IA e risco de integridade: nova arquitetura de controle público

21/06/2026

Os setores e órgãos de gestão de riscos e controle, à semelhança das demais organizações públicas e privadas em geral, atravessam uma transformação silenciosa e acelerada: a delegação, à inteligência artificial, de tudo aquilo que for meramente operacional. Triagem de dados, identificação de padrões, emissão de alertas, tarefas que antes demandavam horas de trabalho humano especializado, são hoje processadas em segundos por sistemas algorítmicos com ganhos inestimáveis de eficiência e segurança.

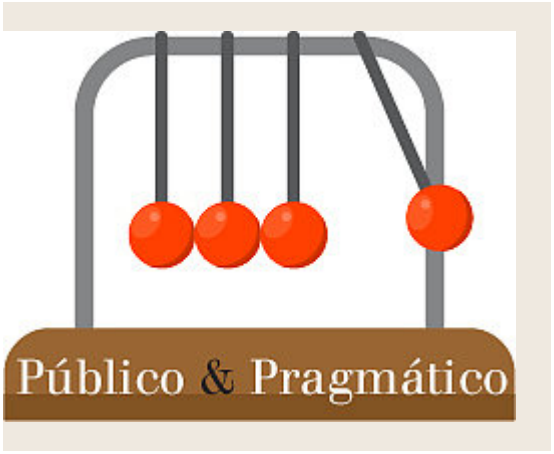
Por exemplo, o sistema ALICE (Analisador de Licitações, Contratos e Editais), desenvolvido pela Controladoria-Geral da União (CGU), que analisa diariamente, de forma automatizada, processos de compras e contratações públicas, é uma das evidências mais eloquentes dessa mudança. A ferramenta analisa milhares de processos de compras, previne inconsistências e gera economia aos cofres públicos. Diante de potenciais riscos e inconsistências, dispara alertas de indícios de irregularidades (*early warnings*) para que seja possível atuar de forma preventiva e tempestiva em processos licitatórios publicados. Atualmente, o Alice realiza a análise a partir dos dados do Portal de Compras do Governo Federal (Compras.gov.br), do Portal de Compras do Banco do Brasil (Licitacoes-e), do Portal de Compras da Caixa Econômica Federal (Licitações Caixa), além das dispensas e inexigibilidades publicadas no Diário Oficial da União (DOU) [1].

Iniciativa semelhante ocorre na China, onde a Comissão Nacional de Desenvolvimento e Reforma (NDRC) estabeleceu um roteiro (NDRC/AI Audit) para que, até o final de 2026, sistemas de inteligência artificial realizem a auditoria integral e automatizada de licitações em nível nacional. O sistema chinês foca na detecção de cartéis locais e conluios, utilizando o processamento de grandes volumes de dados para retirar dos oficiais locais o poder de manipular contratos públicos [2].

Assim, a IA acaba por redimensionar os temas de gestão de risco e controle da administração, como tem redimensionado vários outros temas da administração pública em geral (o modo de se comunicar com a cidadania, prestação de serviços ao cidadão, diagnósticos em saúde pública, planejamento educacional *etc*.). Ao ampliar exponencialmente a capacidade de monitoramento, tornou possível fiscalizar o que antes era simplesmente inviável ou custoso de acompanhar. Mas essa mesma potência coloca novas questões ao Direito Administrativo: como garantir que sistemas tão poderosos operem com transparência, responsabilização e respeito às garantias dos administrados? Como preservar a cadeia de *accountability* quando parte da gestão de risco passa a ser orientada por algoritmos?

Governança de IA

A teoria clássica do controle na administração pública parte de uma premissa simples, não substituída pelas tecnologias emergentes: o poder deve ser exercido por quem tem responsabilidade legalmente atribuída, e isso exige identificação do agente público responsável. Ou seja, saber quem decide é saber quem responde pela decisão. O sistema de *accountability* foi construído sobre essa cadeia de visibilidade [3].



Quando, por exemplo, o ALICE emite um alerta que orienta a abertura de uma auditoria, a decisão formal de abertura do procedimento continua sendo do auditor responsável. O algoritmo não decide, ele alerta, gera um *handoff* para o auditor, entrega uma informação que precisa ser avaliada e sobre a qual se impõe uma decisão humana normativamente atribuída.

