Mistree (2017) reinforce this point by showing, in comparative perspective, that state capacities vary significantly across countries and regions and that this variation explains much of the difference in policy effectiveness, even when formal institutional designs are similar.

These classical dimensions remain relevant for thinking about contemporary municipal administration, but digitalization adds specific requirements related to data, software, information security, contract management, and specialized technical knowledge. In the digital environment, capacity is not limited to possessing infrastructure or usage licenses. It involves understanding public problems in enough depth to know what kind of technological solution is proportionate to them, choosing tools suited to the scale and complexity of the problem, establishing limits on use, and retaining some degree of control over decisions that come to be supported, wholly or partly, by automated systems. A municipality may acquire an advanced artificial intelligence platform and still exhibit low state capacity if it depends entirely on the vendor to operate, interpret, and correct the solution, a situation that is quite common in small and medium-sized administrations.

Digital transformation also reconfigures the relationship between autonomy and cooperation, one of the central axes of the state capacity literature. Local governments frequently need the support of other levels of government, universities, private companies, and civil society organizations to carry out artificial intelligence projects, but they must retain enough capacity to define public purposes and protect the collective interest against commercial interests or external agendas. Productive partnerships differ structurally from relationships of dependency: in cooperation, the state retains the competence to evaluate proposals, negotiate terms, and decide based on its own criteria; in dependency, the state accepts externally defined criteria, with no real possibility of contesting them or replacing the partner.

Digital government, platformization, and vendor dependency

The shift from electronic government to digital government moved the focus away from simply making services available through electronic channels toward data integration, process redesign, and the strategic use of information as a management resource. Dunleavy, Margetts, Bastow, and Tinkler (2006) described this movement as the emergence of a "digital-era governance," characterized by the reintegration of functions previously fragmented by new public management and by the holistic digitalization of processes that come to be designed around the user, rather than merely around the internal structure of the agency. Gil-Garcia, Dawes, and Pardo (2018) note that digital government research has matured by recognizing that technology, organization, and politics mutually determine one another, so that the success of a digital initiative cannot be explained by its technical quality alone.

More recently, international organizations have come to treat the digital transformation of the public sector as a governance problem, not merely an infrastructure one. The Organisation for Economic Co-operation and Development, through its Recommendation on Digital Government Strategies (OECD, 2014) and its Digital Government Policy Framework (OECD, 2020), argues that digitalization should be guided by user-centered design, openness by default, data capacity, and digital leadership, understood as governments' ability to steer, rather than merely follow, technological transformation. This framing is particularly relevant for municipalities, which tend to occupy the position of technology takers in the face of a supplier market that is far more concentrated and better informed than the purchasing administration itself.

The integration of data and processes, however, can occur through proprietary platforms that concentrate technical knowledge and contractual power in the hands of a few vendors. Vendor dependency, discussed in the information systems and public administration literatures under the term vendor lock-in, arises when contracts, data, and internal processes become tied to technologies that are difficult to replace. Municipalities with small teams tend to accept off-the-shelf solutions without demanding adequate documentation, data portability, or knowledge-transfer clauses. Over time, the cost of switching vendors becomes high enough that the administration loses real capacity to negotiate better terms or to discontinue a system that no longer serves the public interest.

The public procurement of artificial intelligence solutions therefore needs to consider the technology's full life cycle, not merely the moment of acquisition. An attractive initial price may conceal recurring costs of licensing, maintenance, integration with other systems, data storage, and model updates. It is also necessary to assess, from the tender stage onward, what rights the administration holds over the data produced, access to audit logs, deadlines for reporting security incidents, and mechanisms for contractual oversight. The idea of local technological sovereignty, in this sense, does not imply that a municipality should internally develop everything it uses, which would be unfeasible for the vast majority of Brazilian municipal administrations, but rather that it should retain decision-making autonomy over the purposes, criteria, and continuity of the systems it adopts.

Data governance and information infrastructure

Artificial intelligence depends, above all, on the quality of the data on which it operates. Local governments frequently accumulate fragmented, duplicated, and outdated databases, the product of decades of isolated systems created to meet the specific demands of each department. Kitchin (2014) draws attention to the fact that data are never "raw": they carry methodological decisions, categories, and omissions that reflect how the institution that produced them viewed its own work. Different departments often use distinct concepts to represent the same people, territories, or services, which hampers the integration of databases and creates inconsistencies that go unnoticed until an automated system amplifies them. Automating fragile databases does not correct this problem; on the contrary, it tends to accelerate errors and widen exclusions that already existed but were previously diluted when decisions depended on case-by-case human analysis.

Data governance, as a field of practice and research, encompasses technical standards, the definition of responsibilities, information security, quality control mechanisms, interoperability among systems, and the ethical use of information. Janssen and van den Hoven (2015) point out that the opening of government data, a topic that has gained momentum over the past decade, needs to be reconciled with privacy and the protection of vulnerable groups: not every piece of public data can be disclosed or combined with other databases without a risk of re-identification or undue exposure of sensitive information. It is therefore necessary to define clearly who produces, validates, accesses, and corrects each type of information, and under what conditions it may circulate among agencies.

The informational capacity of a local government also depends on basic infrastructure and operational continuity, dimensions that tend to receive less attention than the more visible aspects of innovation. Municipal systems need routines for backups, periodic updates, access control, and recovery plans in the event of failures or cyberattacks. The unavailability of an essential database can interrupt services on which the population depends directly, such as health or social assistance registries. The governance of artificial intelligence, in this sense, cannot be treated as an isolated topic; it needs to be integrated with the information and security management that should already exist, regardless of the use of more sophisticated algorithms.

Trustworthy artificial intelligence, explainability, and public value

The international literature on trustworthy artificial intelligence has emphasized principles of lawfulness, ethics, and technical robustness as conditions for automated systems to be acceptable in high social-impact contexts (Floridi et al., 2018). In the public sector, these principles need to be articulated with the notion of public value, developed by Moore (1995) and revisited by Bozeman (2007) in his discussion of public values and the collective interest. An artificial intelligence solution cannot be evaluated solely by criteria of efficiency or statistical accuracy; it must also be examined in terms of equity in the distribution of its outcomes, transparency about how it works, the legitimacy of the process that led to its adoption, and its actual capacity to produce collective benefits, rather than merely internal administrative gains.

