

ARTIGO

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

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

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

Robson Barbosa¹

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

SUBMETIDO	ACEITO	PUBLICADO	DOI
20/03/2025	30/10/2025	05/12/2025	em processamento

RESUMO

Governos locais enfrentam pressão crescente para adotar inteligência artificial na formulação de políticas e na prestação de serviços, mas costumam operar com limitações de pessoal, infraestrutura, dados, coordenação e controle institucional. Este artigo desenvolve um modelo teórico de capacidades estatais voltado à governança responsável da inteligência artificial em governos locais. A abordagem é teórico-integrativa e articula a literatura sobre capacidade estatal, governo digital, governança de dados, valor público e inteligência artificial confiável. O resultado é um conjunto de seis capacidades interdependentes: estratégica, regulatória, técnica, informacional, organizacional e participativa. O modelo mostra que fragilidades em qualquer uma dessas dimensões tendem a produzir dependência de fornecedores, automação fragmentada, decisões opacas e ampliação de desigualdades. A conclusão central é que a maturidade digital não se mede pelo número de ferramentas adotadas, mas pela capacidade pública de selecionar, supervisionar, avaliar, contestar e, quando necessário,

descontinuar sistemas, equilibrando eficiência, direitos, transparência, responsabilização e valor público.

Palavras-chave: capacidades estatais, governo local, transformação digital, inteligência artificial, governança pública

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

INTRODUÇÃO

A incorporação de sistemas baseados em inteligência artificial na administração pública mudou os meios pelos quais governos diagnosticam problemas, organizam serviços e se relacionam com os cidadãos. Ferramentas de classificação, previsão, triagem e atendimento automatizado passaram a integrar processos que até pouco tempo dependiam quase exclusivamente da análise de servidores. No nível local, essas tecnologias já aparecem em saúde, educação, assistência social, mobilidade, fiscalização, tributação e planejamento urbano, muitas vezes chegando antes de qualquer norma municipal específica sobre seu uso.

Os governos municipais ocupam uma posição peculiar nesse processo. São responsáveis pelos serviços mais próximos do cotidiano das pessoas e mantêm contato direto com a população, mas em geral dispõem de menor capacidade fiscal, técnica e informacional do que estados e governo federal. Essa assimetria produz uma tensão entre a pressão por modernização, alimentada tanto por demandas sociais quanto pela oferta comercial de soluções de inteligência artificial, e a capacidade real de selecionar, contratar, supervisionar e avaliar sistemas complexos. Em muitos casos, a adoção de uma ferramenta nasce da iniciativa isolada de uma secretaria, de uma oferta de fornecedor ou de uma necessidade emergencial, sem que exista uma estratégia municipal capaz de orientar decisões futuras.

A digitalização de processos administrativos não é um fenômeno neutro. Sistemas automatizados reorganizam autoridade, alteram fluxos de trabalho e influenciam a distribuição de benefícios e restrições entre os cidadãos. Quando um algoritmo prioriza atendimentos, identifica famílias vulneráveis, calcula riscos de inadimplência ou seleciona áreas para fiscalização, ele passa a incorporar critérios que precisam ser compreensíveis e passíveis de contestação. A opacidade técnica desses sistemas pode transformar escolhas essencialmente administrativas, e portanto políticas, em decisões que parecem objetivas apenas porque vêm de um cálculo, reduzindo a visibilidade de quem, afinal, responde por elas.

Diante desse cenário, este artigo se orienta pela seguinte pergunta: quais capacidades estatais são necessárias para que governos locais governem a

inteligência artificial de maneira responsável e voltada ao interesse público? O objetivo é desenvolver um modelo teórico que articule competências estratégicas, regulatórias, técnicas, informacionais, organizacionais e participativas, e que sirva de referência tanto para análise acadêmica quanto para planejamento institucional. O argumento central é que a maturidade digital depende menos da quantidade de ferramentas implantadas e mais da capacidade da administração de governá-las ao longo de todo o seu ciclo de vida, desde a concepção até a eventual descontinuação.

A relevância do tema decorre de um risco pouco discutido nos debates mais entusiasmados sobre inovação: a possibilidade de que a transformação digital amplie, ao mesmo tempo, eficiência, desigualdades e dependências. Municípios com baixa capacidade institucional correm o risco de se tornar consumidores cativos de soluções proprietárias, incapazes de auditar as próprias decisões administrativas ou de preservar conhecimento técnico quando um contrato termina ou um fornecedor muda de estratégia comercial. O modelo aqui proposto busca oferecer uma estrutura de análise, planejamento e avaliação para essa realidade, sem presumir que toda automação seja desejável nem que toda resistência à automação seja atraso.

O artigo está organizado em sete seções, além desta introdução. A segunda seção apresenta o referencial teórico, articulando a literatura sobre capacidade estatal, governo digital, governança de dados, inteligência artificial confiável e discricionariedade burocrática. A terceira seção descreve o percurso metodológico da revisão integrativa. A quarta seção apresenta o modelo de seis capacidades e discute suas interdependências. A quinta seção detalha o ciclo de vida da inteligência artificial no governo local e os mecanismos de avaliação de risco, contratação e prestação de contas. A sexta seção discute aplicações do modelo, indicadores de maturidade, cooperação federativa, gestão de pessoas, transparência, sustentabilidade financeira e uma agenda de proposições para pesquisa empírica futura. A sétima seção apresenta as considerações finais.

REFERENCIAL TEÓRICO

A fundamentação teórica deste artigo articula cinco corpos de literatura que raramente aparecem integrados em um único modelo: a sociologia histórica e a ciência política sobre capacidade estatal, os estudos de administração pública sobre governo digital e governança de dados, a literatura de ética e governança da inteligência artificial, e os estudos sobre discricionariedade burocrática e street-level bureaucracy. Essa combinação permite tratar a

transformação digital como um processo político e organizacional, e não apenas como uma inovação técnica a ser adotada ou rejeitada.

Capacidades estatais: da tradição clássica à agenda digital

O conceito de capacidade estatal descreve o conjunto de atributos que permitem a um governo formular objetivos, mobilizar recursos e implementar políticas de forma efetiva. A literatura clássica da sociologia histórica trouxe esse conceito de volta ao centro do debate sobre desenvolvimento e governança, enfatizando burocracia profissional, autonomia frente a grupos de interesse, coordenação interna e capacidade de penetração social (Skocpol, 1985; Evans, 1995). Evans (1995) cunhou a expressão "autonomia inserida" para descrever burocracias que combinam competência técnica e insulamento político com vínculos densos com atores sociais e econômicos, uma combinação que permite ao Estado tanto formular políticas coerentes quanto obter informação e cooperação da sociedade.

Fukuyama (2013, 2004) contribuiu para essa tradição ao distinguir capacidade estatal de escopo de atuação estatal e de qualidade da governança. Para o autor, um Estado pode ter escopo amplo, isto é, atuar em muitas áreas, e ainda assim ter baixa capacidade de executar bem o que se propõe a fazer; a força do Estado, nesse sentido, refere-se à sua habilidade de planejar e executar políticas e de fazer cumprir leis de maneira transparente. Essa distinção é útil para pensar governos locais que ampliam rapidamente o escopo de suas ações digitais, adquirindo sistemas de inteligência artificial em diversas áreas, sem que a capacidade de operá-los e supervisioná-los cresça na mesma proporção. Centeno, Kohli, Yashar e Mistree (2017) reforçam esse ponto ao mostrar, em perspectiva comparada, que capacidades estatais variam significativamente entre países e regiões e que essa variação explica boa parte das diferenças na efetividade de políticas públicas, mesmo quando os desenhos institucionais formais são semelhantes.

Essas dimensões clássicas permanecem relevantes para pensar a administração municipal contemporânea, mas a digitalização acrescenta exigências específicas relacionadas a dados, software, segurança da informação, gestão contratual e conhecimento técnico especializado. No ambiente digital, capacidade não se limita a possuir infraestrutura ou licenças de uso. Ela envolve compreender problemas públicos com profundidade suficiente para saber que tipo de solução tecnológica é proporcional a eles, escolher ferramentas adequadas à escala e à

complexidade do problema, estabelecer limites de uso e manter algum grau de controle sobre decisões que passam a ser apoiadas, total ou parcialmente, por sistemas automatizados. Um município pode adquirir uma plataforma avançada de inteligência artificial e, ainda assim, apresentar baixa capacidade estatal quando depende inteiramente do fornecedor para operar, interpretar e corrigir a solução, situação bastante comum em administrações de pequeno e médio porte.

