

(Public) Attention is all you need: **desafios e potencialidades** **da transparência, do Big Data e da IA no** **combate à corrupção¹**

Ciro Moraes dos Reis
Guilherme France

Resumo

Neste texto, investigamos como a transparência, o *Big Data* e a Inteligência Artificial tornam-se protagonistas no combate à corrupção. Revelamos potencialidades: mostramos como dados massivos e algoritmos podem expor irregularidades escondidas no setor público e incentivar uma cidadania mais informada e participativa. Mas não fugimos dos dilemas: apontamos que, apesar das promessas, essas tecnologias trazem riscos – decisões opacas, exclusão digital, vieses algorítmicos. Analisamos o cenário brasileiro, casos concretos, e propomos perguntas inquietantes: estamos preparados para confiar a luta anticorrupção

1 Originalmente, “Attention is All You Need” é o título de um artigo fundamental em IA que introduziu um mecanismo, a “atenção”, que permite que os modelos de IA foquem nas partes mais relevantes de grandes volumes de dados para processar informações de forma eficiente. Trasladamos a metáfora para o espaço público: “atenção pública” é a capacidade coletiva de mirar criticamente os dados que o Estado expõe por transparência e políticas de dados abertos. Nesse mar de planilhas, diários oficiais e APIs (nem sempre navegáveis) ferramentas de IA e Big Data Analytics podem funcionar como bússolas automáticas, rastreando padrões, riscos e anomalias em gastos, licitações e processos, reforçando accountability e participação cidadã apesar de entraves institucionais.

a máquinas? Quais mecanismos de controle, auditoria e envolvimento social precisamos construir? Ao navegar por desafios regulatórios e institucionais, o texto convida o leitor a refletir: a era digital pode reinventar a transparência, mas cada passo precisa ser calculado – com ética, governança e audácia.

Abstract

In this text, we investigate how transparency, Big Data, and Artificial Intelligence are becoming key players in the fight against corruption. We reveal their potential: we show how massive data and algorithms can expose hidden irregularities in the public sector and encourage a more informed and participatory citizenry. But we do not shy away from the dilemmas: we point out that, despite the promises, these technologies carry risks – opaque decisions, digital exclusion, algorithmic biases. We analyze the Brazilian scenario, concrete cases, and propose unsettling questions: are we prepared to entrust the fight against corruption to machines? What control, audit, and social involvement mechanisms do we need to build? As it navigates regulatory and institutional challenges, the text invites the reader to reflect: the digital age can reinvent transparency, but each step must be calculated – with ethics, governance, and audacity.

1. Introdução

A promessa das tecnologias digitais na administração pública veio acompanhada de um paradoxo inquietante: à medida que mais dados são coletados e mais algoritmos são aplicados, menos claros se tornam os processos de decisão que moldam nossas vidas. Com o entusiasmo pela inovação vieram também novos riscos de injustiça, opacidade e exclusão digital. Como enfrentar essas tensões, garantindo que as ferramentas de inteligência artificial e *Big Data* se traduzam efetivamente em integridade, transparência e responsabilidade democrática?

Este artigo analisa como as Tecnologias Digitais Anticorrupção (ACTs), em especial *Big Data Analytics* (BDA) e iniciativas que usam algoritmos de machine learning, estão redesenhando a luta contra a corrupção no Brasil. O objetivo é refletir criticamente sobre ferramentas que, embora poderosas, podem reproduzir desigualdades se não vierem acompanhadas de governança sólida e participação cidadã efetiva.

Mais do que celebrar as conquistas tecnológicas, o texto busca fornecer uma visão realista e provocativa sobre as condições necessárias para que *Big Data* e Inteligência Artificial efetivamente promovam a democracia e não acabem reproduzindo, de forma ampliada, os mesmos problemas que deveriam resolver.

Em vez de apresentar receitas prontas, essas reflexões pretendem lançar luz sobre aspectos frequentemente negligenciados, estimulando uma agenda pública capaz de transformar retórica em mudança concreta. Na primeira seção, reconstrói-se a corrupção como fenômeno estrutural, articulando conceitos relacionados a direitos humanos, combate à corrupção e o papel das tecnologias anti-corrupção (ACTs) como arranjos sociotécnicos que combinam código, dados e poder. A segunda seção analisa o contexto regulatório brasileiro, da Constituição Federal (art. 5º e art. 37) e convenções internacionais até a LAI, LGPD, Lei do Governo Digital e o decreto da OGP, destacando avanços, tensões e lacunas na governança de dados. Na terceira seção, são apresentados estudos de caso de *Big Data* e machine learning e generative AI, nacionais e internacionais, com seus objetivos, resultados e limitações. A quarta seção aprofunda os desafios técnicos e institucionais que persistem: qualidade e interoperabilidade dos dados, vieses algorítmicos, fragmentação de responsabilidades e insegurança jurídica. Percebe-se por que as ACTs ainda não alcançam seu potencial. Por fim, a conclusão sintetiza aprendizados, identifica lacunas regulatórias e propõe reflexões para uma governança democrática e auditável de tecnologias anticorrupção, capaz de articular transparência, participação cidadã e rigor técnico.