The explainability of automated systems has both a technical and an institutional dimension, which tend to be treated separately but need to converge in administrative practice. A model may be statistically complex, and yet the public decision that results from it must be justified in language comprehensible to those affected by it. A citizen affected by an automated, or partly automated, decision should be able to know which criteria were relevant to the outcome, who is institutionally responsible for that decision, and how to request its review. Raji et al. (2020) describe this gap as a problem of algorithmic accountability: systems are frequently deployed without mechanisms that allow for external auditing or the correction of errors identified after use. Without these elements, automation tends to weaken procedural rights that public administration had already recognized before the introduction of any algorithm.

The robustness of an artificial intelligence system also needs to include a notion of social performance, in addition to purely technical performance. A system may function exactly as specified and still produce unjust outcomes when that specification failed to adequately consider the diversity of the populations served. Performance tests should assess the system's behavior across different groups and territories, rather than relying only on overall averages that may mask impacts concentrated on specific populations. Responsible governance therefore requires continuous monitoring throughout the entire operation of the system, not only during the approval phase, along with the institutional capacity to discontinue a use when identified harms outweigh the expected benefits.

Algorithmic bias and bureaucratic discretion

Algorithms trained on historical data tend to reproduce inequalities present in the very administrative practices that generated that data. Incomplete records about certain neighborhoods or social groups can lead a system to underestimate real needs, simply because those needs were never adequately recorded in the past. Seemingly neutral variables, such as address or service history, can function as indirect proxies for income, race, or territory, reproducing patterns of discrimination without this appearing explicitly in the system's design (O'Neil, 2016; Eubanks, 2018). Eubanks (2018) documents, in North American case studies, how automated screening systems in social policy tend to intensify surveillance over already vulnerable populations, while publicly presenting themselves as tools of efficiency and objectivity.

The street-level bureaucracy literature offers an important vocabulary for understanding this transformation. Lipsky (2010) showed that front-line workers, such as social workers, teachers, and police officers, exercise considerable discretion when applying public policies in concrete situations, filling gaps that general rules cannot anticipate. Bovens and Zouridis (2002) described the shift from street-level bureaucracies to system-level bureaucracies, in which decisions that once depended on an individual worker's judgment come to be mediated, or even replaced, by standardized information systems. More recently, Busch and Henriksen (2018) systematically reviewed the literature on digital discretion, concluding that information technologies do not eliminate administrative discretion but shift it to other points in the process, often less visible than the moment of direct interaction with the citizen.

Replacing traditional bureaucratic discretion with automated rules does not eliminate value-laden choices; it merely shifts them to earlier stages, such as defining the problem to be solved, selecting the variables that make up the model, and configuring the parameters that determine what counts as an acceptable outcome. Bullock (2019) argues that this shift of discretion to the system-design stage requires public policy professionals to participate actively in the development and evaluation of artificial intelligence tools, not only in their final approval, lest substantive decisions be made by technical teams lacking the necessary knowledge of the social context in which the system will be applied. Barth and Arnold (1999) had already anticipated, in one of the earliest articles on the topic in North American public administration, that expert systems tend to reduce the discretion perceived by civil servants without necessarily reducing the actual discretion embedded in the system, a point that remains relevant more than two decades later.

Human oversight of automated systems needs to be effective, not merely formal. Placing a civil servant at the end of the decision process, tasked with approving the outcome suggested by the system, is not sufficient when that worker lacks the information, time, or institutional autonomy to override the automated recommendation. Peeters (2020) calls this phenomenon the limited agency of the human operator vis-à-vis the algorithm, a situation in which oversight becomes symbolic because the cost of disagreeing with the system, in terms of time and required justification, is systematically higher than the cost of simply accepting its recommendation. Responsible governance therefore needs to clearly define which situations require mandatory human review, formally record cases of divergence between worker and system, and institutionally protect the exercise of professional judgment, including when that judgment contradicts the automated recommendation.

METHODOLOGICAL APPROACH

The study uses the integrative theoretical review as its research strategy, an approach recognized as suitable for building conceptual models that bring together distinct fields of knowledge that do not yet engage systematically with one another (Torraco, 2016). Unlike a systematic review aimed at synthesizing empirical evidence on a specific question, the integrative review allows for combining theoretical literatures from different traditions to propose a new analytical framework, capable of guiding both academic reflection and administrative practice.

Contributions were gathered from four main areas: public administration and the historical sociology of the state, with an emphasis on the state capacity literature; information systems and digital government, including studies on platformization and vendor dependency; algorithmic ethics and governance, focused on trustworthy artificial intelligence and accountability; and studies on bureaucratic discretion, both in the classical tradition and in the more recent literature on digital bureaucracy.

The construction of the model proceeded through four successive stages. The first stage identified, from the literature reviewed, the main limitations local governments face in adopting digital technologies, with particular attention to constraints of personnel, budget, data, and internal coordination. The second stage cross-referenced the classical dimensions of state capacity, such as autonomy, coordination, and social embeddedness, with the specific requirements imposed by the life cycle of artificial intelligence systems, from conception to eventual discontinuation. The third stage formulated, from this cross-referencing, six specific capacities, each corresponding to a coherent set of resources, routines, and institutional relationships necessary for responsible governance. The fourth stage examined the interdependencies among these capacities and the risks of cascading failures, where weakness in one dimension compromises the functioning of the others.

References published up to 2024 were used as a priority, in keeping with the editorial deadline of the volume to which this article was originally submitted. It should be noted that the model does not aim to empirically measure the maturity of specific municipalities at this stage of the research; its purpose is to offer consistent conceptual categories that can be operationalized into indicators and data-collection instruments by subsequent research, including case studies, survey research, and comparative analyses among municipalities of different sizes and regions.