A transformação digital também reconfigura a relação entre autonomia e cooperação, um dos eixos centrais da literatura de capacidade estatal. Governos locais precisam, com frequência, do apoio de outras esferas de governo, universidades, empresas privadas e organizações da sociedade civil para viabilizar projetos de inteligência artificial, mas devem preservar capacidade suficiente para definir finalidades públicas e proteger o interesse coletivo diante de interesses comerciais ou de agendas externas. Parcerias produtivas diferem estruturalmente de relações de dependência: na cooperação, o Estado mantém competência para avaliar propostas, negociar termos e decidir com base em critérios próprios; na dependência, o Estado aceita critérios definidos externamente, sem real possibilidade de contestação ou de substituição do parceiro.

Governo digital, plataformização e dependência de fornecedores

A passagem do governo eletrônico para o governo digital deslocou o foco da simples disponibilização de serviços em canais eletrônicos para a integração de dados, o redesenho de processos e o uso estratégico da informação como recurso de gestão. Dunleavy, Margetts, Bastow e Tinkler (2006) descreveram esse movimento como o surgimento de uma "governança da era digital", caracterizada pela reintegração de funções antes fragmentadas pela nova gestão pública e pela digitalização holística de processos que passam a ser desenhados em função do usuário, e não apenas da estrutura interna do órgão. Gil-Garcia, Dawes e Pardo (2018) observam que a pesquisa em governo digital amadureceu ao reconhecer que tecnologia, organização e política se determinam mutuamente, de modo que o sucesso de uma iniciativa digital não pode ser explicado apenas por sua qualidade técnica.

Mais recentemente, organismos internacionais passaram a tratar a transformação digital do setor público como um problema de governança, e não apenas de infraestrutura. A Organização para a Cooperação e Desenvolvimento Econômico, por meio de sua Recomendação sobre Estratégias de Governo Digital (OECD, 2014) e do Marco de Política de Governo Digital (OECD, 2020), defende que a digitalização deve ser

orientada por design centrado no usuário, abertura por padrão, capacidade de dados e liderança digital, entendida como a habilidade dos governos de dirigir, e não apenas seguir, a transformação tecnológica. Esse enquadramento é particularmente relevante para municípios, que costumam ocupar posição de tomadores de tecnologia diante de um mercado de fornecedores muito mais concentrado e mais bem-informado do que a própria administração compradora.

A integração de dados e processos, contudo, pode ocorrer por meio de plataformas proprietárias que concentram conhecimento técnico e poder contratual nas mãos de poucos fornecedores. A dependência de fornecedores, discutida na literatura de sistemas de informação e de administração pública sob o termo *vendor lock-in*, surge quando contratos, dados e processos internos ficam vinculados a tecnologias de difícil substituição. Municípios com equipes reduzidas tendem a aceitar soluções de prateleira sem exigir documentação suficiente, portabilidade de dados ou cláusulas de transferência de conhecimento. Com o tempo, o custo de trocar de fornecedor se torna elevado o suficiente para que a administração perca capacidade real de negociar condições melhores ou de descontinuar um sistema que já não atende ao interesse público.

A contratação pública de soluções de inteligência artificial precisa, portanto, considerar o ciclo de vida completo da tecnologia, e não apenas o momento da aquisição. Um preço inicial atrativo pode ocultar custos recorrentes de licenciamento, manutenção, integração com outros sistemas, armazenamento de dados e atualização de modelos. Também é necessário avaliar, desde o edital, quais são os direitos da administração sobre os dados produzidos, o acesso a registros de auditoria, os prazos de comunicação de incidentes de segurança e os mecanismos de fiscalização contratual. A ideia de soberania tecnológica local, nesse sentido, não implica que o município deva desenvolver internamente tudo o que usa, algo inviável para a grande maioria das administrações municipais brasileiras, mas sim que deva conservar autonomia decisória sobre finalidades, critérios e continuidade dos sistemas que adota.

Governança de dados e infraestrutura informacional

A inteligência artificial depende, antes de tudo, da qualidade dos dados sobre os quais opera. Governos locais frequentemente acumulam bases de dados fragmentadas, duplicadas e desatualizadas, produto de décadas de sistemas isolados criados para atender demandas específicas de cada secretaria. Kitchin (2014) chama atenção para o fato de que dados nunca

são "crus": eles carregam decisões metodológicas, categorias e omissões que refletem a forma como a instituição que os produziu enxergava seu próprio trabalho. Secretarias diferentes costumam usar conceitos distintos para representar as mesmas pessoas, territórios ou serviços, o que dificulta a integração de bases e cria inconsistências que passam despercebidas até que um sistema automatizado as amplifique. A automação de bases de dados frágeis não corrige esse problema; ao contrário, tende a acelerar erros e ampliar exclusões que já existiam, mas que ficavam diluídas quando as decisões dependiam de análise humana caso a caso.

A governança de dados, como campo de prática e de pesquisa, compreende padrões técnicos, definição de responsabilidades, segurança da informação, mecanismos de controle de qualidade, interoperabilidade entre sistemas e uso ético das informações. Janssen e van den Hoven (2015) destacam que a abertura de dados governamentais, tema que ganhou força na última década, precisa ser conciliada com privacidade e proteção de grupos vulneráveis: nem todo dado público pode ser divulgado ou combinado com outras bases sem risco de reidentificação ou de exposição indevida de informações sensíveis. É necessário, portanto, definir claramente quem produz, valida, acessa e corrige cada tipo de informação, e sob quais condições ela pode circular entre órgãos.

A capacidade informacional de um governo local também depende de infraestrutura básica e de continuidade operacional, dimensões que costumam receber menos atenção do que os aspectos mais visíveis da inovação. Sistemas municipais precisam de rotinas de cópia de segurança, atualização periódica, controle de acesso e planos de recuperação diante de falhas ou ataques cibernéticos. A indisponibilidade de uma base de dados essencial pode interromper serviços dos quais a população depende diretamente, como cadastros de saúde ou de assistência social. A governança da inteligência artificial, nesse sentido, não pode ser tratada como um tema isolado; ela precisa estar integrada à gestão de informação e de segurança que já deveria existir, independentemente do uso de algoritmos mais sofisticados.

Inteligência artificial confiável, explicabilidade e valor público

A literatura internacional sobre inteligência artificial confiável tem enfatizado princípios de legalidade, ética e robustez técnica como condições para que sistemas automatizados sejam aceitáveis em contextos de alto impacto social (Floridi et al., 2018). No setor público, esses princípios precisam ser articulados com a noção de valor público, desenvolvida por

Moore (1995) e retomada por Bozeman (2007) em sua discussão sobre valores públicos e interesse coletivo. Uma solução de inteligência artificial não pode ser avaliada apenas por critérios de eficiência ou de acurácia estatística; ela precisa ser examinada também segundo equidade na distribuição de seus resultados, transparência sobre seu funcionamento, legitimidade do processo que levou à sua adoção e capacidade efetiva de produzir benefícios coletivos, e não apenas ganhos administrativos internos.

A explicabilidade de sistemas automatizados possui uma dimensão técnica e outra institucional, que costumam ser tratadas separadamente, mas que precisam convergir na prática administrativa. Um modelo pode ser estatisticamente complexo e, ainda assim, a decisão pública que dele resulta precisa ser justificada em linguagem compreensível para quem é afetado por ela. O cidadão atingido por uma decisão automatizada, ou parcialmente automatizada, deve poder saber quais critérios foram relevantes para o resultado, quem é o responsável institucional por essa decisão e como solicitar sua revisão. Raji et al. (2020) descrevem essa lacuna como um problema de responsabilização algorítmica: sistemas são frequentemente implantados sem mecanismos que permitam auditoria externa ou correção de erros identificados após o uso. Sem esses elementos, a automação tende a enfraquecer direitos procedimentais que a administração pública já reconhecia antes da introdução de qualquer algoritmo.