2. Fundamentação teórica

A corrupção, no contexto deste artigo, será entendida como “o *abuso de poder confiado para ganhos privados*” (TRANSPARENCY INTERNATIONAL, 2019). É essencial perceber que essa definição não se limita às práticas mais visíveis e tipificadas como criminosas, como o peculato, o suborno e a lavagem de dinheiro. Ao contrário, ela abrange também fenômenos menos evidentes, como a chamada corrupção institucional, na qual grupos de interesse se apropriam e manipulam normas legais para favorecer objetivos privados (ODILLA, 2024).

Compreender a corrupção como fenômeno estrutural exige deslocar o foco da simples ilicitude de atos individuais para o modo como interesses privados capturam e moldam o processo de produção normativa, criando e reforçando estruturas legais que legitimam privilégios. Nesse quadro, práticas formalmente legais podem funcionar como mecanismos de captura do Estado, alinhando normas a interesses específicos em detrimento do bem público. A literatura sobre *state capture* mostra como grupos influentes se apropriam dos espaços decisórios para institucionalizar privilégios e manter barreiras ao controle social. É aí que entra o papel da transparência: ao expor o texto das normas, os bastidores das negociações e as redes de influência, a divulgação de dados deixa de ser mero procedimento burocrático e passa a atuar como instrumento de desmonte crítico dessas engrenagens. Tornar visíveis os fluxos normativos e as relações de poder eleva drasticamente o custo político da corrupção, pois submete à fiscalização pública não apenas os atos de desvio, mas o próprio processo de criação e manutenção de privilégios.

As Tecnologias Digitais Anticorrupção (ACTs) são arranjos socio-técnicos nos quais código, dados, normas e poder se entrelaçam, não mero kit de soluções neutras (ODILLA, 2024). Sem participação cidadã no desenho e na auditoria desses sistemas, o risco é trocar velhas sombras por opacidades ainda mais sofisticadas. Sua eficácia repousa no *Big Data Analytics*, capaz de extrair padrões e alertar fraudes a partir de

grandes volumes, variedades e velocidades de dados – desde notas fiscais até posts em redes sociais (HENRIQUES, 2021). Contudo, quando faltam padronização, veracidade e governança, o BDA apenas amplifica inconsistências, mostrando que a promessa analítica depende de dados íntegros e gestão orientada à precisão e integridade (KÖBIS & RAHWAN 2021).

Nesse sentido, os Dados Governamentais Abertos (Open Government Data – OGD) aparecem como uma pré-condição crítica. Regulamentado inicialmente pela Lei de Acesso à Informação (LAI, Lei nº 12.527/2011) e pela Política de Dados Abertos (Decreto nº 8.777/2016), o paradigma do OGD determina que os dados governamentais não devem ser apenas acessíveis, mas também reutilizáveis, interoperáveis e compreensíveis. Entretanto, para além da simples publicação formal dos dados, há que se garantir que estes sejam inteligíveis e semanticamente enriquecidos, o que requer uma governança técnica rigorosa ainda incipiente em diversas instâncias governamentais brasileiras (KLEIN et al., 2018). Dados abertos que não são compreendidos pela sociedade falham na sua função fundamental de democratizar a informação e fortalecer o controle público.

Em última instância, combater a corrupção por meio de tecnologias digitais vai além da eficiência operacional e instrumental; implica também manter uma postura crítica e reflexiva permanente sobre como os dados e algoritmos podem acabar reproduzindo ou até intensificando desigualdades sociais e políticas que deveriam enfrentar. Apenas uma articulação cuidadosa entre transparência, participação cidadã e rigoroso controle técnico-jurídico pode garantir que o uso de *Big Data* e IA realmente sirva ao interesse público e não a renovadas formas de dominação digital.

3. Contexto brasileiro: regulação, governança de dados e digitalização

O cenário regulatório brasileiro relacionado ao acesso à informação pública, governança digital e abertura de dados constitui um ver-

dadeiro mosaico de normas e práticas cuja interação gera tanto avanços significativos quanto profundas contradições.

A arquitetura institucional que regula o acesso e uso de dados públicos no Brasil está alicerçada em fundamentos constitucionais e internacionais de direitos humanos e anticorrupção. No plano constitucional, destacam-se o direito de acesso à informação (art. 5º, XXXIII) e o habeas data (art. 5º, LXXII), bem como o princípio da publicidade (art. 37, caput) da CF/1988. Em âmbito internacional, a Convenção Americana de Direitos Humanos (Pacto de San José, 1969) e a Convenção das Nações Unidas contra a Corrupção (UNCAC, 2003) reforçam a transparência como elemento central tanto de proteção de direitos, notadamente o direito de acesso à informação, quanto de combate à corrupção.

No nível infraconstitucional, compõem o mosaico normativo a Lei de Acesso à Informação – LAI (Lei nº 12.527/2011), que operacionaliza o direito constitucional à informação; a Lei Geral de Proteção de Dados – LGPD (Lei nº 13.709/2018), que define regras de privacidade e governança de dados pessoais; e a Lei do Governo Digital (Lei nº 14.129/2021), que estabelece padrões de interoperabilidade e serviços “digitais por padrão”, além do Decreto nº 10.160/2019, que normatiza a adesão brasileira à Parceria para Governo Aberto (OGP). Cada um desses instrumentos traz visões muitas vezes distintas sobre transparência, privacidade, compartilhamento e governança de dados, resultando em um panorama regulatório multifacetado cuja complexidade frequentemente dificulta a implementação efetiva das políticas públicas (FILGUEIRAS & LUI, 2023; CRISTÓVAM & HAHN, 2020).