MODEL OF STATE CAPACITIES FOR THE RESPONSIBLE GOVERNANCE OF LOCAL AI

The proposed model comprises six interdependent capacities: strategic, regulatory, technical, informational, organizational, and participatory. Each corresponds to a specific set of resources, routines, and institutional relationships, but none of them, in isolation, produces responsible governance. The absence or weakness of one dimension limits the functioning of the others: technical competence without strategic direction tends to produce automation without clear priorities; well-drafted regulation, but without reliable data and qualified personnel to apply it, tends to remain merely formal; and social participation without a minimum of explainability does not guarantee real control over the systems, because the citizens consulted have no way of assessing what is being decided in their name.

Strategic capacity

Strategic capacity refers to the administration's ability to define public problems, set priorities, and recognize limits on the use of technology. Local governments need to distinguish situations in which artificial intelligence offers a real benefit from situations in which simpler process improvements, such as integrating forms or revising a service workflow, would be sufficient and cheaper. The starting point of any project should be the identified public need, not the availability of a tool on the market or pressure from an external funding call that ties resources to the adoption of a particular technology.

Municipal digital transformation plans need to establish clear objectives, project selection criteria, and monitoring indicators. Artificial intelligence projects must be aligned with the broader municipal planning and with existing sectoral policies, preventing automation initiatives from becoming islands disconnected from the government's strategy. Strategic definition also needs to explicitly indicate which areas are considered high risk and which types of decision should not be fully delegated to automated systems, such as decisions that directly affect freedom, health, or access to essential subsistence benefits.

Strategic capacity also includes a long-term vision dimension that tends to be neglected in contexts of frequent political turnover. Artificial intelligence systems create costs and dependencies that persist well beyond the duration of an elected term. The continuity of these initiatives requires adequate documentation, multi-year budget forecasting, and the preservation of accumulated institutional knowledge, so that changes in political leadership do not eliminate essential controls or interrupt projects that were producing positive results, simply for reasons of administrative discontinuity.

Regulatory capacity

At the local level, regulatory capacity manifests itself in internal rules, contracts, operating protocols, and audit procedures. Even when general legislation on data protection or artificial intelligence is set at the national level, the municipality remains responsible for defining how specific systems will be acquired, deployed, and used within its territory. These internal rules need to be proportional to the risk of each application, avoiding both a complete absence of control and the creation of unnecessary bureaucratic requirements for low-impact systems.

Administrative contracts involving artificial intelligence must explicitly set out the system's purpose, expected performance criteria, transparency and security requirements, and the rules governing ownership and portability of the data involved. They also need to define contractual liability for failures, the vendor's obligation to report relevant changes in how the system functions, and guaranteed access to the evidence needed for external audits. Generic commercial confidentiality clauses, common in technology contracts, cannot serve as an absolute barrier to public control over decisions that directly affect citizens.

Algorithmic impact assessment can be used as a prior instrument before deploying applications considered sensitive. This type of assessment should identify the populations potentially affected, the data to be used, the risks involved, the planned mitigation measures, and the appeal channels available to anyone who feels harmed by an automated decision. Systems classified as low risk may follow a simplified assessment process, provided that a minimum level of documentation is preserved that allows for later reconstruction of how and why that tool was adopted.

Technical capacity

Technical capacity requires the administration to have civil servants or teams capable of understanding, at least in broad terms, the architecture of the systems used, the origin and limitations of the data involved, and the metrics used to assess their performance. It is not necessary for every worker involved to be a programmer or a data scientist, but the responsible managers need sufficient technical literacy to question vendors in a qualified manner and to critically interpret the results a system presents.

Interdisciplinary teams tend to be better suited than exclusively technological units for running artificial intelligence projects in the public sector. Professionals from information technology, legal counsel, public policy design, personal data protection, and the substantive areas directly affected should participate jointly in relevant decisions. This diversity of perspectives reduces blind spots that a technically homogeneous team tends not to notice, and brings the system's design closer to the concrete context in which it will be applied.

Training and retaining qualified professionals are structural challenges for most Brazilian municipalities, especially small and medium-sized ones, which compete at a disadvantage with the private sector for this kind of talent. Cooperation with universities, intermunicipal consortia, and other levels of government can expand the resources available. It is important, however, that these partnerships produce effective knowledge transfer into the municipal administration; permanent dependence on external consultants, even technically competent ones, preserves the very institutional fragility the model seeks to reduce.

Informational capacity

Informational capacity covers the quality, integration, protection, and governance of the data used by the administration. Before applying any artificial intelligence system, the municipality needs to know its own databases, identify their gaps, and understand the criteria that guided their historical construction. Low-quality data are not automatically corrected by the sophistication of a statistical model; on the contrary, more sophisticated models tend to amplify errors present in low-quality data, producing results that appear reliable precisely because they are presented with numerical precision.

Interoperability among systems requires shared technical and semantic standards. Two departments within the same city government may use the same term for different concepts, or different terms for the same concept, which hampers any attempt at database integration. Without prior harmonization of these standards, integration attempts tend to produce inconsistencies that only become visible once they have already caused some administrative harm. Data catalogs, metadata dictionaries, and common identifiers across systems are relatively simple instruments, but ones of great practical importance for reducing this kind of problem.

The privacy of personal data needs to be built into system design from the outset, rather than treated as a later legal-compliance check. Data collection carried out to enable an artificial intelligence system needs to be necessary and proportionate to its stated purpose. Access to that data should be controlled through clear permission profiles, and sensitive data cannot be reused for purposes incompatible with the one that justified its original collection, unless there is a specific legal basis and a prior assessment of the risks involved.

Organizational capacity

Organizational capacity refers to internal coordination, work routines, and the institutional culture that sustain the day-to-day operation of artificial intelligence systems. Deploying any such system alters tasks and responsibilities within the administration. When this change is not openly discussed with the teams involved, it tends to generate silent resistance, improper use of the tool, or, at the opposite extreme, excessive trust in automated recommendations that should be questioned more carefully.