A robustez de um sistema de inteligência artificial também precisa incluir uma noção de desempenho social, além do desempenho puramente técnico. Um sistema pode funcionar exatamente conforme sua especificação e, ainda assim, produzir resultados injustos quando essa especificação não considerou adequadamente a diversidade dos públicos atendidos. Testes de desempenho devem avaliar o comportamento do sistema em diferentes grupos e territórios, e não apenas médias gerais que podem esconder impactos concentrados em populações específicas. A governança responsável exige, portanto, monitoramento contínuo ao longo de toda a operação do sistema, e não apenas na fase de homologação, além da capacidade institucional de interromper um uso quando os danos identificados superarem os benefícios esperados.

Viés algorítmico e discricionariedade burocrática

Algoritmos treinados com dados históricos tendem a reproduzir desigualdades presentes nas próprias práticas administrativas que geraram esses dados. Registros incompletos sobre determinados bairros ou grupos sociais podem levar um sistema a subestimar necessidades reais,

simplesmente porque essas necessidades nunca foram adequadamente registradas no passado. Variáveis aparentemente neutras, como o endereço ou o histórico de atendimentos, podem funcionar como indicadores indiretos de renda, raça ou território, reproduzindo padrões de discriminação sem que isso apareça de forma explícita no desenho do sistema (O'Neil, 2016; Eubanks, 2018). Eubanks (2018) documenta, em estudos de caso norte-americanos, como sistemas automatizados de triagem em políticas sociais tendem a intensificar a vigilância sobre populações já mais vulneráveis, ao mesmo tempo em que se apresentam publicamente como ferramentas de eficiência e objetividade.

A literatura sobre street-level bureaucracy oferece um vocabulário importante para entender essa transformação. Lipsky (2010) mostrou que servidores da linha de frente, como agentes de assistência social, professores e policiais, exercem discricionariedade considerável ao aplicar políticas públicas em situações concretas, preenchendo lacunas que normas gerais não conseguem antecipar. Bovens e Zouridis (2002) descreveram a passagem de burocracias de linha de frente para burocracias de sistema, nas quais decisões que antes dependiam do julgamento individual de um servidor passam a ser mediadas, ou até substituídas, por sistemas de informação padronizados. Mais recentemente, Busch e Henriksen (2018) revisaram sistematicamente a literatura sobre discricionariedade digital, concluindo que tecnologias de informação não eliminam a discricionariedade administrativa, mas a deslocam para outros pontos do processo, muitas vezes menos visíveis do que o momento do atendimento direto ao cidadão.

A substituição da discricionariedade burocrática tradicional por regras automatizadas não elimina escolhas de valor; ela apenas as desloca para etapas anteriores, como a definição do problema a ser resolvido, a seleção das variáveis que compõem o modelo e a configuração dos parâmetros que definem o que conta como um resultado aceitável. Bullock (2019) argumenta que essa transferência de discricionariedade para o momento do design do sistema exige que profissionais de políticas públicas participem ativamente do desenvolvimento e da avaliação de ferramentas de inteligência artificial, e não apenas de sua homologação final, sob pena de que decisões substantivas sejam tomadas por equipes técnicas sem o conhecimento necessário sobre o contexto social em que o sistema será aplicado. Barth e Arnold (1999) já haviam antecipado, em um dos primeiros artigos sobre o tema na administração pública norte-americana, que sistemas especialistas tendem a reduzir a discricionariedade percebida pelos servidores sem necessariamente reduzir a discricionariedade real embutida no sistema, um ponto que permanece atual mais de duas décadas depois.

A supervisão humana sobre sistemas automatizados precisa ser efetiva, não apenas formal. Inserir um servidor ao final do processo de decisão, com a função de aprovar o resultado sugerido pelo sistema, não é suficiente quando esse servidor não dispõe de informação, tempo ou autonomia institucional para contrariar a recomendação automatizada. Peeters (2020) chama esse fenômeno de agência limitada do operador humano diante do algoritmo, situação em que a supervisão se torna simbólica porque o custo de discordar do sistema, em termos de tempo e de justificativa exigida, é sistematicamente maior do que o custo de simplesmente aceitar sua recomendação. A governança responsável precisa, portanto, definir com clareza quais situações exigem revisão humana obrigatória, registrar formalmente os casos de divergência entre servidor e sistema, e proteger institucionalmente o exercício do julgamento profissional, inclusive quando esse julgamento contraria a recomendação automatizada.

PERCURSO METODOLÓGICO

O estudo utiliza a revisão teórica integrativa como estratégia de pesquisa, abordagem reconhecida como adequada à construção de modelos conceituais que articulam campos de conhecimento distintos e que ainda não dialogam de forma sistemática entre si (Torraco, 2016). Diferentemente de uma revisão sistemática voltada à síntese de evidências empíricas sobre uma pergunta específica, a revisão integrativa permite combinar literaturas teóricas de tradições diferentes para propor uma estrutura analítica nova, capaz de orientar tanto a reflexão acadêmica quanto a prática administrativa.

Foram reunidas contribuições de quatro áreas principais: administração pública e sociologia histórica do Estado, com ênfase na literatura sobre capacidade estatal; sistemas de informação e governo digital, incluindo estudos sobre plataformização e dependência de fornecedores; ética e governança algorítmica, com foco em inteligência artificial confiável e responsabilização; e estudos sobre discricionariedade burocrática, tanto na tradição clássica quanto na literatura mais recente sobre burocracia digital.

O percurso de construção do modelo ocorreu em quatro etapas sucessivas. A primeira etapa identificou, a partir da literatura revisada, as principais limitações enfrentadas por governos locais na adoção de tecnologias digitais, com atenção especial a restrições de pessoal, orçamento, dados e coordenação interna. A segunda etapa cruzou as dimensões clássicas de capacidade estatal, como autonomia, coordenação e inserção social, com as exigências específicas impostas pelo ciclo de vida de sistemas de

inteligência artificial, desde a concepção até a eventual descontinuação. A terceira etapa formulou, a partir desse cruzamento, seis capacidades específicas, cada uma correspondente a um conjunto coerente de recursos, rotinas e relações institucionais necessárias à governança responsável. A quarta etapa examinou as interdependências entre essas capacidades e os riscos de falhas em cascata, quando a fragilidade em uma dimensão compromete o funcionamento das demais.

Foram utilizadas prioritariamente referências publicadas até 2024, respeitando o marco editorial do volume em que este artigo foi originalmente submetido. É importante registrar que o modelo não pretende medir empiricamente a maturidade de municípios específicos nesta etapa da pesquisa; seu propósito é oferecer categorias conceituais consistentes, capazes de serem operacionalizadas em indicadores e instrumentos de coleta de dados por pesquisas subsequentes, inclusive estudos de caso, levantamentos survey e análises comparativas entre municípios de diferentes portes e regiões.

MODELO DE CAPACIDADES ESTATAIS PARA A GOVERNANÇA RESPONSÁVEL DA IA LOCAL

O modelo proposto é composto por seis capacidades interdependentes: estratégica, regulatória, técnica, informacional, organizacional e participativa. Cada uma corresponde a um conjunto específico de recursos, rotinas e relações institucionais, mas nenhuma delas produz, isoladamente, governança responsável. A ausência ou fragilidade de uma dimensão limita o funcionamento das demais: competência técnica sem direção estratégica tende a produzir automação sem prioridades claras; regulação bem redigida, mas sem dados confiáveis e pessoal qualificado para aplicá-la, tende a permanecer apenas formal; e participação social sem explicabilidade mínima não garante controle real sobre os sistemas, porque os cidadãos consultados não têm como avaliar o que está sendo decidido em seu nome.

Capacidade estratégica

A capacidade estratégica refere-se à habilidade da administração de definir problemas públicos, estabelecer prioridades e reconhecer limites para o uso de tecnologia. Governos locais precisam distinguir situações em que a inteligência artificial oferece um benefício real de situações em que melhorias mais simples de processo, como a integração de formulários ou a revisão de um fluxo de atendimento, seriam suficientes e mais baratas. O ponto de partida de qualquer projeto deveria ser a necessidade pública

identificada, e não a disponibilidade de uma ferramenta no mercado ou a pressão de um edital de financiamento externo que condiciona recursos à adoção de determinada tecnologia.

Planos municipais de transformação digital precisam estabelecer objetivos claros, critérios de seleção de projetos e indicadores de acompanhamento. Os projetos de inteligência artificial devem estar alinhados ao planejamento municipal mais amplo e às políticas setoriais já existentes, evitando que iniciativas de automação se tornem ilhas desconectadas da estratégia de governo. A definição estratégica também precisa indicar, de forma explícita, quais são as áreas consideradas de alto risco e quais tipos de decisão não deveriam ser integralmente delegados a sistemas automatizados, como decisões que afetam diretamente a liberdade, a saúde ou o acesso a benefícios essenciais de subsistência.