A LAI, promulgada em 2011, representou um salto qualitativo no reconhecimento do direito à informação pública como pilar essencial da democracia e da accountability estatal. Inspirada nos princípios de governo aberto, estabeleceu que a transparência é regra e o sigilo é exceção, exigindo dos órgãos públicos a divulgação proativa de informações em formatos acessíveis e reutilizáveis (POSSAMAI & GONZATTI DE SOUZA, 2020). Pesquisas, porém, mostram que burocracias ainda negam infor-

mações alegando “trabalho adicional” ou divulgando arquivos fechados, prática que preserva ampla margem discricionária e compromete o acesso efetivo (ARAÚJO & SOUZA, 2011; POSSAMAI & GONZATTI DE SOUZA, 2020). Isso revela que a implementação da LAI é marcada por uma considerável margem de discricionariedade administrativa, que frequentemente opera em favor da opacidade e em detrimento da transparência efetiva.

A Lei do Governo Digital (14.129/2021) tenta impulsionar serviços “digitais por padrão” e dados em formato aberto, mas esbarra em falta de padrões técnicos, metadados precários, integração frágil entre bases e resistência cultural ao compartilhamento (CRISTÓVAM & HAHN, 2020; POSSAMAI & GONZATTI DE SOUZA, 2020). Sem APIs e identificadores comuns, a interoperabilidade (elemento central para governança de dados) segue um dos maiores gargalos da máquina pública (FRANCO, 2024).

Mais recentemente, avançaram as discussões sobre um marco regulatório específico para a Inteligência Artificial, materializadas em torno do PL 2338/2023. Sinalizam um esforço em alinhar o Brasil às melhores práticas internacionais, prevendo mecanismos como relatórios de impacto algorítmico e auditorias públicas sobre sistemas baseados em IA, com destaque à inteligência artificial generativa. Todavia, enquanto a regulação federal permanece em disputa legislativa, alguns estados já adotaram legislações próprias, criando uma fragmentação regulatória que potencialmente pode gerar assimetrias operacionais e jurídicas (FILGUEIRAS & LUI, 2023).

Válidas de nota são as iniciativas estaduais nessa frente. O estado de Alagoas inovou com a promulgação da Lei nº 9.095/2023, que estabelece parâmetros para o uso ético, transparente e eficiente de ferramentas de IA pelos órgãos da administração pública estadual visando a melhoria dos serviços prestados à população. De maneira semelhante, o Paraná promulgou a Lei nº 22.324/2025. Outra legislação de destaque vem de Goiás, cuja Lei Complementar nº 205/2025 busca fomentar a inovação com o uso de AI.

Em suma, integrar LAI, LGPD, Lei do Governo Digital e um futuro marco de IA exige superar resistências históricas, carências técnicas e a fragmentação normativa. A governança digital brasileira ainda lida com dados mal padronizados, infraestrutura limitada e baixa participação cívica, um conjunto que mantém opacidade onde se prometia transparência.

4. Exemplos de aplicações tecnológicas no combate à corrupção

O potencial das tecnologias digitais, em especial *Big Data* e Inteligência Artificial (IA), na luta contra a corrupção pode ser observado em diversos exemplos práticos, tanto no contexto brasileiro quanto internacional. Essas experiências oferecem insights importantes sobre como algoritmos podem traduzir dados abertos em ações concretas de responsabilização pública.

É importante distinguir tecnicamente entre Machine Learning (ML) e Generative AI (GenAI): ML engloba algoritmos estatísticos – como regressão, SVM, árvores de decisão e redes neurais profundas – que aprendem a mapear entradas em saídas ou agrupar dados não rotulados, exigindo engenharia manual de características e conjuntos de dados estruturados. Em contraste, GenAI dedica-se à síntese autônoma de conteúdo (texto, imagem, áudio, vídeo) por meio de arquiteturas de deep learning avançadas – sobretudo Transformers com mecanismos de atenção autoregressiva (VASWANI et al., 2017), produzindo resultados originais e criativos. A maior parte das iniciativas descritas a seguir utiliza pipelines de *Big Data* e técnicas tradicionais de ML supervisionado ou heurístico, sem incorporar os modelos fundacionais e o autoaprendizado contínuo típicos da IA generativa.

Exemplos internacionais fortalecem esse panorama. A Ucrânia revolucionou sua transparência pública com a plataforma **ProZorro**, um sistema aberto que permite monitoramento em tempo real das compras públicas, alertando automaticamente jornalistas e sociedade civil sobre

possíveis irregularidades. Na Hungria, o projeto **Red Flags** emprega algoritmos preditivos para pontuar contratos públicos com base em dezenas de indicadores de risco, enquanto no Reino Unido a plataforma **Spend Network** demonstra como a adoção de APIs abertas e identificadores únicos aumenta significativamente a transparência e facilita a investigação jornalística e cívica.