Every significant artificial intelligence project needs clearly designated leads in the business, technical, and internal control areas. Governance of these projects should not be confined to the information technology department, even though it retains a central role in operations. The substantive departments, which know the practical consequences of each automated decision on the lives of the people served, need to closely monitor performance indicators and recorded incidents.

Institutional learning depends on systematic records, periodic evaluation, and a willingness to correct course when necessary. Identified failures should produce concrete changes in processes and training programs, not merely administrative justifications for maintaining the status quo. A punitive organizational culture tends to encourage the concealment of incidents, because workers fear reprisals for reporting problems; a responsible organizational culture, by contrast, clearly distinguishes between an isolated error, repeated negligence, and a legitimate learning process, treating each situation proportionately.

Participatory capacity

Participatory capacity makes it possible to include citizens, social movements, and civil society organizations in defining and evaluating artificial intelligence systems. Applications that directly affect access to social benefits, the intensity of inspection, or the quality of public services need to take into account the concrete experiences of the groups affected by these decisions. Well-conducted public consultations can reveal practical barriers that would hardly emerge from the analysis of administrative data alone.

Explainability aimed at the general public requires accessible language, different from that used in internal technical reports. Detailed technical documentation is important for audit purposes, but it does not replace public communication directed at the ordinary citizen. The municipality needs to explain, in understandable terms, what the system's purpose is, what general criteria guide its decisions, what risks have been identified, and what channels exist for contesting a result considered unfair.

The right of appeal needs to be operational, not merely provided for formally in some little-known regulation. Citizens affected by an automated decision must be able to request its review by a person qualified to re-examine the case, without this process requiring specialized technical knowledge on the part of the person appealing. Systematic records of these appeals, including their outcomes, help the administration identify recurring error patterns and continuously improve the system, turning individual challenges into a source of collective learning.

Interdependence and cascading failures

The six capacities described form, in practice, a single system, not a set of independent items on a checklist. A procurement process may be legally sound and, at the same time, technically fragile; data may be well integrated technically and still reproduce social exclusions that were never discussed with the affected groups; teams may be technically qualified and lack the institutional autonomy to halt a problematic project. The governance maturity of a municipality therefore depends on the relative balance among these six dimensions, not on the isolated performance of any one of them.

Cascading failures occur when weakness in one capacity feeds problems in another. Low-quality data produce inadequate administrative outcomes; low technical competence prevents that inadequacy from being identified in time; opaquely drafted contracts hinder correcting the problem even after it is identified; and the absence of an effective appeal channel turns an isolated error into lasting harm for whoever was affected by it. Systemic analysis of these causal chains makes it possible to identify, in each concrete case, the most critical points for priority intervention.

The model suggests that investments in digital transformation should be prioritized according to the weakest link in the chain, not according to whatever technological novelty is available on the market. In many municipalities, acquiring a new technology is, objectively, less urgent than improving the quality of existing data, reviewing current contracts, or investing in staff training. Digital maturity, from this perspective, is not a race to see how many systems have been deployed, but a gradual process of institutional consolidation that needs to advance in a balanced way across the six capacities.

THE LIFE CYCLE OF ARTIFICIAL INTELLIGENCE IN LOCAL GOVERNMENT

Responsible governance needs to accompany every phase of an artificial intelligence system's life cycle: conception, procurement, development, deployment, operation, and, eventually, discontinuation. In the conception phase, the public problem to be addressed is clearly defined and solution alternatives, technological or otherwise, are evaluated. In the procurement phase, the technical, contractual, and governance requirements that will guide acquisition are established. In the development phase, the available data and the behavior of the proposed models are tested before any use in production. In the deployment phase, the people who will operate the system and the procedures that will guide its day-to-day functioning are prepared.

During the operation phase, the system needs to be monitored continuously. Changes in input data over time, a phenomenon known in the technical literature as data drift, can progressively reduce a model's performance without this being immediately noticed by the responsible managers. Recorded incidents, formal complaints, and automated decisions that were overturned on appeal are important indicators that something may be functioning inadequately. Evaluating a system in operation should not be limited to its technical availability, that is, to the mere fact that it is running without interruption.

Discontinuing a system is a legitimate, and sometimes necessary, part of its governance. Systems that have lost their original purpose, that have become technically obsolete, or that produce risks disproportionate to the benefits they generate should be shut down, even if this means publicly acknowledging that a previous investment did not produce the expected results. In this process, the administration needs to preserve the relevant data, the historical records of decisions made, and the continuity of the public service that depended on that system, preventing the discontinuation of a tool from translating into an interruption of service to the population.

Risk assessment and proportionality

Not all artificial intelligence applications in the public sector present the same level of risk. A virtual assistant that provides general information about service hours is qualitatively different from a system that denies access to a social benefit or that prioritizes certain addresses for tax inspection. The intensity of governance controls should track the potential impact of the decision, the scale of people affected, and the concrete possibility of irreversible harm.

Risk-assessment criteria may consider, among other factors, the fundamental rights potentially affected, the social vulnerability of the populations involved, the degree of decision-making autonomy assigned to the system, the quality of the data used, and the reversibility of the decision made. Yeung (2018) describes algorithmic regulation precisely as a spectrum of regulatory intensity that should vary according to these factors, rather than treating all automation with the same yardstick. Applications classified as high risk require independent prior assessment, rigorous testing before deployment, and reinforced human oversight throughout operation.

Proportionality in governance avoids two equally problematic extremes: a complete absence of control, which exposes the administration and citizens to avoidable risks, and indiscriminate bureaucracy, which treats low-risk systems with the same rigor demanded for high-impact ones, consuming scarce resources without a corresponding gain in safety. Local governments with limited resources necessarily need to concentrate their control efforts where the identified risks are actually greatest.

Public procurement and vendor management

Procurement is a decisive moment in the trajectory of an artificial intelligence system in the public sector, because many of the future limitations on governance are already set at this stage, even though their effects only appear years later. Tenders and terms of reference should precisely describe the public problem to be solved, the data involved, and the performance, security, and transparency requirements. Procurement driven by brand name or by generic promises of efficiency tends to reduce competition among vendors and to hinder any objective technical evaluation of the proposals received.