A capacidade estratégica inclui, ainda, uma dimensão de visão de longo prazo que costuma ser negligenciada em contextos de alternância política frequente. Sistemas de inteligência artificial criam custos e dependências que permanecem muito além da duração de um mandato eletivo. A continuidade dessas iniciativas exige documentação adequada, previsão orçamentária plurianual e preservação de conhecimento institucional acumulado, de modo que mudanças na gestão política não eliminem controles essenciais nem interrompam projetos que estavam produzindo resultados positivos apenas por motivos de descontinuidade administrativa.

Capacidade regulatória

No nível local, a capacidade regulatória se manifesta em normas internas, contratos, protocolos operacionais e procedimentos de auditoria. Mesmo quando a legislação geral sobre proteção de dados ou sobre inteligência artificial é definida em âmbito nacional, o município continua responsável por definir como sistemas específicos serão adquiridos, implantados e utilizados em seu território. Essas regras internas precisam ser proporcionais ao risco de cada aplicação, evitando tanto a ausência completa de controle quanto a criação de exigências burocráticas desnecessárias para sistemas de baixo impacto.

Os contratos administrativos que envolvem inteligência artificial devem prever, de forma explícita, a finalidade do sistema, os critérios de desempenho esperado, os requisitos de transparência e segurança, e as regras de propriedade e portabilidade dos dados envolvidos. Também precisam definir responsabilidade contratual por falhas, obrigação do fornecedor de comunicar alterações relevantes no funcionamento do sistema

e garantia de acesso a evidências necessárias para auditoria externa. Cláusulas genéricas de confidencialidade comercial, comuns em contratos de tecnologia, não podem funcionar como barreira absoluta ao controle público sobre decisões que afetam diretamente os cidadãos.

A avaliação de impacto algorítmico pode ser utilizada como instrumento prévio à implantação de aplicações consideradas sensíveis. Esse tipo de avaliação deve identificar os públicos potencialmente afetados, os dados que serão utilizados, os riscos envolvidos, as medidas de mitigação previstas e os canais de recurso disponíveis para quem se sentir prejudicado por uma decisão automatizada. Sistemas classificados como de baixo risco podem seguir um processo simplificado de avaliação, desde que preservem um mínimo de documentação que permita reconstituir, posteriormente, como e por que aquela ferramenta foi adotada.

Capacidade técnica

A capacidade técnica exige que a administração conte com servidores ou equipes capazes de compreender, ao menos em suas linhas gerais, a arquitetura dos sistemas utilizados, a origem e as limitações dos dados envolvidos e as métricas usadas para avaliar seu desempenho. Não é necessário que todos os servidores envolvidos sejam programadores ou cientistas de dados, mas os gestores responsáveis precisam possuir letramento técnico suficiente para questionar fornecedores de forma qualificada e para interpretar criticamente os resultados apresentados por um sistema.

Equipes interdisciplinares tendem a ser mais adequadas do que núcleos exclusivamente tecnológicos para conduzir projetos de inteligência artificial no setor público. Profissionais de tecnologia da informação, de assessoria jurídica, de formulação de políticas públicas, de proteção de dados pessoais e das áreas finalísticas diretamente afetadas deveriam participar conjuntamente das decisões relevantes. Essa diversidade de perspectivas reduz pontos cegos que uma equipe tecnicamente homogênea tende a não perceber, e aproxima o desenho do sistema do contexto concreto em que ele será aplicado.

A capacitação e a retenção de profissionais qualificados são desafios estruturais para a maioria dos municípios brasileiros, especialmente aqueles de pequeno e médio porte, que competem em desvantagem com o setor privado pela atração desse tipo de talento. A cooperação com universidades, consórcios intermunicipais e outras esferas de governo pode ampliar os recursos disponíveis. É importante, no entanto, que essas parcerias

produzam efetiva transferência de conhecimento para dentro da administração municipal; a dependência permanente de consultorias externas, ainda que tecnicamente competentes, mantém a fragilidade institucional que o modelo busca justamente reduzir.

Capacidade informacional

A capacidade informacional abrange qualidade, integração, proteção e governança dos dados utilizados pela administração. Antes de aplicar qualquer sistema de inteligência artificial, o município precisa conhecer suas próprias bases de dados, identificar suas lacunas e compreender os critérios que orientaram sua construção histórica. Dados de baixa qualidade não são corrigidos automaticamente pela sofisticação de um modelo estatístico; ao contrário, modelos mais sofisticados tendem a amplificar erros presentes em dados de baixa qualidade, produzindo resultados que parecem confiáveis justamente por serem apresentados com precisão numérica.

A interoperabilidade entre sistemas exige padrões técnicos e semânticos compartilhados. Duas secretarias de uma mesma prefeitura podem usar o mesmo termo para conceitos diferentes, ou termos diferentes para o mesmo conceito, o que dificulta qualquer tentativa de integração de bases. Sem harmonização prévia desses padrões, tentativas de integração tendem a produzir inconsistências que só se tornam visíveis quando já causaram algum dano administrativo. Catálogos de dados, dicionários de metadados e identificadores comuns entre sistemas são instrumentos relativamente simples, mas de grande importância prática para reduzir esse tipo de problema.

A privacidade dos dados pessoais precisa ser incorporada desde o desenho dos sistemas, e não como uma verificação posterior de conformidade legal. As coletas de dados realizadas para viabilizar um sistema de inteligência artificial precisam ser necessárias e proporcionais à finalidade declarada. O acesso a esses dados deve ser controlado por perfis de permissão claros, e dados sensíveis não podem ser reutilizados para finalidades incompatíveis com aquela que justificou sua coleta original, sem que exista base legal específica e avaliação prévia dos riscos envolvidos.

Capacidade organizacional

A capacidade organizacional refere-se à coordenação interna, às rotinas de trabalho e à cultura institucional que sustentam a operação cotidiana de sistemas de inteligência artificial. A implantação de qualquer sistema desse

tipo altera tarefas e responsabilidades dentro da administração. Quando essa mudança não é discutida abertamente com as equipes envolvidas, tende a haver resistência silenciosa, uso inadequado da ferramenta ou, no extremo oposto, confiança excessiva em recomendações automatizadas que deveriam ser questionadas com mais cuidado.

Cada projeto relevante de inteligência artificial precisa contar com responsáveis claramente designados nas áreas de negócio, técnica e de controle interno. A governança desses projetos não deveria ficar restrita ao departamento de tecnologia da informação, ainda que este mantenha papel central na operação. As secretarias finalísticas, que conhecem as consequências práticas de cada decisão automatizada sobre a vida das pessoas atendidas, precisam acompanhar de perto os indicadores de desempenho e os incidentes registrados.

A aprendizagem institucional depende de registros sistemáticos, avaliação periódica e disposição para corrigir rumo quando necessário. Falhas identificadas deveriam produzir mudanças concretas em processos e em programas de treinamento, e não apenas justificativas administrativas para manter o status quo. Uma cultura organizacional punitiva tende a incentivar a ocultação de incidentes, porque servidores temem represálias por reportar problemas; uma cultura organizacional responsável, por outro lado, distingue claramente entre erro pontual, negligência reiterada e processo legítimo de aprendizado, tratando cada situação de forma proporcional.

Capacidade participativa

A capacidade participativa permite incluir cidadãos, movimentos sociais e organizações da sociedade civil na definição e na avaliação de sistemas de inteligência artificial. Aplicações que afetam diretamente o acesso a benefícios sociais, a intensidade da fiscalização ou a qualidade de serviços públicos precisam levar em conta as experiências concretas dos grupos impactados por essas decisões. Consultas públicas bem conduzidas podem revelar barreiras práticas que dificilmente apareceriam apenas na análise de dados administrativos.

A explicabilidade voltada ao público em geral exige uma linguagem acessível, diferente daquela usada em relatórios técnicos internos. Documentação técnica detalhada é importante para fins de auditoria, mas não substitui a comunicação pública dirigida ao cidadão comum. O município precisa informar, em termos compreensíveis, qual é a finalidade do sistema, quais critérios gerais orientam suas decisões, quais riscos foram

identificados e quais canais existem para contestar um resultado considerado injusto.