No Brasil, iniciativas como a **Operação Serenata de Amor** destacam-se pela combinação entre ativismo social e análise de dados públicos. O **bot Rosie**, desenvolvido por programadores voluntários, automatizou a auditoria da Cota Parlamentar, resultando em reembolsos indevidos identificados e devolvidos por parlamentares (ODILLA & MATTONI, 2023).

Criada em 2016 por desenvolvedores brasileiros em reação ao contexto político pós-impeachment da presidente Dilma Rousseff, a Rosie é um bot inovador baseado em algoritmos de *machine learning*. Seu objetivo central foi identificar gastos suspeitos realizados por deputados federais por meio da análise sistemática dos dados públicos disponibilizados pela Câmara dos Deputados, focando em reembolsos provenientes da cota para o exercício da atividade parlamentar (CEAP). Ela utilizava técnicas avançadas de cruzamento de dados para apontar automaticamente despesas possivelmente irregulares, tais como refeições incompatíveis com o consumo individual ou gastos realizados em locais improváveis ou distantes das atividades parlamentares. Os resultados eram divulgados via Twitter, aumentando assim a visibilidade pública das ações dos deputados e contribuindo para a pressão política e social por maior transparência.

Houve ainda uma parceria com a Câmara dos Deputados para padronizar os dados, a criação do portal Jarbas, a abertura do código no GitHub e um canal no Telegram com mais de 600 voluntários, evidenciando forte mobilização cívica. Em termos de impacto, Rosie obteve destaque ao gerar casos concretos em que deputados devolveram valores aos cofres públicos, resultado direto da pressão pública causada pela divulgação automática dos gastos questionáveis. Por exemplo, Rosie si-

nalizou casos suspeitos como deputados solicitando reembolso por refeições em um restaurante em seu estado de origem no mesmo dia e horário em que estavam votando na capital federal (ODILLA & MATTONI, 2023). Outros exemplos incluem gastos irregulares com bebidas alcoólicas e o reembolso de despesas que indicavam o consumo de 12 kg de comida em uma única refeição (ODILLA, 2023).

Entretanto, o projeto enfrentou desafios significativos, incluindo resistência institucional por parte dos parlamentares e dificuldades financeiras para manutenção e atualização tecnológica. Além disso, a dependência de plataformas externas, especialmente o Twitter, expôs vulnerabilidades que resultaram no bloqueio temporário da conta do bot, destacando riscos associados à terceirização da infraestrutura tecnológica.

Outro caso é o **Ligado nos Políticos**, primeira iniciativa brasileira de *linked data* sobre agentes eletivos, que integra bases do TSE, Senado, Câmara, Transparência Brasil e Lei da Ficha Limpa usando web crawling e screen scraping para gerar conjuntos RDF com URIs abertas (ARAÚJO & SOUZA, 2011). Ferramentas anticorrupção dependem de cruzar dados de fontes diversas: doações do TSE, contratos no ComprasNet, sanções no TCU. Se cada órgão publica planilhas diferentes, analistas gastam tempo limpando e alinhando colunas e os algoritmos aprendem menos. Ao seguir os princípios de dados abertos do W3C, a plataforma possibilita interligação, reuso e cruzamento automático com outros bancos, elevando o patamar de transparência e permitindo análises avançadas sobre atuação e financiamento político.

Apesar dessa abordagem inovadora, o projeto enfrentou limitações técnicas significativas, incluindo uma cobertura inicial restrita a políticos federais, ausência de identificadores únicos padronizados nas fontes originais e inconsistências nos formatos e atualizações dos dados. A clareza e acessibilidade dos dados publicados ajudam a elevar o padrão de transparência política e possibilitam análises mais sofisticadas por parte de pesquisadores e jornalistas.

A esfera governamental brasileira também apresenta casos interessantes. A Controladoria Geral da União implementou o sistema ALICE,

um modelo de IA voltado à fiscalização preventiva de licitações públicas, capaz de cruzar informações societárias e financeiras para identificar irregularidades com alto grau de precisão, economizando milhões de reais em potenciais danos aos cofres públicos. O sistema Alice marca um avanço institucional significativo na utilização de inteligência artificial para fiscalização de licitações e contratos públicos.

Tecnicamente, Alice é um RPA (Robotic Process Automation). Trata-se de um robô digital treinado para executar tarefas manuais e repetitivas, reduzindo os custos, a carga de trabalho e os erros dos humanos. Foi concebida para que as pessoas sejam alocadas em questões mais importantes, que exigem percepção e decisão humana (OLIVEIRA et al., 2022). Ela utiliza algoritmos que cruzam editais de licitação com bases externas de dados sobre a idoneidade e capacidade técnica das empresas concorrentes, produzindo alertas automáticos sobre possíveis riscos de corrupção ou fraude. Essa abordagem tem permitido que a CGU, com parceria do TCU, aumente consideravelmente sua eficiência na fiscalização preventiva e na identificação proativa de irregularidades em contratos públicos, mesmo diante de limitações humanas e financeiras.