Proofs of concept can be useful for reducing uncertainty about a solution before a long-term contract, provided they are conducted with clear evaluation criteria and do not, in practice, create premature dependency on the vendor being tested. The municipality needs to retain control over the data used during these experiments and prevent informal pilots from being automatically converted into long-term contracts without due bidding process and without a rigorous evaluation of the results obtained.

Contract management continues after the contract is signed, and this phase tends to receive less institutional attention than the acquisition phase. Agreed service levels, recorded incidents, updates carried out, and any changes in how the model functions need to be systematically monitored by the administration. Outsourcing the operation of a system does not, under any circumstances, transfer public responsibility for the decisions that result from it; the vendor may be held civilly liable for contractual failures, but political and administrative responsibility remains with the public authority.

Auditing, monitoring, and accountability

Audits of artificial intelligence systems can examine the quality of the data used, the system's technical performance, information security, the existence of discriminatory biases, and compliance with applicable legal and contractual rules. The frequency of these audits should reflect the application's risk level and how often the system is updated. Systems that learn continuously from new data, or that receive frequent updates from the vendor, need mechanisms for ongoing control, not just a one-off audit carried out at the time of initial approval.

The indicators used to monitor a system should include, among others, overall accuracy rates, error rates disaggregated by social group and territory, the number of decisions successfully contested, average response time to appeals, recorded incidents, and the satisfaction level of the users served. Metrics aggregated at the municipal level can hide impacts concentrated in specific neighborhoods or populations, which is why disaggregating indicators is as important as having them in the first place.

Accountability for automated systems requires consistent documentation over time. Records of system versions, relevant decisions made about how the system functions, and human interventions carried out during its operation allow the entire process leading to a given result to be reconstructed later. Transparency about these records must respect legitimate requirements of information security and personal data protection, but these restrictions cannot be used to turn a vendor's trade secret into an absolute barrier against public oversight.

Participation, contestation, and social control

Municipal councils, public hearings, and specific consultations can accompany the deployment of artificial intelligence applications considered relevant to the community. This participation needs to take place before the system is deployed and continue throughout the entire period of its evaluation, not only after the decision to adopt a particular tool has already been made and formalized. Informing the population only after the decision has been settled significantly reduces the legitimacy of the whole process, even if the communication itself is technically accurate.

The contestation channels available to citizens need to be accessible, free of charge, and capable of generating a response within a reasonable time. People affected by an automated decision should not be required to understand the system's technical details in order to request its review; the burden of technically justifying the decision made falls on the public authority, not on the citizen who feels harmed by it. This reversal of the burden of explanation, implicitly present in much of the literature on administrative due process, takes on new relevance in the face of systems whose technical complexity can easily intimidate someone seeking to challenge a decision.

Internal control bodies, municipal ombudsman's offices, and data protection authorities can collaborate in complementary ways within this arrangement of social control. Distributing these functions across different bodies, rather than concentrating them within a single department responsible for both operating and overseeing the system, creates institutional checks and balances and reduces the risk of excessive concentration of decision-making power over technologies that directly affect people's lives.

APPLYING THE MODEL AND MATURITY INDICATORS

The model developed here can support municipal diagnostics of the current stage of artificial intelligence governance. At an initial stage, projects tend to be isolated, dependent on specific vendors, and without any coordination among departments. At an intermediate stage, internal rules and technical teams already exist, but their integration across different areas of the administration remains uneven. At an advanced stage, the governance of artificial intelligence systems is embedded in strategic planning and in the municipality's regular mechanisms of internal and external control.

Strategic indicators may include the existence of a formal digital transformation plan, explicit project-selection criteria, and demonstrable alignment with existing sectoral policies. Regulatory indicators involve the existence of appropriate contract models, algorithmic impact assessments actually carried out, and formally established appeal channels. Technical indicators examine the composition and autonomy of the teams responsible for operating the systems. Informational indicators assess the quality, interoperability, and level of protection of the databases used.

Organizational indicators include the existence of formal coordination mechanisms among areas, regular training programs, and structured processes of institutional learning drawn from recorded incidents. Participatory indicators observe the frequency and quality of consultations carried out, the level of proactive transparency maintained by the administration, and the effectiveness of the contestation channels available to the population. It is important to stress that this evaluation should guide the continuous improvement of municipal governance, not merely serve to produce a simplified ranking among city governments, which would hardly capture the real differences in context among them.

Municipalities interested in advancing this diagnosis can begin with a complete inventory of the automated systems already in use and the databases that feed them. Knowing what is already in operation, even if informally or poorly documented, is a necessary condition for any serious attempt to govern these technologies. From this inventory, the administration can classify the risks involved, define clear institutional owners for each system, and prioritize corrections for the most urgent cases.

Intermunicipal consortia and networks can share specialists and common technical standards among different city governments. This cooperation tends to reduce individual costs and strengthen each municipality's bargaining power with vendors, provided each party's responsibilities remain clearly defined throughout the cooperative arrangement.

Portfolio planning and project prioritization

Local governments tend to begin their digital transformation journey through scattered projects, driven by immediate administrative demands or by one-off external funding opportunities. This trajectory, although understandable given the budget constraints faced by most municipalities, can generate overlap among similar systems and the adoption of solutions that bear little relation to the public priorities actually defined by the administration. Strategic capacity needs to turn this scattered set of initiatives into a coherent portfolio, one that allows costs, risks, contractual dependencies, and expected benefits to be compared across different projects competing for scarce resources.

Project prioritization should consider public materiality, that is, its actual relevance to the collective interest. Problems with high social impact, a large volume of service demand, or persistent difficulty coordinating across areas may justify deeper analytical investment. It is important, however, that the use of artificial intelligence not be presumed to be the automatic solution to these problems. In many concrete cases, simply revising administrative forms, integrating already existing databases, or simplifying bureaucratic procedures offers a greater benefit and a lower risk than adopting a sophisticated automated system.