O direito de recurso precisa ser operacional, e não apenas previsto formalmente em algum regulamento pouco conhecido. Cidadãos afetados por uma decisão automatizada devem poder solicitar sua revisão por uma pessoa qualificada para reexaminar o caso, sem que esse processo exija conhecimento técnico especializado da parte de quem recorre. Registros sistemáticos desses recursos, incluindo seu resultado, ajudam a administração a identificar padrões recorrentes de erro e a aperfeiçoar continuamente o sistema, transformando a contestação individual em fonte de aprendizado coletivo.

Interdependência e falhas em cascata

As seis capacidades descritas formam, na prática, um sistema único, e não um conjunto de itens independentes de uma lista de verificação. Uma contratação pode estar juridicamente correta e ser, ao mesmo tempo, tecnicamente frágil; dados podem estar bem integrados tecnicamente e, ainda assim, reproduzir exclusões sociais que nunca foram discutidas com os grupos afetados; equipes podem ser tecnicamente qualificadas e não possuir autonomia institucional para interromper um projeto problemático. A maturidade de governança de um município depende, portanto, do equilíbrio relativo entre essas seis dimensões, e não do desempenho isolado de qualquer uma delas.

Falhas em cascata ocorrem quando a fragilidade em uma capacidade alimenta problemas em outra. Dados de baixa qualidade produzem resultados administrativos inadequados; baixa competência técnica impede que essa inadequação seja identificada a tempo; contratos redigidos de forma opaca dificultam a correção do problema mesmo depois de identificado; e a ausência de canal de recurso efetivo transforma um erro pontual em dano persistente para quem foi afetado por ele. A análise sistêmica dessas cadeias de causalidade permite identificar, em cada caso concreto, quais são os pontos mais críticos para intervenção prioritária.

O modelo sugere que investimentos em transformação digital deveriam ser priorizados conforme o elo mais frágil da cadeia, e não conforme a novidade tecnológica disponível no mercado. Em muitos municípios, a aquisição de uma nova tecnologia é, objetivamente, menos urgente do que melhorar a qualidade dos dados existentes, revisar contratos vigentes ou investir em formação de pessoal. A maturidade digital, sob essa perspectiva, não é uma corrida por quantidade de sistemas implantados, mas um processo gradual

de consolidação institucional que precisa avançar de forma equilibrada entre as seis capacidades.

CICLO DE VIDA DA INTELIGÊNCIA ARTIFICIAL NO GOVERNO LOCAL

A governança responsável precisa acompanhar todas as fases do ciclo de vida de um sistema de inteligência artificial: concepção, contratação, desenvolvimento, implantação, operação e, eventualmente, descontinuação. Na fase de concepção, define-se com clareza o problema público a ser enfrentado e avaliam-se alternativas de solução, tecnológicas ou não. Na fase de contratação, estabelecem-se os requisitos técnicos, contratuais e de governança que orientarão a aquisição. Na fase de desenvolvimento, testam-se os dados disponíveis e o comportamento dos modelos propostos antes de qualquer uso em produção. Na fase de implantação, preparam-se as pessoas que operarão o sistema e os procedimentos que orientarão seu funcionamento cotidiano.

Durante a fase de operação, o sistema precisa ser monitorado de forma contínua. Mudanças nos dados de entrada ao longo do tempo, fenômeno conhecido na literatura técnica como deriva de dados, podem reduzir progressivamente o desempenho de um modelo sem que isso seja percebido de imediato pelos gestores responsáveis. Incidentes registrados, reclamações formais e decisões automatizadas que foram revertidas por recurso são indicadores importantes de que algo pode estar funcionando de forma inadequada. A avaliação de um sistema em operação não deveria se limitar à sua disponibilidade técnica, isto é, ao fato de estar simplesmente funcionando sem interrupções.

A descontinuação de um sistema é parte legítima, e às vezes necessária, de sua governança. Sistemas que perderam sua finalidade original, que se tornaram tecnicamente obsoletos ou que produzem riscos desproporcionais aos benefícios que geram deveriam ser encerrados, ainda que isso represente reconhecer publicamente que um investimento anterior não produziu os resultados esperados. Nesse processo, a administração precisa preservar os dados relevantes, os registros históricos de decisões tomadas e a continuidade do serviço público que dependia daquele sistema, evitando que a descontinuação de uma ferramenta se traduza em interrupção do atendimento à população.

Avaliação de risco e proporcionalidade

Nem todas as aplicações de inteligência artificial no setor público apresentam o mesmo nível de risco. Um assistente virtual que fornece

informações gerais sobre horários de atendimento é qualitativamente diferente de um sistema que nega o acesso a um benefício social ou que prioriza determinados endereços para fiscalização tributária. A intensidade dos controles de governança deveria acompanhar o impacto potencial da decisão, a escala de pessoas afetadas e a possibilidade concreta de dano irreversível.

Critérios de avaliação de risco podem considerar, entre outros fatores, os direitos fundamentais potencialmente afetados, a vulnerabilidade social dos públicos envolvidos, o grau de autonomia decisória atribuído ao sistema, a qualidade dos dados utilizados e a reversibilidade da decisão tomada. Yeung (2018) descreve a regulação algorítmica justamente como um espectro de intensidade regulatória que deveria variar conforme esses fatores, em vez de tratar toda automação com a mesma régua. Aplicações classificadas como de alto risco exigem avaliação prévia independente, testes rigorosos antes da implantação e supervisão humana reforçada durante toda a operação.

A proporcionalidade na governança evita dois extremos igualmente problemáticos: a ausência completa de controle, que expõe a administração e os cidadãos a riscos evitáveis, e a burocracia indiscriminada, que trata sistemas de baixo risco com o mesmo rigor exigido para sistemas de alto impacto, consumindo recursos escassos sem ganho correspondente de segurança. Governos locais com recursos limitados precisam, necessariamente, concentrar seus esforços de controle onde os riscos identificados são efetivamente maiores.

Contratação pública e gestão de fornecedores

A contratação é um momento decisivo na trajetória de um sistema de inteligência artificial no setor público, porque muitas das limitações futuras de governança já ficam definidas nessa etapa, ainda que seus efeitos só apareçam anos depois. Editais e termos de referência deveriam descrever com precisão o problema público a ser resolvido, os dados envolvidos, os requisitos de desempenho, segurança e transparência. A contratação orientada por marca comercial ou por promessas genéricas de eficiência tende a reduzir a concorrência entre fornecedores e a dificultar qualquer avaliação técnica objetiva das propostas recebidas.

Provas de conceito podem ser úteis para reduzir a incerteza sobre uma solução antes de um contrato de longo prazo, desde que sejam conduzidas com critérios claros de avaliação e não criem, na prática, uma dependência prematura do fornecedor testado. O município precisa manter controle sobre os dados utilizados durante esses experimentos e impedir que pilotos

informais sejam convertidos automaticamente em contratos de longa duração sem o devido processo licitatório e sem avaliação criteriosa dos resultados obtidos.

A gestão contratual continua após a assinatura do contrato, e essa fase costuma receber menos atenção institucional do que a fase de aquisição. Níveis de serviço acordados, incidentes registrados, atualizações realizadas e eventuais alterações no funcionamento do modelo precisam ser acompanhados de forma sistemática pela administração. A terceirização da operação de um sistema não transfere, em hipótese alguma, a responsabilidade pública pelas decisões que dele resultam; o fornecedor pode ser responsabilizado civilmente por falhas contratuais, mas a responsabilidade política e administrativa permanece do lado do poder público.

Auditoria, monitoramento e prestação de contas

Auditorias de sistemas de inteligência artificial podem examinar a qualidade dos dados utilizados, o desempenho técnico do sistema, a segurança da informação, a existência de vieses discriminatórios e a conformidade com normas legais e contratuais aplicáveis. A periodicidade dessas auditorias deveria refletir o nível de risco da aplicação e a frequência com que o sistema é atualizado. Sistemas que aprendem continuamente com novos dados, ou que recebem atualizações frequentes por parte do fornecedor, precisam de mecanismos de controle contínuo, e não apenas de uma auditoria pontual realizada no momento da homologação inicial.