Institucionalmente, a ferramenta representa um marco na modernização dos órgãos de controle, potencializando sua capacidade técnica e reforçando o compromisso com a transparência e integridade na gestão pública. Contudo, Alice ainda enfrenta desafios. Um problema primário é a grande quantidade de alertas incorretos (DE QUADROS, 2022). Há também questões relacionadas à qualidade e integração de dados de diferentes fontes governamentais, necessitando de melhorias constantes em sua arquitetura tecnológica e atualização periódica dos modelos de aprendizado.

De forma semelhante, a CGU utiliza o sistema MARA (Mapeamento de Risco de Corrupção na Administração Pública Federal), combinando aprendizado de máquina e análise de redes para detectar fraudes em convênios federais, ampliando significativamente a eficiência das auditorias realizadas pela instituição (ODILLA, 2024).

Outras iniciativas ilustram a amplitude das ACTs brasileiras: o **Victor**, desenvolvido pelo STF em parceria com a UnB, usa IA para triagem de recursos extraordinários, acelerando o fluxo judicial; o **Bem-Te-Vi**, do TST, aplica modelos preditivos à gestão de processos e à detecção de conflitos de interesse; o **Cérebro**, do CADE, minera dados de licitações para sinalizar condutas anticoncorrenciais; o **Spia**, da SSPDS-CE, emprega reconhecimento visual para localizar veículos e pessoas procuradas; e o **Zello**, do TCU, funciona como chatbot no WhatsApp, oferecendo consultas processuais automatizadas (PEREIRA & GUERRA, 2023).

Esses exemplos revelam um ecossistema variado: há ganhos claros com *Big Data* e ML, mas a adoção de IA generativa ainda é incipiente. Sempre que amparadas por governança de dados e transparência algorítmica, essas soluções podem ampliar a fiscalização e fortalecem a confiança pública. Persistem, porém, barreiras estruturais, que discutiremos a seguir.

5. Problemas e limites das aplicações tecnológicas

Apesar dos avanços mencionados, o uso de tecnologias digitais anticorrupção esbarra em diversos obstáculos estruturais e éticos. A Transparência Internacional – Brasil publicou em dezembro de 2021 o **Índice de Transformação Digital e Integridade (ITDI)**, que avaliou o nível de maturidade de iniciativas digitais de todos os governos estaduais brasileiros em quatro eixos: Transparência e Dados Abertos; Engajamento e Participação; Serviços e Contas Públicas; Segurança da Informação e Proteção de Dados. Utilizou-se de 128 indicadores para aferir práticas que fortalecem integridade e combate à corrupção. Os resultados mostraram que, apesar de acelerações pontuais durante a pandemia de Covid-19, nenhum estado alcançou o nível mais alto de maturidade, com disparidades regionais marcantes entre Sul/Sudeste e Norte/Nordeste. Desde 2022, o **Índice de Transparência e Governança Pública (ITGP)** vem aferindo o grau de abertura dos governos estaduais e do Distrito Federal; na edição inaugural, o estudo contemplou ainda

180 prefeituras, avaliadas em parceria com organizações locais, revelando avanços pontuais, mas distantes de um padrão nacional consistente de transparência.

Ambas as ferramentas instituíram um mecanismo de monitoramento, fornecendo diagnósticos precisos e recomendações em dados abertos, plataformas digitais e canais de participação. Contudo, elas também escancaram fragilidades estruturais persistentes: baixa interoperabilidade de sistemas, inconsistências de metadados, resistência cultural à governança de dados e carência de profissionais especializados, que limitam o alcance real da digitalização como instrumento eficaz de combate à corrupção no Brasil.

Neste contexto, destaca-se o problema da **qualidade dos dados**. Muitas bases públicas são disponibilizadas sem padronização ou metadados adequados, o que limita severamente a eficácia das análises automatizadas. Essa fragmentação representa um desafio técnico persistente, exigindo constante esforço para limpeza e normalização das informações antes que qualquer análise significativa possa ocorrer (CORRÊA et al., 2024). Existem bases abertas ao público em PDFs escaneados ou planilhas sem metadados, comprometendo análises sérias. A falta de metadados ou dicionários de dados que facilitem a compreensão e o uso também é um problema (MADURO-ABREU et al., 2020).

Outro problema relevante é o fenômeno da **injustiça algorítmica**, exemplificado pelo caso do **Laranjômetro** no Brasil. Desenvolvida para identificar candidaturas fictícias, essa ferramenta demonstrou um viés significativo contra mulheres negras jovens, refletindo e amplificando preconceitos históricos presentes nos dados eleitorais utilizados para treinamento dos modelos preditivos (ODILLA, 2024). Esse exemplo ressalta a necessidade urgente de auditorias éticas rigorosas e de mecanismos transparentes para mitigar vieses.

A falta de **transparência algorítmica** também constitui uma barreira significativa. Muitas soluções tecnológicas, sobretudo aquelas desenvolvidas por instituições governamentais como o TCU, não oferecem acesso público aos seus códigos-fonte e critérios decisórios internos,

criando “caixas pretas” cujo funcionamento é difícil ou impossível de verificar externamente (ODILLA, 2024). Propostas como conselhos de auditoria algorítmica pública (auditability boards) têm sido sugeridas para enfrentar esse problema, mas ainda não avançaram significativamente no contexto brasileiro.