Ademais, para além de identificação de agente público responsável, ainda que abastecido de informações reunidas por sistemas motorizados por IA, outro elemento se faz importante, qual seja, a vigilância sobre o risco de *algorithmic capture.*

A captura algorítmica é o risco de que sistemas automatizados de controle sejam mal calibrados, operem com vieses não detectados ou sejam manipulados para ignorar exatamente o que deveriam fiscalizar. É

aqui que a corrupção do futuro pode germinar. Por exemplo, um algoritmo treinado predominantemente em dados de irregularidades de pequeno porte pode ser menos sensível a esquemas sofisticados ainda não catalogados. Um sistema cujos parâmetros não são suficientemente auditáveis pode produzir resultados que ninguém consegue contestar, por má-fé ou mesmo por opacidade técnica [4].

A IA como vetor de risco à integridade institucional

Se, internamente, o desafio é calibrar e auditar os sistemas de gestão de risco e controle algorítmicos, externamente há uma dimensão igualmente relevante: a IA generativa já opera como instrumento ativo de ataque à integridade institucional e à segurança dos negócios (públicos e privados), e os órgãos ou setores de gestão de riscos precisam estar preparados para esse cenário.

Para ilustrar os riscos associados, em 2024, um funcionário financeiro de uma multinacional em Hong Kong autorizou quinze transferências bancárias que somaram US$ 25,6 milhões após uma videoconferência com o que lhe pareceu ser o diretor financeiro e colegas da empresa. Todos eram falsos, figuras sintéticas geradas por IA com voz e imagem clonadas de pessoas reais [5]. No mesmo período, no Reino Unido, circulou vídeo criado sinteticamente em que o primeiro-ministro Keir Starmer promovia um esquema de investimento [6]. No Brasil, da mesma forma, a Advocacia-Geral da União (AGU) notificou o TikTok para que a rede social removesse um vídeo criado sinteticamente por inteligência artificial em que o então ministro da Fazenda, Fernando Haddad, proferia informação inexistente relacionadas à política econômica [7].

Para a administração pública, essa dimensão é particularmente sensível. A confiança pública no Estado opera, em grande medida, sobre a presunção de autenticidade dos atos e comunicações institucionais. Um edital, uma instrução normativa, uma nota oficial, todos pressupõem que quem os emitiu é quem diz ser, e que o conteúdo não foi adulterado.

A IA generativa coloca essa presunção sob tensão crescente, exigindo que os sistemas de controle incorporem essa realidade em seus modelos de gestão de risco de integridade, bem como desenvolvam agilidade institucional para respostas tempestivas. Como demonstrado no caso brasileiro, a eficácia do controle em ambientes digitais hiperconectados não reside apenas na detecção do ilícito, mas na capacidade do Estado de interromper o dano em tempo quase real, impedindo que a desinformação sintética se consolide como verdade social ou gere prejuízos financeiros irreversíveis às organizações e aos cidadãos [8].

Gestão de risco, IA e o cenário regulatório: o Brasil no contexto global

As tecnologias emergem e, inevitavelmente, o mundo regulatório também se move. O AI Act europeu classifica como de alto risco os sistemas de IA usados em controle, licitação e administração pública, impondo obrigações de transparência, rastreabilidade, supervisão humana e gestão de riscos [9].

A norma europeia parte de um reconhecimento que merece atenção, qual seja, sistemas de controle automatizado concentram poder, e poder concentrado exige regulação proporcional, não para limitar sua eficácia, mas para sustentá-la no tempo.

O Brasil avança nessa direção. O PL 2.338/2023, aprovado no Senado e em tramitação na Câmara, propõe categorização de sistemas por nível de risco e obrigações de transparência algorítmica. Da mesma fora, a Resolução CNJ nº 615/2025 estabelece governança de IA no Poder Judiciário, incluindo rastreabilidade de dados, controle de versões e acesso a relatórios de auditoria. São avanços concretos que colocam o Brasil em trajetória adequada [10].

O passo seguinte, ainda em aberto, é estender essa lógica de governança especificamente aos sistemas de IA utilizados na gestão de riscos e no controle público.

Por uma nova arquitetura de gestão de riscos e controle

Limitar a adoção da IA na gestão de riscos e no controle público seria um equívoco estratégico e institucional. O desafio é construir, com a mesma seriedade com que se desenvolvem os sistemas, a arquitetura de governança que os sustente.

Essa arquitetura deve articular quatro dimensões. A primeira é a auditabilidade real dos algoritmos: os critérios que geram alertas, as trilhas de análise e os dados de treinamento precisam ser transparentes, documentados e compreensíveis para os gestores que deles dependem, para os administrados que por eles são afetados e para observadores externos. A segunda é a

responsabilização clara na cadeia humana: a orientação algorítmica não pode funcionar como escudo de irresponsabilidade, quem define os parâmetros, quem alimenta os dados, quem interpreta os alertas e toma as decisões correspondentes precisa responder pelo resultado. A terceira é a supervisão humana efetiva do funcionamento [11], com capacidade real de reconhecer eventuais insuficiências, inconsistências e *automation bias* [12] como riscos institucionais tão relevantes quanto o desvio que se quer prevenir. A quarta é a autenticação institucional robusta e capacidade de pronta resposta, ou seja, protocolos que preservem a integridade da comunicação pública diante da capacidade crescente de fabricação sintética de voz, imagem e texto, protegendo tanto os cidadãos quanto as próprias instituições e a reação em caso de ataques.