Os indicadores usados para monitorar um sistema deveriam incluir, entre outros, taxas gerais de acurácia, taxas de erro desagregadas por grupo social e território, número de decisões contestadas com sucesso, tempo médio de resposta a recursos, incidentes registrados e nível de satisfação dos usuários atendidos. Métricas agregadas em nível municipal podem ocultar impactos concentrados em bairros ou populações específicas, razão pela qual a desagregação dos indicadores é tão importante quanto sua própria existência.

A prestação de contas sobre sistemas automatizados requer documentação consistente ao longo do tempo. Registros de versões do sistema, decisões relevantes tomadas sobre seu funcionamento e intervenções humanas realizadas durante sua operação permitem reconstruir, posteriormente, todo o processo que levou a um determinado resultado. A transparência sobre esses registros deve respeitar exigências legítimas de segurança da informação e de proteção de dados pessoais, mas essas restrições não

podem ser usadas para transformar o segredo comercial de um fornecedor em uma barreira absoluta contra a fiscalização pública.

Participação, contestação e controle social

Conselhos municipais, audiências públicas e consultas específicas podem acompanhar a implantação de aplicações de inteligência artificial consideradas relevantes para a comunidade. Essa participação precisa ocorrer antes da implantação do sistema e continuar durante todo o período de sua avaliação, e não apenas depois que a decisão de adotar determinada ferramenta já foi tomada e formalizada. Informar a população somente após a decisão estar consolidada reduz significativamente a legitimidade de todo o processo, ainda que a comunicação em si seja tecnicamente correta.

Os canais de contestação disponíveis aos cidadãos precisam ser acessíveis, gratuitos e capazes de gerar resposta em prazo razoável. As pessoas afetadas por uma decisão automatizada não deveriam ser obrigadas a compreender detalhes técnicos do sistema para poder solicitar sua revisão; o ônus de justificar tecnicamente a decisão tomada é do poder público, não do cidadão que se sente prejudicado por ela. Essa inversão do ônus da explicação, presente implicitamente em boa parte da literatura sobre devido processo administrativo, ganha nova relevância diante de sistemas cuja complexidade técnica pode facilmente intimidar quem busca contestar uma decisão.

Órgãos de controle interno, ouvidorias municipais e autoridades de proteção de dados podem colaborar de forma complementar nesse arranjo de controle social. A distribuição dessas funções entre diferentes instâncias, em vez de concentrá-las em um único departamento responsável tanto pela operação quanto pela fiscalização do sistema, cria contrapesos institucionais e reduz o risco de concentração excessiva de poder decisório sobre tecnologias que afetam diretamente a vida da população.

APLICAÇÃO DO MODELO E INDICADORES DE MATURIDADE

O modelo aqui desenvolvido pode apoiar diagnósticos municipais sobre o estágio atual de governança da inteligência artificial. Em um estágio inicial, os projetos costumam ser isolados, dependentes de fornecedores específicos e sem qualquer coordenação entre secretarias. Em um estágio intermediário, normas internas e equipes técnicas já existem, mas sua integração ainda é desigual entre diferentes áreas da administração. Em um estágio avançado, a governança de sistemas de inteligência artificial está

incorporada ao planejamento estratégico e aos mecanismos regulares de controle interno e externo do município.

Indicadores estratégicos podem incluir a existência de um plano formal de transformação digital, critérios explícitos de seleção de projetos e alinhamento demonstrável com as políticas setoriais vigentes. Indicadores regulatórios envolvem a existência de modelos contratuais adequados, de avaliações de impacto algorítmico realizadas e de canais de recurso formalmente instituídos. Indicadores técnicos examinam a composição e a autonomia das equipes responsáveis pela operação dos sistemas. Indicadores informacionais avaliam a qualidade, a interoperabilidade e o nível de proteção das bases de dados utilizadas.

Indicadores organizacionais incluem a existência de mecanismos formais de coordenação entre áreas, programas regulares de treinamento e processos estruturados de aprendizagem institucional a partir de incidentes registrados. Indicadores participativos observam a frequência e a qualidade de consultas realizadas, o nível de transparência ativa mantido pela administração e a efetividade dos canais de contestação disponíveis à população. É importante destacar que essa avaliação deveria orientar a melhoria contínua da governança municipal, e não servir apenas para produzir um ranking simplificado entre prefeituras, que dificilmente captaria as diferenças reais de contexto entre elas.

Municípios interessados em avançar nesse diagnóstico podem começar por um inventário completo dos sistemas automatizados já em uso e das bases de dados que os alimentam. Conhecer o que já está em operação, ainda que de forma informal ou pouco documentada, é condição necessária para qualquer tentativa séria de governar essas tecnologias. A partir desse inventário, a administração pode classificar os riscos envolvidos, definir responsáveis institucionais claros para cada sistema e priorizar correções nos casos mais urgentes.

Consórcios e redes intermunicipais podem compartilhar especialistas e padrões técnicos comuns entre diferentes prefeituras. Essa cooperação tende a reduzir custos individuais e a fortalecer o poder de negociação de cada município diante de fornecedores, desde que as responsabilidades de cada ente permaneçam claramente definidas ao longo de todo o arranjo cooperativo.

Planejamento de portfólio e priorização de projetos

Governos locais tendem a iniciar sua trajetória de transformação digital por meio de projetos dispersos, motivados por demandas administrativas imediatas ou por oportunidades pontuais de financiamento externo. Essa trajetória, ainda que compreensível diante das restrições orçamentárias enfrentadas pela maioria dos municípios, pode gerar sobreposição de sistemas semelhantes e a adoção de soluções que pouco se relacionam com as prioridades públicas efetivamente definidas pela administração. A capacidade estratégica precisa transformar esse conjunto disperso de iniciativas em um portfólio coerente, que permita comparar custos, riscos, dependências contratuais e benefícios esperados entre diferentes projetos concorrentes por recursos escassos.

A priorização de projetos deveria considerar sua materialidade pública, isto é, sua relevância efetiva para o interesse coletivo. Problemas com alto impacto social, grande volume de atendimentos ou dificuldade persistente de coordenação entre áreas podem justificar um investimento analítico mais aprofundado. É importante, no entanto, que o uso de inteligência artificial não seja presumido como solução automática para esses problemas. Em muitos casos concretos, a simples revisão de formulários administrativos, a integração de bases de dados já existentes ou a simplificação de procedimentos burocráticos oferece um benefício maior e um risco menor do que a adoção de um sistema automatizado sofisticado.

O portfólio municipal deveria registrar também projetos interrompidos, pilotos experimentais e aplicações adquiridas isoladamente por secretarias específicas, sem passar por qualquer processo de avaliação centralizada. A ausência de um inventário unificado impede que a administração conheça, de forma completa, seus próprios compromissos contratuais e os dados que efetivamente processa. Um cadastro municipal de sistemas automatizados constitui instrumento básico, e de baixo custo de implementação, para qualquer estratégia de governança digital, além de poder apoiar diretamente iniciativas de transparência pública voltadas à população.

A análise de portfólio precisa ser periódica, e não realizada uma única vez no início de um projeto de transformação digital. Mudanças regulatórias, tecnológicas e orçamentárias alteram, com o tempo, a conveniência relativa das soluções já implantadas. Projetos que apresentavam valor claro em seu momento inicial podem se tornar redundantes ou incompatíveis com novos padrões técnicos adotados posteriormente pela administração. A revisão periódica do portfólio permite realocar recursos escassos e evitar a manutenção de sistemas obsoletos apenas em razão do investimento já

realizado no passado, armadilha conhecida na literatura de gestão como custo irrecuperável.

Cooperação federativa e consórcios intermunicipais

A desigualdade de capacidade institucional entre municípios brasileiros, amplamente documentada pela literatura de federalismo e políticas públicas, sugere a necessidade de arranjos cooperativos específicos para a governança da inteligência artificial. Consórcios intermunicipais podem compartilhar especialistas, infraestrutura de tecnologia da informação, padrões comuns de dados e serviços de auditoria independente. Essa estratégia tende a reduzir custos individuais e a ampliar o poder de negociação de cada município diante de grandes fornecedores de tecnologia, benefício especialmente relevante para municípios de pequeno e médio porte, que isoladamente têm pouca capacidade de barganha contratual.