Além disso, a **fragmentação institucional** dificulta a implementação ampla e coordenada dessas tecnologias. Cada órgão público, seja federal, estadual ou municipal, opera sistemas próprios, frequentemente incompatíveis entre si, dificultando estratégias integradas de combate à corrupção. Somado a isso, existe uma escassez crônica de recursos humanos qualificados, como cientistas de dados e analistas especializados, necessária para sustentar e desenvolver continuamente essas ferramentas (RAGAZZO et al., 2021).

A partir da perspectiva da governança, esses desafios regulatórios apontam para problemas sistêmicos mais profundos relacionados à gestão estratégica da informação pública. “Governança digital” exige mais que digitalizar rotinas: implica reformar instituições para que abertura e compartilhamento de dados se tornem pilares de uma administração democrática e responsiva. Entretanto, um forte path-dependency mantém normas e práticas legadas que travam mudanças profundas, perpetuando a ideia de que cada órgão “possui” seus dados (FILGUEIRAS & LUI, 2023). Essa cultura, somada à fragmentação administrativa, compromete a interoperabilidade e reduz o alcance de políticas de dados integrados.

Além dos impasses normativos, a governança digital brasileira esbarra num conjunto entrelaçado de barreiras: a falta de servidores com formação contínua em ciência de dados limita a limpeza e a publicação de bases interoperáveis (HENRIQUES, 2021); sistemas legados sem APIs padronizadas confinam iniciativas de IA e *Big Data* a pilotos locais (CRISTÓVAM & HAHN, 2020); e, quando os dados enfim chegam ao público, jargões técnicos e interfaces pouco intuitivas transformam portais em meros repositórios, enfraquecendo o controle social (POSSAMAI & GONZATTI DE SOUZA, 2020).

Essa combinação multifacetada de fatores torna ainda mais urgente uma abordagem estratégica e integrada, na qual investimentos técnicos sejam acompanhados por reformas institucionais profundas, capacitação continuada e fomento à participação cívica ativa, como condição essencial para que a abertura e governança digital cumpram sua promessa democrática e transformadora.

A ambiguidade normativa e regulatória, notadamente o aparente conflito entre a Lei de Acesso à Informação (LAI) e a Lei Geral de Proteção de Dados (LGPD), gera insegurança jurídica sobre a abertura e o compartilhamento de dados públicos. Isso cria um “espaço ‘cinzento’ legal” que dificulta o acesso e o compartilhamento de dados, podendo inclusive levar ao uso indevido da LGPD para restringir o acesso a informações públicas, mesmo quando não envolvem dados pessoais.

Essa ambiguidade é agravada pela lacuna regulatória específica para a Inteligência Artificial no setor público e pelo fato de que o Brasil ainda carece de um arcabouço regulatório robusto e abrangente para a área (ODILLA, 2025). Essa situação contribui para a incerteza na aplicação de tecnologias de IA no combate à corrupção, sendo necessária a criação urgente de frameworks regulatórios robustos para a governança de AI-ACTS.

Esses desafios apontam para uma conclusão crítica: embora a tecnologia represente um potencial transformador significativo, sua real efetividade depende de uma governança robusta, ética e transparente, que inclua múltiplos atores sociais e institucionais. Sem esses mecanismos de governança, o risco é transformar promessas tecnológicas em ferramentas que perpetuam desigualdades e obscuridades, ao invés de combatê-las.

6. Conclusões

Após percorrer casos, normas e impasses, fica claro que integridade digital não virá de “soluções mágicas”; exige enfrentar questões estruturais que transformem retórica de inovação em prática responsável.

A ausência de uma orientação federal clara, combinada à resistência institucional de muitos órgãos públicos em abrir e compartilhar dados de maneira consistente, impede a criação de uma governança digital efetivamente integrada e transparente em escala nacional (FRANCO, 2024).

Primeiramente, é preciso pensar na governança efetivamente pública das ACTs (ODILLA, 2025). A experiência sugere que quando métricas, logs e parâmetros permanecem internos, a confiança pública converte-se em ato de fé. Propostas legislativas que preveem relatórios de impacto algorítmico e instâncias participativas sinalizam um possível avanço, ainda que suas implicações práticas demandem debate contínuo.

Desafios de interoperabilidade também se impõem (FRANCO, 2024). Apesar de progressos em dados abertos, a indefinição sobre quem padroniza identificadores, APIs e dicionários de dados mantém bases desconectadas, o que compromete análises e amplia riscos de viés.

Em paralelo, a possibilidade de contestação pública dos sistemas merece atenção. A divulgação integral de código pode esbarrar em restrições legítimas, mas descrições técnicas detalhadas, conjuntos de teste anonimizados e auditorias independentes parecem constituir um patamar mínimo para escrutínio democrático (ARAÚJO & SOUZA, 2011).

Essas medidas, contudo, exigem investimento continuado em equipes diversas e bem formadas. A literatura indica que homogeneidade profissional tende a reproduzir preconceitos nos modelos; diversidade, portanto, é condição para justiça algorítmica, não mero adorno.