The municipal portfolio should also record discontinued projects, experimental pilots, and applications acquired in isolation by specific departments, without going through any centralized evaluation process. The absence of a unified inventory prevents the administration from fully knowing its own contractual commitments and the data it actually processes. A municipal register of automated systems is a basic instrument, low-cost to implement, for any digital governance strategy, and one that can also directly support public transparency initiatives aimed at the population.

Portfolio analysis needs to be periodic, not carried out only once at the start of a digital transformation project. Regulatory, technological, and budgetary changes alter, over time, the relative merit of already-deployed solutions. Projects that offered clear value at their outset may become redundant or incompatible with new technical standards subsequently adopted by the administration. Periodic portfolio review allows scarce resources to be reallocated and prevents the maintenance of obsolete systems merely because of investment already made in the past, a trap known in the management literature as sunk cost.

Federative cooperation and intermunicipal consortia

The inequality in institutional capacity among Brazilian municipalities, widely documented in the federalism and public policy literature, points to the need for specific cooperative arrangements for artificial intelligence governance. Intermunicipal consortia can share specialists, information technology infrastructure, common data standards, and independent audit services. This strategy tends to reduce individual costs and to expand each municipality's bargaining power with major technology vendors, a benefit especially relevant for small and medium-sized municipalities, which individually have little contractual leverage.

Intermunicipal cooperation, however, needs to preserve clear governance mechanisms. Sharing the same technology platform among different city governments does not eliminate each party's individual responsibility for the administrative decisions made with the support of that platform. Cooperation contracts and protocols must precisely define ownership of the data involved, access rules among participating municipalities, information security responsibilities, shared financing, and the procedures for any party to exit the cooperative arrangement. A consortium without these clearly defined rules risks merely shifting the problem of dependency on a private vendor to an equivalent dependency on a poorly governed regional structure.

States and the federal government can also support municipalities through common technical standards, training programs, and shared computing and storage infrastructure. Standardization promoted by higher levels of government facilitates interoperability among different municipal systems, but it needs to allow room for adaptation to local realities. Excessively uniform models, designed from the standpoint of major capital cities, may ignore relevant territorial and institutional differences present in smaller municipalities or regions with very different characteristics. Federative cooperation tends to be more productive when it combines a common technical reference with real implementation autonomy for each municipality.

Public universities and research institutes can contribute technical training, independent system evaluation, and applied research directed at the concrete needs of local administration. These partnerships need to be formalized in a way that preserves both the academic independence of the institutions involved and the practical usefulness of the results produced for municipal management. Short-term projects, common in this type of cooperation, should produce adequate documentation and effective knowledge transfer, so that the results obtained remain available to the administration even after the partnership formally ends.

People management, work, and organizational change

The adoption of automated systems concretely alters the day-to-day work of public servants. Tasks may be redistributed among different positions, decision criteria may change without clear notice to those who apply them, and new forms of monitoring the worker's own performance may emerge as a side effect of deploying a system. Jarrahi (2018) argues that the relationship between artificial intelligence and human labor is best understood as complementary rather than substitutive in most organizational contexts, but this complementarity is only realized when the transformation is planned with attention to its impact on the people involved, rather than treated merely as a software-deployment project.

Front-line workers hold contextual knowledge about how public services actually function that is rarely fully captured in available administrative data. Ignoring this tacit knowledge reduces the final quality of the deployed system and tends to increase the workers' own resistance to the new tool. The active participation of these professionals in the design, testing, and ongoing evaluation of the system makes it possible to identify exceptions, ambiguities, and effects not anticipated by the technical team responsible for development.

The training offered to civil servants needs to go beyond simply teaching how to use the system's interface. Managers and direct operators need to understand the tool's purpose, its known limitations, the indicators used to assess its performance, and the procedures available for appeal in case of error. They also need to be able to recognize concrete situations in which the automated recommendation is clearly not appropriate to the case at hand. This training needs to be updated whenever there are relevant changes in how the system functions, rather than offered only once, at the time of its initial deployment.

Organizational change management needs to communicate honestly both the objectives and the real limits of each automation project. Promises of broad replacement of human tasks or of automatic and immediate efficiency gains tend to generate insecurity and distrust among workers, especially when those promises fail to materialize in practice. The administration should clearly explain how responsibilities will be redistributed under the new tool and what institutional mechanisms will protect the exercise of professional judgment by the workers involved. This internally built legitimacy is a necessary, though not sufficient, condition for responsible use of the adopted technology.

Proactive transparency and public communication

Transparency about the automated systems used by the administration needs to be carefully planned, not treated as a bureaucratic obligation fulfilled only formally. Simply publishing source code or detailed technical documents may be insufficient to ensure real social understanding of how a system works, while a generic announcement that the city government "uses artificial intelligence" in a given area, without further detail, does not allow for any effective control by the population. Information that is genuinely useful for social oversight includes the system's specific purpose, the agency responsible for its operation, the type of data used, the populations potentially affected, and the concrete channels available for contestation.

The transparency required must respect legitimate requirements of security, privacy, and intellectual property, but these restrictions cannot completely eliminate the accountability owed to the population. In practice, a municipality can disclose the general criteria used by a system and the results of its periodic evaluations without exposing specific personal data or technical vulnerabilities exploitable by malicious third parties. Contracts with vendors must ensure, from their initial drafting, that the trade secrecy invoked by the company does not prevent an adequate explanation of the public decisions supported by its technology.

Periodic performance reports can present operating indicators, recorded incidents, complaints received, and corrective measures actually adopted. Deetz (2001), discussing communication between organizations and their stakeholders, argues that genuine communicative processes require real space for dissenting voices, not just the one-way transmission of information favorable to the issuing institution; this perspective applies directly to public communication about automated systems, which should also include identified problems, not only positive results. Regularly publishing this type of report strengthens institutional learning and allows municipal councils, the local press, and society at large to follow the results obtained over time. The deliberate omission of failures tends to undermine public trust even more severely when these problems eventually become known through other sources, such as journalistic reports or complaints from affected citizens.