A cooperação intermunicipal, no entanto, precisa preservar mecanismos claros de governança. O compartilhamento de uma mesma plataforma tecnológica entre diferentes prefeituras não elimina a responsabilidade individual de cada ente sobre as decisões administrativas tomadas com apoio dessa plataforma. Contratos e protocolos de cooperação devem definir com precisão a propriedade dos dados envolvidos, as regras de acesso entre municípios participantes, as responsabilidades de segurança da informação, o financiamento compartilhado e os procedimentos previstos para saída de qualquer ente do arranjo cooperativo. Um consórcio sem essas regras claramente definidas corre o risco de apenas deslocar o problema da dependência de um fornecedor privado para uma dependência equivalente em relação a uma estrutura regional mal governada.

Estados e a União também podem apoiar municípios por meio de padrões técnicos comuns, programas de capacitação e infraestrutura compartilhada de computação e armazenamento. A padronização promovida por esferas superiores de governo facilita a interoperabilidade entre sistemas municipais distintos, mas precisa admitir margem de adaptação às realidades locais. Modelos excessivamente uniformes, pensados a partir da realidade de grandes capitais, podem ignorar diferenças territoriais e institucionais relevantes presentes em municípios menores ou em regiões com características muito distintas. A cooperação federativa tende a ser mais produtiva quando combina uma referência técnica comum com autonomia real de implementação por parte de cada município.

Universidades públicas e institutos de pesquisa podem contribuir com formação técnica, avaliação independente de sistemas e pesquisa aplicada

voltada às necessidades concretas da administração local. Essas parcerias precisam ser formalizadas de modo a preservar simultaneamente a independência acadêmica das instituições envolvidas e a utilidade prática dos resultados produzidos para a gestão municipal. Projetos de curta duração, comuns nesse tipo de cooperação, deveriam produzir documentação adequada e efetiva transferência de conhecimento, de modo que os resultados obtidos permaneçam disponíveis à administração mesmo depois do encerramento formal da parceria.

Gestão de pessoas, trabalho e mudança organizacional

A adoção de sistemas automatizados altera de maneira concreta o trabalho cotidiano dos servidores públicos. Tarefas podem ser redistribuídas entre diferentes cargos, critérios de decisão podem mudar sem aviso claro para quem os aplica, e novas formas de monitoramento do próprio desempenho do servidor podem surgir como efeito colateral da implantação de um sistema. Jarrahi (2018) argumenta que a relação entre inteligência artificial e trabalho humano é melhor compreendida como complementar do que como substitutiva na maioria dos contextos organizacionais, mas essa complementaridade só se realiza quando a transformação é planejada com atenção ao impacto sobre as pessoas envolvidas, e não tratada apenas como um projeto de implantação de software.

Servidores da linha de frente possuem conhecimento contextual sobre o funcionamento real dos serviços públicos que raramente está integralmente registrado nos dados administrativos disponíveis. Ignorar esse conhecimento tácito reduz a qualidade final do sistema implantado e tende a aumentar a resistência dos próprios servidores à nova ferramenta. A participação ativa desses profissionais no desenho, nos testes e na avaliação contínua do sistema permite identificar exceções, ambiguidades e efeitos não antecipados pela equipe técnica responsável pelo desenvolvimento.

A capacitação oferecida aos servidores precisa ir além do simples treinamento sobre o uso da interface do sistema. Gestores e operadores diretos precisam compreender a finalidade da ferramenta, suas limitações conhecidas, os indicadores usados para avaliar seu desempenho e os procedimentos disponíveis para recurso em caso de erro. Também precisam ser capazes de reconhecer situações concretas em que a recomendação automatizada claramente não é adequada ao caso em análise. Esse treinamento precisa ser atualizado sempre que houver mudanças relevantes no funcionamento do sistema, e não oferecido apenas uma vez, no momento de sua implantação inicial.

A gestão da mudança organizacional precisa comunicar, de forma honesta, tanto os objetivos quanto os limites reais de cada projeto de automação. Promessas de substituição ampla de tarefas humanas ou de ganhos de eficiência automáticos e imediatos tendem a gerar insegurança e desconfiança entre os servidores, especialmente quando essas promessas não se confirmam na prática. A administração deveria explicar claramente como as responsabilidades serão redistribuídas com a nova ferramenta e quais mecanismos institucionais protegerão o exercício do julgamento profissional dos servidores envolvidos. Essa legitimidade construída internamente é condição necessária, embora não suficiente, para um uso responsável da tecnologia adotada.

Transparência ativa e comunicação pública

A transparência sobre sistemas automatizados utilizados pela administração precisa ser planejada com cuidado, e não tratada como uma obrigação burocrática cumprida apenas formalmente. A simples publicação de códigos-fonte ou de documentos técnicos detalhados pode ser insuficiente para garantir compreensão social real sobre o funcionamento de um sistema, enquanto a divulgação genérica de que a prefeitura "utiliza inteligência artificial" em determinada área, sem maiores detalhes, não permite qualquer controle efetivo por parte da população. Informações verdadeiramente úteis para o controle social incluem a finalidade específica do sistema, o órgão responsável por sua operação, o tipo de dados utilizados, os públicos potencialmente afetados e os canais concretos disponíveis para contestação.

A transparência necessária deve respeitar exigências legítimas de segurança, privacidade e propriedade intelectual, mas essas restrições não podem eliminar por completo a prestação de contas devida à população. O município pode, na prática, divulgar os critérios gerais utilizados por um sistema e os resultados de suas avaliações periódicas sem expor dados pessoais específicos ou vulnerabilidades técnicas exploráveis por terceiros mal-intencionados. Contratos com fornecedores devem assegurar, desde sua redação inicial, que o segredo comercial invocado pela empresa não impeça a explicação adequada das decisões públicas apoiadas por sua tecnologia.

Relatórios periódicos de desempenho podem apresentar indicadores de funcionamento, incidentes registrados, reclamações recebidas e medidas corretivas efetivamente adotadas. Deetz (2001), ao discutir a comunicação entre organizações e seus públicos de interesse, argumenta que processos comunicativos genuínos exigem espaço real para vozes divergentes, e não apenas a transmissão unidirecional de informações favoráveis à instituição

emissora; essa perspectiva se aplica diretamente à comunicação pública sobre sistemas automatizados, que deveria incluir também os problemas identificados, não apenas os resultados positivos. A publicação regular desse tipo de relatório fortalece a aprendizagem institucional e permite que conselhos municipais, imprensa local e a sociedade em geral acompanhem os resultados obtidos ao longo do tempo. A omissão deliberada de falhas tende a reduzir a confiança pública de forma ainda mais severa quando esses problemas eventualmente se tornam conhecidos por outras fontes, como reportagens jornalísticas ou denúncias de cidadãos afetados.

A comunicação sobre esses temas precisa utilizar linguagem acessível ao cidadão comum, sem jargão técnico desnecessário. Termos especializados, quando inevitáveis, devem ser explicados em linguagem simples, e exemplos concretos podem demonstrar de forma didática como o sistema efetivamente participa do processo decisório em questão. Essa transparência relevante, e não meramente formal, permite que as pessoas compreendam o real papel da tecnologia em suas vidas e identifiquem com clareza a autoridade pública responsável por cada decisão que as afeta.

Sustentabilidade financeira e custo do ciclo de vida

A sustentabilidade financeira é, também, um componente relevante da capacidade estatal para governar tecnologias digitais. Projetos de inteligência artificial envolvem custos de integração com sistemas legados, limpeza e tratamento de dados, poder computacional, licenciamento de software, segurança da informação, treinamento de pessoal e auditoria periódica. Uma avaliação baseada exclusivamente no preço de aquisição inicial tende a subestimar significativamente os compromissos financeiros futuros assumidos pela administração, podendo produzir uma forma sutil de dependência orçamentária ao longo dos anos seguintes.

O custo total do ciclo de vida de um sistema precisa ser comparado tanto ao benefício público esperado quanto às alternativas disponíveis, incluindo alternativas não tecnológicas. Soluções mais simples podem ser objetivamente mais adequadas quando oferecem maior transparência e exigem uma manutenção compatível com a capacidade real da equipe municipal disponível. A escolha automática pelo modelo tecnicamente mais sofisticado não é necessariamente a decisão mais racional do ponto de vista da gestão pública, especialmente quando os ganhos marginais de desempenho obtidos exigem uma infraestrutura desproporcional ao porte e às necessidades reais do município.