Por fim, permanece essencial uma questão fundamental de filosofia política: de quem é a responsabilidade quando sistemas algorítmicos falham? Atualmente, essa responsabilidade se perde num labirinto entre técnicos terceirizados, gestores públicos e cidadãos prejudicados.

Se negligenciarmos esses pontos, corremos o risco de ver as tecnologias automatizadas reforçando desigualdades existentes ou, na melhor hipótese, entregando uma ilusão tecnológica que confunde eficiência aparente com justiça real. Se, ao contrário, forem levadas a sério, poderemos direcionar nossa atenção coletiva (sem truques retóricos) para questões que o poder geralmente prefere manter fora do alcance do

olhar público. Transparência, então, poderá finalmente evoluir de mera retórica para uma expressão genuína e tangível de responsabilidade democrática e atenção pública.

Bibliografia

ARAÚJO, L. R., & SOUZA, J. F. Aumentando a transparência do governo por meio da transformação de dados governamentais abertos em dados ligados. **Revista Eletrônica de Sistemas de Informação**, 10(1), Artigo 7, 2011. <https://doi.org/10.5329/RESI.2011.1001007>

CORRÊA, A. H. M., OLIVEIRA, E. R. B. L., & AZEVEDO, D. C. Transparência governamental e proteção de dados: Uma análise integrada sob a perspectiva da LGPD. **Ciências Sociais Aplicadas**, 29(140), p. 1-24, 2024. <https://doi.org/10.69849/revistaft/ra10202411302127>

CONTROLADORIA-GERAL DA UNIÃO. (s.d.). **Projeto ALICE – Auditoria baseada em IA**. <https://www.gov.br/cgu/pt-br/assuntos/auditoria-e-fiscalizacao/alice>

CRISTÓVAM, J. S. da S., & HAHN, T. M. Administração pública orientada por dados: Governo aberto e infraestrutura nacional de dados abertos. **Revista de Direito Administrativo e Gestão Pública**, 1(1), p. 1-24, 2020.

DE QUADROS DANTAS, Daniela, and MARTINS, Leandro Barbosa. “Estudo Sobre a Eficácia E a Eficiência Do Uso Da Ferramenta Alice Como Fundamento Para a Prevenção E O Combate à Corrupção No âmbito Da Controladoria-Geral Da União.” **Revista Da CGU**, revista.cgu.gov.br/Cadernos_CGU/article/view/599.

FILGUEIRAS, F., & LUI, L. Designing data governance in Brazil: An institutional analysis. **Policy Design and Practice**, 6(1), p. 41-56, 2023. <https://doi.org/10.1080/25741292.2022.2065065>

FRANCO, F. S. R. Brazilian federal public administration, open data, technologies and the right to information: The main strategies for the creation of broader public values. **Brazilian Journal of Law, Technology and Innovation**, 2(1), p. 144-178, 2024.

GOMES, H. S. Como as robôs Alice, Sofia e Mônica ajudam o TCU a caçar irregularidades em licitações. G1, 2018, 31 jul. <https://g1.globo.com/economia/tecnologia/noticia/como-as-robos-alice-sofia-e-monica-ajudam-o-tcu-a-cacar-irregularidades-em-licitacoes.ghtml>

GOVERNO DO REINO UNIDO. **The Procurement Act 2023: A short guide for suppliers**, 2025. <https://www.gov.uk/government/publications/procurement-act-2023-short-guides/the-procurement-act-2023-a-short-guide-for-suppliers-html>

HENRIQUES, A. C. V. **Big data analytics para o desenvolvimento humano: Um estudo no governo federal brasileiro** [Tese de doutorado, Fundação Getulio Vargas], 2021.

K-MONITOR. (s.d.). **Red Flags** [HU/EN]. <https://k-monitor.hu/activities/20170202-red-flags-hu-en>

KLEIN, R. H., KLEIN, D. C. B., & LUCIANO, E. M. Identificação de mecanismos para ampliação da transparência em portais de dados abertos: Uma análise no contexto brasileiro. *Contextus*, 17(1), p. 203-226, 2018. <https://doi.org/10.19094/contextus.v17i1.32834>

KÖBIS, N., STARKE, C., & RAHWAN, I. Artificial intelligence as an anti-corruption tool (AI-ACT): **Potentials and pitfalls for top-down and bottom-up approaches**. arXiv, 2021. <https://arxiv.org/abs/2102.11567>

MADURO-ABREU, A., LITRE, G., SANTOS, L., SÁTIRO, G. S., ÁVILA, K., OLIVEIRA, J. E., & CASTRO, D. Transparência da informação pública no Brasil: Uma análise da acessibilidade de big data. *Revista Eletrônica de Comunicação, Informação & Inovação em Saúde*, 14(1), 2020. <https://doi.org/10.29397/reciis.v14i1.1690>

ODILLA, F. **Inside the Car Wash**: The narrative of a corruption scandal in Brazil. Trabalho apresentado na 66 Conference of the Political Studies Association, Brighton, Reino Unido, março de 2016.

ODILLA, F. Bots against corruption: Exploring the benefits and limitations of AI-based anti-corruption technology. *Crime, Law and Social Change*, 80, p. 353-396, 2023. <https://doi.org/10.1007/s10611-023-10091-0>

ODILLA, F. From concerned citizens to civic bots: The bottom-up fight against corruption in Brazil from a longitudinal perspective. In: MATTONI, A. (Ed.), **Digital media and grassroots anti-corruption** (p. 24-48). Edward Elgar, 2024a.