Communication on these topics needs to use language accessible to the ordinary citizen, without unnecessary technical jargon. Specialized terms, when unavoidable, should be explained in plain language, and concrete examples can illustrate, in a didactic way, how the system actually participates in the decision-making process in question. This meaningful, rather than merely formal, transparency allows people to understand the real role technology plays in their lives and to clearly identify the public authority responsible for each decision that affects them.

Financial sustainability and life-cycle cost

Financial sustainability is also a relevant component of state capacity for governing digital technologies. Artificial intelligence projects involve costs of integration with legacy systems, data cleaning and processing, computing power, software licensing, information security, staff training, and periodic auditing. An evaluation based exclusively on the initial acquisition price tends to significantly underestimate the future financial commitments assumed by the administration, potentially producing a subtle form of budgetary dependency in the years that follow.

The total life-cycle cost of a system needs to be compared both against the expected public benefit and against available alternatives, including non-technological ones. Simpler solutions may be objectively more appropriate when they offer greater transparency and require maintenance compatible with the real capacity of the available municipal team. Automatically choosing the technically most sophisticated model is not necessarily the most rational decision from a public management standpoint, especially when the marginal performance gains obtained require infrastructure disproportionate to the municipality's size and real needs.

Long-term budget forecasting needs to account, from the outset of the project, for both update costs and the eventual costs of discontinuing the system. Systems that do not receive sufficient resources for their ongoing maintenance tend to degrade technically over time and may nevertheless remain in use out of sheer administrative inertia, even after they stop producing the expected results. The administration should also set aside specific resources for potential data migration and vendor replacement, thereby reducing the risk of technological lock-in that is difficult to reverse in the future.

Shared purchasing among municipalities and the adoption of open technical standards can reduce individual acquisition costs, but they also need to be evaluated against each participating administration's real operating capacity. The savings obtained at the initial acquisition stage lose practical value when the chosen solution requires permanent external technical support that the municipal team cannot replace. The financial sustainability of an artificial intelligence project ultimately depends on a careful balance among acquisition cost, available internal technical knowledge, and the quality of governance established over the system.

Theoretical propositions and an agenda for empirical validation

The model developed in this article makes it possible to formulate specific propositions to guide future empirical research. The first proposition holds that the adoption of artificial intelligence will produce greater public value when there is effective alignment between a municipality's strategic capacity and its informational capacity. Without a clear definition of the public problem to be solved and without administrative data of adequate quality, a system's good technical performance is unlikely to translate into a real improvement in public policy, even if the project's technical reports present technically satisfactory indicators.

The second proposition holds that regulatory capacity moderates the risks associated with both vendor dependency and the opacity of automated decisions. Municipalities that develop more robust contracts, systematically carry out algorithmic impact assessments, and maintain effectively operational appeal mechanisms tend to preserve greater institutional autonomy over time, even in the face of vendors that are technically dominant in their market. The third proposition indicates that technical capacity and organizational capacity have a complementary effect on each other: technically qualified teams produce little practical result when they lack adequate access to necessary information, recognized institutional authority, or effective integration with the substantive areas responsible for the public policy in question.

The fourth proposition directly relates social participation and institutional legitimacy. Systems that affect fundamental rights or the distribution of essential public services tend to achieve greater social acceptance when the populations potentially affected actively participate in the system's design, receive clear explanations of how it works, and have effective appeal channels available when they feel harmed. It is important to note that social participation does not eliminate legitimate conflicts of interest among different groups affected by the same public policy, but it makes the criteria used and the institutional responsibilities involved considerably more visible to everyone concerned.

These four propositions can be examined empirically through in-depth case studies, survey research applied to municipal managers, and comparative analyses across different local administrations. Longitudinal research would make it possible to track the evolution of the six capacities described before and after the deployment of specific artificial intelligence systems, offering more robust evidence about causal relationships than purely cross-sectional studies. Comparative studies conducted in municipalities of different population sizes and fiscal capacities could identify functionally equivalent institutional combinations, that is, distinct arrangements that produce similar governance outcomes, as well as the practical limits of the different intermunicipal cooperation strategies discussed earlier. The empirical agenda arising from this model should also include analysis of distributive outcomes across different social groups, costs actually incurred, and experiences reported both by public servants and by citizens directly affected by the systems evaluated.

Building quantitative indicators from this model requires considerable methodological caution. Simple measures of the presence or absence of a given formal instrument can easily conflate bureaucratic formalization with actual institutional capacity. A formally published municipal policy does not guarantee its consistent practical application; a team formally assigned to a given function may lack the resources needed to perform it; and a formally existing appeal channel may, in practice, be inaccessible to most of the citizens who would need it. Empirical validation of the proposed model should therefore combine documentary analysis, in-depth qualitative interviews, direct observation of administrative processes, and analysis of performance data actually produced by the systems evaluated.

CONCLUDING REMARKS

This article developed a theoretical model of state capacities aimed at the responsible governance of artificial intelligence in local governments. The six proposed capacities, strategic, regulatory, technical, informational, organizational, and participatory, help explain why the adoption of the same technology produces such different administrative outcomes across different municipalities, and why the sheer volume of investment in digital tools is not, by itself, a good indicator of governance quality.

The main conclusion of this work is that the digital maturity of a local government does not correspond to the number of automated systems it uses. It lies, rather, in the institutional capacity to define clear purposes for each system adopted, maintain effective control over its functioning over time, assess its real impacts on the population served, and, when necessary, discontinue uses that prove inadequate. Responsible governance, as discussed throughout this article, combines administrative efficiency with respect for rights, transparency about the criteria used, and effective generation of public value, without automatically privileging any one of these dimensions over the others.

The model further shows that the institutional weaknesses discussed in this article are fundamentally interdependent. Consistent improvement in governance requires a systemic diagnosis, capable of identifying where the weakest links lie in each specific municipal context, and sustained institutional investment over time, not only at moments of greater political visibility. Purchasing advanced technology without simultaneously developing the institutional capacities needed to govern it tends to widen, rather than reduce, the municipality's structural dependence on external vendors.