A previsão orçamentária de longo prazo precisa contemplar, desde o início do projeto, tanto os custos de atualização quanto os custos eventuais de descontinuação do sistema. Sistemas que não recebem recursos suficientes para sua manutenção contínua tendem a se degradar tecnicamente com o tempo e, ainda assim, podem permanecer em uso apenas por inércia administrativa, mesmo depois de deixarem de produzir os resultados esperados. A administração também deveria reservar recursos específicos para eventual migração de dados e substituição de fornecedor, reduzindo assim o risco de um aprisionamento tecnológico difícil de reverter no futuro.

Compras compartilhadas entre municípios e a adoção de padrões técnicos abertos podem reduzir custos individuais de aquisição, mas precisam ser avaliadas também segundo a capacidade real de operação de cada administração participante. A economia obtida na fase inicial de aquisição perde valor prático quando a solução escolhida exige suporte técnico externo permanente que a equipe municipal não consegue substituir. A sustentabilidade financeira de um projeto de inteligência artificial depende, em última análise, de um equilíbrio cuidadoso entre custo de aquisição, conhecimento técnico interno disponível e qualidade da governança estabelecida sobre o sistema.

Proposições teóricas e agenda de validação empírica

O modelo desenvolvido neste artigo permite formular proposições específicas para orientar pesquisas empíricas futuras. A primeira proposição sustenta que a adoção de inteligência artificial produzirá maior valor público quando houver alinhamento efetivo entre a capacidade estratégica e a capacidade informacional do município. Sem uma definição clara do problema público a ser resolvido e sem dados administrativos de qualidade adequada, o bom desempenho técnico de um sistema dificilmente se converte em melhoria real de política pública, ainda que os relatórios técnicos do projeto apresentem indicadores tecnicamente satisfatórios.

A segunda proposição sustenta que a capacidade regulatória modera os riscos associados tanto à dependência de fornecedores quanto à opacidade das decisões automatizadas. Municípios que desenvolvem contratos mais robustos, realizam avaliações de impacto algorítmico de forma sistemática e mantêm mecanismos de recurso efetivamente operacionais tendem a preservar maior autonomia institucional ao longo do tempo, mesmo diante de fornecedores tecnicamente dominantes em seu mercado de atuação. A terceira proposição indica que capacidade técnica e capacidade organizacional apresentam um efeito complementar entre si: equipes

tecnicamente qualificadas geram pouco resultado prático quando não possuem acesso adequado às informações necessárias, autoridade institucional reconhecida ou integração efetiva com as áreas finalísticas responsáveis pela política pública em questão.

A quarta proposição relaciona diretamente participação social e legitimidade institucional. Sistemas que afetam direitos fundamentais ou a distribuição de serviços públicos essenciais tendem a obter maior aceitação social quando os públicos potencialmente impactados participam ativamente do desenho do sistema, recebem explicações claras sobre seu funcionamento e dispõem de canais de recurso efetivos quando se sentem prejudicados. A participação social, importante ressaltar, não elimina conflitos de interesse legítimos entre diferentes grupos afetados por uma mesma política pública, mas torna os critérios utilizados e as responsabilidades institucionais envolvidas consideravelmente mais visíveis para todos os envolvidos.

Essas quatro proposições podem ser examinadas empiricamente por meio de estudos de caso aprofundados, levantamentos survey aplicados a gestores municipais e análises comparativas entre diferentes administrações locais. Pesquisas longitudinais permitiriam acompanhar a evolução das seis capacidades descritas antes e depois da implantação de sistemas específicos de inteligência artificial, oferecendo evidência mais robusta sobre relações causais do que estudos exclusivamente transversais. Estudos comparativos conduzidos em municípios de diferentes portes populacionais e capacidades fiscais poderiam identificar combinações institucionais funcionalmente equivalentes, isto é, arranjos distintos que produzem resultados semelhantes de governança, além dos limites práticos das diferentes estratégias de cooperação intermunicipal discutidas anteriormente. A agenda empírica decorrente deste modelo também deveria incluir análise de resultados distributivos entre diferentes grupos sociais, custos efetivamente incorridos e experiências relatadas tanto por servidores públicos quanto por cidadãos diretamente afetados pelos sistemas avaliados.

A construção de indicadores quantitativos a partir deste modelo exige cautela metodológica considerável. Medidas simples de presença ou ausência de determinado instrumento formal podem facilmente confundir formalização burocrática com capacidade institucional efetiva. Uma política municipal formalmente publicada não garante sua aplicação prática consistente; uma equipe formalmente designada para determinada função pode não dispor dos recursos necessários para desempenhá-la; e um canal de recurso formalmente existente pode, na prática, ser inacessível à maioria dos cidadãos que dele precisariam. A validação empírica do modelo proposto

deveria, portanto, combinar análise documental, entrevistas qualitativas em profundidade, observação direta de processos administrativos e análise de dados de desempenho efetivamente produzidos pelos sistemas avaliados.

CONSIDERAÇÕES FINAIS

Este artigo desenvolveu um modelo teórico de capacidades estatais voltado à governança responsável da inteligência artificial em governos locais. As seis capacidades propostas, estratégica, regulatória, técnica, informacional, organizacional e participativa, ajudam a explicar por que a adoção de uma mesma tecnologia produz resultados administrativos tão distintos entre municípios diferentes, e por que o simples volume de investimento em ferramentas digitais não é, por si só, um bom indicador de qualidade de governança.

A principal conclusão deste trabalho é que a maturidade digital de um governo local não corresponde à quantidade de sistemas automatizados que ele utiliza. Ela reside, isso sim, na capacidade institucional de definir finalidades claras para cada sistema adotado, manter controle efetivo sobre seu funcionamento ao longo do tempo, avaliar seus impactos reais sobre a população atendida e, quando necessário, interromper usos que se revelam inadequados. A governança responsável, tal como discutida ao longo deste artigo, combina eficiência administrativa com respeito a direitos, transparência sobre critérios utilizados e geração efetiva de valor público, sem privilegiar automaticamente nenhuma dessas dimensões em detrimento das demais.

O modelo evidencia, ainda, que as fragilidades institucionais discutidas neste artigo são fundamentalmente interdependentes entre si. A melhoria consistente da governança exige diagnóstico sistêmico, capaz de identificar onde estão os elos mais fracos em cada contexto municipal específico, e investimento institucional sustentado ao longo do tempo, e não apenas em momentos de maior visibilidade política. Comprar tecnologia avançada sem desenvolver, paralelamente, as capacidades institucionais necessárias para governá-la tende a ampliar, em vez de reduzir, a dependência estrutural do município em relação a fornecedores externos.

Como limitação central, é preciso reconhecer que este estudo é teórico e não apresenta validação empírica direta das relações propostas. Pesquisas futuras podem operacionalizar os indicadores aqui discutidos em instrumentos concretos de coleta de dados, comparar sistematicamente diferentes municípios brasileiros e acompanhar projetos específicos de

inteligência artificial ao longo de vários anos de operação. Estudos de caso aprofundados poderiam examinar, com mais detalhe do que foi possível neste artigo, contratos administrativos reais, bases de dados efetivamente utilizadas, processos de participação social conduzidos e incidentes concretamente registrados em diferentes contextos municipais.

Para a administração pública brasileira, a implicação prática central deste artigo é que a inteligência artificial deveria ser tratada como uma política de Estado sujeita a governança contínua, e não como um conjunto de projetos técnicos pontuais geridos isoladamente por cada secretaria interessada. O controle institucional sobre esses sistemas precisa existir antes de sua implantação, durante toda sua operação e também depois, quando chega o momento de avaliar se determinado sistema deveria continuar em uso, ser reformulado ou simplesmente descontinuado. Somente esse tipo de controle contínuo permite que a transformação digital dos governos locais cumpra sua promessa de maior eficiência sem comprometer direitos, ampliar desigualdades já existentes ou aprofundar a dependência de municípios em relação a atores externos ao interesse público que deveriam servir.

REFERÊNCIAS

- 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., Cows, 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>

CONTRIBUIÇÃO DO AUTOR

Robson Barbosa: conceituação, metodologia, investigação, análise formal, redação do manuscrito original e revisão e edição.

DECLARAÇÕES EDITORIAIS

Financiamento: o estudo não recebeu financiamento externo.

Conflito de interesses: o autor declara não haver conflito de interesse.

Uso de inteligência artificial: não declarado pelo autor.