ODILLA, F. Unfairness in AI anti-corruption tools: Main drivers and consequences. *Minds and Machines*, 34, 2024b. <https://doi.org/10.1007/s11023-024-09688-8>

ODILLA, F. **The digitalisation of anti-corruption in Brazil**: Scandals, reforms, and innovation. Routledge, 2025.

ODILLA, F., & VELOSO, C. Citizens and their bots that sniff corruption: Using digital technology to monitor and expose politicians who misuse public money. *American Behavioral Scientist*, 2024.

OLIVEIRA, Tiago Chaves, DA ROCHA, Andre Luiz Monteiro, DE REZENDE, Matheus Scatolino. Alice: Desafios, resultados e perspectivas da ferramenta de auditoria contínua de compras públicas governamentais com uso de inteligência artificial. *Revista da CGU*, [S. l.], v. 14, n. 26, 2022. DOI: 10.36428/revistadacgu.v14i26.530. Disponível em: https://revista.cgu.gov.br/Revista_da_CGU/article/view/530. Acesso em: 1 jul. 2025.

PEREIRA, A., & GUERRA, K. N. **Projeto de uma robô “#” auxiliar da captação de doadores de sangue da Hemominas: Reflexões iniciais sobre a IA no SUS.** Anais do Congresso Brasileiro de Informática em Saúde, 1-8, agosto de 2023.

POSSAMAI, A. J., & GONZATTI DE SOUZA, V. Transparência e dados abertos governamentais: Possibilidades e desafios a partir da Lei de Acesso à Informação. **Administração Pública e Gestão Social**, 12(2), p. 1-22, 2020. <https://doi.org/10.21171/apgs.v12i2.3515>

PROZORRO. (s.d.). About. <https://prozorro.gov.ua/en/about>

RAGAZZO, C. E. J., BALERONI, M. R. C., & LEITE JÚNIOR, D. W. M. Limites ao acesso de autoridades públicas a big data: Evolução legislativa e governança regulatória. **Revista da Faculdade de Direito da UFPR**, 66 (2), p. 9-30, 2021.

RIBEIRO, H. V., LOPES, D. D., PESSA, A. A. B., MARTINS, A. F., DA CUNHA, B. R., GONÇALVES, S., PERC, M. **Deep learning criminal networks.** arXiv, 2023. <https://arxiv.org/abs/2304.08457>

SPEND NETWORK. (s.d.). **Global procurement data and analysis.** <https://www.spendnetwork.com/>

TRANSPARENCY INTERNATIONAL. **What is corruption?** 11 nov 2019. <https://www.transparency.org/en/what-is-corruption>

TRANSPARÊNCIA INTERNACIONAL – BRASIL. **Índice de Transformação Digital e Integridade avalia estados brasileiros**, maio de 2025. <https://transparenciainternacional.org.br/posts/indice-de-transformacao-digital-e-integridade-avalia-estados-brasileiros>

TRANSPARÊNCIA INTERNACIONAL – BRASIL. (s.d.). **Índice de Transparência e Governança Pública (ITGP).** <https://transparenciainternacional.org.br/itgp>

TRANSPARÊNCIA INTERNACIONAL UCRÂNIA. **How Red Flags help Hungarians detect risky procurements**, 2019. <https://ti-ukraine.org/en/news/how-red-flags-help-hungarians-detect-risky-procurements/>

TRANSPARÊNCIA INTERNACIONAL UCRÂNIA. **DOZORRO 5 years**, 2022. <https://ti-ukraine.org/en/project/dozorro-5-years/>

VASWANI, A., SHAZEER, N., PARMAR, N., USZKOREIT, J., JONES, L., GOMEZ, A. N., KAISER, Ł., & POLOSUKHIN, I. Attention is all you need. In: **Proceedings of the 31st Conference on Neural Information Processing Systems** (p. 5998-6008), 2017. Curran Associates. <https://arxiv.org/abs/1706.03762>

Ciro Moraes dos Reis é Mestre em Gestão Pública e Governança pela London School of Economics e Advogado pela Universidade Federal da Bahia, com atuação como Coordenador na Transparência Internacional – Brasil e ex-Gerente de Parcerias no TikTok. Especialista em impacto social e segurança digital, liderou iniciativas contra desinformação e discurso de ódio durante eleições brasileiras. Sua carreira multifacetada abrange governança, prática jurídica e tecnologia, combinando expertise em políticas públicas, transparência e transformação digital.

Guilherme France · Advogado e Gerente do Centro de Conhecimento Anticorrupção da Transparência Internacional Brasil. Doutorando em Ciência Política no Instituto de Estudos Sociais e Políticos da Universidade do Estado do Rio de Janeiro (UERJ). Mestre em Direito Internacional (UERJ) e Mestre em História, Política e Bens Culturais (CPDOC/FGV). Possui graduação em Direito também pela UERJ e em Relações Internacionais pela Pontifícia Universidade Católica do Rio de Janeiro.