As a central limitation, it must be acknowledged that this study is theoretical and does not present direct empirical validation of the relationships proposed. Future research can operationalize the indicators discussed here into concrete data-collection instruments, systematically compare different Brazilian municipalities, and track specific artificial intelligence projects over several years of operation. In-depth case studies could examine, in more detail than was possible in this article, real administrative contracts, databases actually used, social participation processes conducted, and incidents concretely recorded in different municipal contexts.

For Brazilian public administration, the central practical implication of this article is that artificial intelligence should be treated as a matter of state policy subject to continuous governance, not as a set of one-off technical projects managed in isolation by each interested department. Institutional control over these systems needs to exist before their deployment, throughout their entire operation, and also afterward, when the time comes to assess whether a given system should continue in use, be redesigned, or simply be discontinued. Only this kind of continuous control allows the digital transformation of local governments to fulfill its promise of greater efficiency without compromising rights, widening existing inequalities, or deepening municipalities' dependence on external actors who should instead serve the public interest.

REFERENCES

Barth, T. J., & Arnold, E. (1999). Artificial intelligence and administrative discretion: Implications for public administration. American Review of Public Administration, 29(4), 332-351. https://doi.org/10.1177/02750749922064463

Bovens, M., & Zouridis, S. (2002). From street-level to system-level bureaucracies: How information and communication technology is transforming administrative discretion and constitutional control. Public Administration Review, 62(2), 174-184. https://doi.org/10.1111/0033-3352.00168

Bozeman, B. (2007). Public values and public interest: Counterbalancing economic individualism. Georgetown University Press.

Bullock, J. B. (2019). Artificial intelligence, discretion, and bureaucracy. American Review of Public Administration, 49(7), 751-761. https://doi.org/10.1177/0275074019856123

Busch, P. A., & Henriksen, H. Z. (2018). Digital discretion: A systematic literature review of ICT and street-level discretion. Information Polity, 23(1), 3-28. https://doi.org/10.3233/IP-170050

Centeno, M. A., Kohli, A., Yashar, D. J., & Mistree, D. (2017). States in the developing world. Cambridge University Press.

Deetz, S. (2001). Conceptualizing organization-stakeholder communication. Management Communication Quarterly, 15(1), 51-60. https://doi.org/10.1177/0893318901151003

Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2006). New public management is dead - Long live digital-era governance. Journal of Public Administration Research and Theory, 16(3), 467-494. https://doi.org/10.1093/jopart/mui057

Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.

Evans, P. (1995). Embedded autonomy: States and industrial transformation. Princeton University Press.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People - An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689-707. https://doi.org/10.1007/s11023-018-9482-5

Fukuyama, F. (2004). State-building: Governance and world order in the 21st century. Cornell University Press.

Fukuyama, F. (2013). What is governance? Governance, 26(3), 347-368. https://doi.org/10.1111/gove.12035

Gil-Garcia, J. R., Dawes, S. S., & Pardo, T. A. (2018). Digital government and public management research: Finding the crossroads. Public Management Review, 20(5), 633-646. https://doi.org/10.1080/14719037.2017.1327181

Janssen, M., & van den Hoven, J. (2015). Big and open linked data (BOLD) in government: A challenge to transparency and privacy? Government Information Quarterly, 32(4), 363-368. https://doi.org/10.1016/j.giq.2015.11.007

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. Business Horizons, 61(4), 577-586. https://doi.org/10.1016/j.bushor.2018.03.007

Kitchin, R. (2014). The data revolution: Big data, open data, data infrastructures and their consequences. Sage.

Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. Telecommunications Policy, 44(6), 101976. https://doi.org/10.1016/j.telpol.2020.101976

Lipsky, M. (2010). Street-level bureaucracy: Dilemmas of the individual in public services (30th anniversary ed.). Russell Sage Foundation.

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

Morozov, E. (2013). To save everything, click here: The folly of technological solutionism. PublicAffairs.

Organisation for Economic Co-operation and Development. (2014). Recommendation of the Council on Digital Government Strategies. OECD Publishing.

Organisation for Economic Co-operation and Development. (2020). The OECD Digital Government Policy Framework: Six dimensions of a digital government. OECD Digital Government Studies. https://doi.org/10.1787/f64fed2a-en

O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown.

Peeters, R. (2020). The agency of algorithms: Understanding human-algorithm interaction in administrative decision-making. Information Polity, 25(4), 507-522. https://doi.org/10.3233/IP-200253

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 33-44. https://doi.org/10.1145/3351095.3372842

Schuppan, T. (2009). E-government in developing countries: Experiences from sub-Saharan Africa. Government Information Quarterly, 26(1), 118-127. https://doi.org/10.1016/j.giq.2008.01.006

Skocpol, T. (1985). Bringing the state back in: Strategies of analysis in current research. In P. Evans, D. Rueschemeyer, & T. Skocpol (Eds.), Bringing the state back in (pp. 3-43). Cambridge University Press.

Torraco, R. J. (2016). Writing integrative literature reviews: Using the past and present to explore the future. Human Resource Development Review, 15(4), 404-428. https://doi.org/10.1177/1534484316671606

Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector - Applications and challenges. International Journal of Public Administration, 42(7), 596-615. https://doi.org/10.1080/01900692.2018.1498103

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. Regulation & Governance, 12(4), 505-523. https://doi.org/10.1111/rego.12158

Zuiderwijk, A., Chen, Y.-C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review. Government Information Quarterly, 38(3), 101577. https://doi.org/10.1016/j.giq.2021.101577

AUTHOR CONTRIBUTION

Robson Barbosa: conceptualization, methodology, investigation, formal analysis, original manuscript writing, and review and editing.

EDITORIAL DECLARATIONS

Funding: the study received no external funding.

Conflict of interest: the author declares no conflict of interest.

Use of artificial intelligence: not declared by the author.