Conclusão

O ALICE aprimorou a contratação pública gerando resultados tangíveis. Esses resultados não são apenas um indicador de eficiência, é uma demonstração de que a IA, bem aplicada, pode ser um dos mais poderosos instrumentos de gestão de riscos, de proteção do patrimônio público e de fortalecimento da integridade institucional já desenvolvidos.

Ao mesmo tempo, os eventos fraudulentos mencionados iluminam as duas faces de uma mesma realidade, que a IA amplifica tanto a capacidade de proteger quanto a capacidade de atacar a integridade. O que determina o resultado será, sempre, a qualidade da governança que acompanhará essas ferramentas.

---

[1] BRASIL. Controladoria-Geral da União. Painel Alice: Análise de Licitações e Editais. Brasília: CGU, 2024. Disponível em: https://www.gov.br/cgu/pt-br/assuntos/auditoria-e-fiscalizacao/alice. Acesso em: 1 maio 2026.

[2] CHINA. National Development and Reform Commission (NDRC). Report on the Implementation of the AI Plus Initiative in Public Procurement. Pequim: NDRC, 2026. Disponível em: https://en.ndrc.gov.br/. Acesso em: 1 maio 2026.

[3] BOVENS, Mark. Analysing and Assessing Accountability: A Conceptual Framework. European Law Journal, v. 13, n. 4, p. 447–468, 2007.

[4] A metáfora da “caixa-preta algorítmica” foi consagrada por PASQUALE, Frank. The Black Box Society: The Secret Algorithms That Control Money and Information. Cambridge: Harvard University Press, 2015. No mesmo sentido, em diagnóstico empírico de falhas em sistemas algorítmicos do setor público, ver EUBANKS, Virginia. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin’s Press, 2018.

[5] KANE, Jennifer. Hong Kong firm loses $25m after staffer tricked by deepfake video call. The Guardian, 5 fev. 2024. Disponível em: https://www.theguardian.com/world/2024/feb/05/hong-kong-company-deepfake-video-conference-call-scam. Acesso em: 1 maio 2026.

[6] RESPONSIBLE AI COLLABORATIVE. Incident 607: Deepfake Video Circulating of British Labour Leader Keir Starmer Touting an Investment Scheme. AI Incident Database, 2024. Disponível em: https://incidentdatabase.ai/cite/607/. Acesso em: 1 maio 2026.

[7] TIKTOK remove vídeo fake de Haddad após notificação da AGU. CNN Brasil, São Paulo, 14 mar. 2024. Política. Disponível em: https://www.cnnbrasil.com.br/politica/tiktok-remove-video-fake-de-haddad-apos-notificacao-da-agu/. Acesso em: 1 maio 2026.

[8] PAWELEC, Maria. Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions. Digital Society, v. 1, n. 2, art. 19, 2022.

[9] Para análise crítica do AI Act europeu, ver VEALE, Michael; ZUIDERVEEN BORGESIUS, Frederik. Demystifying the Draft EU Artificial Intelligence Act: Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach. Computer Law Review International, v. 22, n. 4, p. 97–112, 2021.

[10] Sobre a abertura da caixa-preta algorítmica como exigência jurídica de transparência decisória no Direito brasileiro, ver FREITAS, Juarez; FREITAS, Thomas Bellini. Direito e Inteligência Artificial em Defesa do Humano. Belo Horizonte: Fórum, 2020. Especificamente sobre machine learning aplicado à Administração Pública brasileira, ver FIGUEIREDO, Carla Regina Bortolaz de; CABRAL, Flávio Garcia. Inteligência artificial: machine learning na

Administração Pública. International Journal of Digital Law, Curitiba, v. 1, n. 1, p. 79–96, jan./abr. 2020.

[11] GREEN, Ben. The Flaws of Policies Requiring Human Oversight of Government Algorithms. Computer Law & Security Review, v. 45, 2022.

[12] ALON-BARKAT, Saar; BUSUIOC, Madalina. Human–AI Interactions in Public Sector Decision Making: "Automation Bias" and "Selective Adherence" to Algorithmic Advice. Journal of Public Administration Research and Theory, v. 33, n. 1, p. 153–169, 2023.