

IA GENERATIVA E LGPD: O DESAFIO DA TRANSPARÊNCIA ALGORÍTMICA

GENERATIVE AI AND LGPD: THE CHALLENGE OF ALGORITHMIC TRANSPARENCY

INTELIGENCIA ARTIFICIAL GENERATIVA Y LGPD: EL DESAFÍO DE LA TRANSPARENCIA ALGORÍTMICA

Marise Miglioli Lorusso¹

DOI: 10.54899/dcs.v23i93.6340

Recibido: 27/07/2026 | Aceptado: 30/07/2026 | Publicación en línea: 15/08/2026.

RESUMO

O avanço exponencial da Inteligência Artificial Generativa (IAGen) reconfigurou de maneira indelével o debate sobre a proteção de dados no Brasil, evidenciando tensões estruturais entre o fomento à inovação tecnológica e as garantias consagradas na Lei Geral de Proteção de Dados Pessoais (LGPD). O presente artigo tem como objetivo analisar os obstáculos técnicos e jurídicos que dificultam a garantia do direito à revisão de decisões automatizadas (Art. 20 da LGPD) frente a sistemas que operam sob a lógica de "caixas pretas" (*black boxes*). Metodologicamente, adotou-se uma abordagem qualitativa de natureza exploratória e descritiva, fundamentada na triangulação entre análise documental, revisão integrativa da literatura e estudo de casos múltiplos, contemplando os episódios envolvendo as empresas Meta e X (Grok), a implementação do programa Smart Sampa, o uso de IA no Judiciário e a plataformização da educação pública. Os resultados demonstram a existência de um choque estrutural profundo entre a exigência jurídica analógica de transparência e a opacidade matemática inerente aos Modelos de Linguagem de Larga Escala (LLMs). Conclui-se que a adoção de um modelo fiscalizatório puramente reativo ou a transposição mimética de normas do Norte Global são insuficientes para o contexto socioeconômico brasileiro. Torna-se imprescindível a implementação de governança experimental (*sandboxes* regulatórios), a exigência de Avaliações de Impacto Algorítmico e a aplicação mandatória da metodologia *Privacy by Design* desde a concepção do código, resguardando os direitos fundamentais sem obstar o desenvolvimento tecnológico soberano.

Palavras-chave: Inteligência Artificial Generativa. LGPD. Transparência Algorítmica. Governança Digital. *Privacy by Design*.

ABSTRACT

The exponential advancement of Generative Artificial Intelligence (GenAI) has indelibly reconfigured the debate on data protection in Brazil, highlighting structural tensions between fostering technological innovation and the guarantees enshrined in the General Data Protection Law (LGPD). This article aims to analyze the technical and legal obstacles that hinder the

¹ Pós-doutora em Ciência da Informação pela Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP), Marília, São Paulo, Brasil. E-mail: marise.m.lorusso@unesp.br Orcid: <https://orcid.org/0000-0002-5755-726X>

guarantee of the right to review automated decisions (Art. 20 of the LGPD) in the face of systems operating under the logic of "black boxes." Methodologically, a qualitative, exploratory, and descriptive approach was adopted, based on the triangulation of documentary analysis, integrative literature review, and multiple case studies, encompassing the episodes involving the companies Meta and X (Grok), the implementation of the Smart Sampa program, the use of AI in the Judiciary, and the platformization of public education. The results demonstrate the existence of a deep structural clash between the analog legal requirement for transparency and the mathematical opacity inherent to Large Language Models (LLMs). It is concluded that the adoption of a purely reactive supervisory model or the mimetic transposition of norms from the Global North are insufficient for the Brazilian socioeconomic context. It becomes essential to implement experimental governance (regulatory sandboxes), require Algorithmic Impact Assessments, and mandate the application of the Privacy by Design methodology from the code's conception, safeguarding fundamental rights without hindering sovereign technological development.

Keywords: Generative Artificial Intelligence. LGPD. Algorithmic Transparency. Digital Governance. Privacy by Design.

RESUMEN

El avance exponencial de la Inteligencia Artificial Generativa (IAGen) ha reconfigurado de manera indeleble el debate sobre la protección de datos en Brasil, evidenciando tensiones estructurales entre el fomento a la innovación tecnológica y las garantías consagradas en la Ley General de Protección de Datos Personales (LGPD). El presente artículo tiene como objetivo analizar los obstáculos técnicos y jurídicos que dificultan la garantía del derecho a la revisión de decisiones automatizadas (Art. 20 de la LGPD) frente a sistemas que operan bajo la lógica de "cajas negras" (black boxes). Metodológicamente, se adoptó un enfoque cualitativo de naturaleza exploratoria y descriptiva, fundamentado en la triangulación entre análisis documental, revisión integradora de la literatura y estudio de casos múltiples, contemplando los episodios que involucran a las empresas Meta y X (Grok), la implementación del programa Smart Sampa, el uso de IA en el Poder Judicial y la plataformización de la educación pública. Los resultados demuestran la existencia de un choque estructural profundo entre la exigencia jurídica analógica de transparencia y la opacidad matemática inherente a los Modelos de Lenguaje de Gran Escala (LLMs). Se concluye que la adopción de un modelo de fiscalización puramente reactivo o la transposición mimética de normas del Norte Global son insuficientes para el contexto socioeconómico brasileño. Se hace imprescindible la implementación de gobernanza experimental (sandboxes regulatorios), la exigencia de Evaluaciones de Impacto Algorítmico y la aplicación obligatoria de la metodología Privacy by Design desde la concepción del código, salvaguardando los derechos fundamentales sin obstaculizar el desarrollo tecnológico soberano..

Palabras clave: Inteligencia Artificial Generativa. LGP. Transparencia Algorítmica. Gobernanza Digital. *Privacy by Design*.



Esta obra está bajo una [Licencia Creative Commons Atribución- NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/)

INTRODUÇÃO

A evolução tecnológica é impulsionada pelas necessidades humanas e pelas transformações das estruturas produtivas. O campo da Inteligência Artificial (IA), cujas raízes teóricas remontam aos trabalhos seminais de Alan Turing no meio do século XX e à histórica Conferência de Dartmouth em 1956, consolidou-se, nas últimas décadas, como o vetor central da Quarta Revolução Industrial (Moretti; Zuffo, 2025). A partir do final do ano de 2022, observou-se uma disrupção sem precedentes com a popularização da Inteligência Artificial Generativa (IAGen), um subcampo voltado à criação de conteúdos sintéticos inéditos, que reconfigurou de maneira profunda os debates globais sobre privacidade, direitos autorais, ética e governança algorítmica (Holderegger; Duarte, 2025; Silva; Domingues, 2024).

No Brasil, a inserção transversal dessas tecnologias tem crescido exponencialmente. Organizações dos setores público e privado adotam ferramentas generativas para otimizar processos, redigir documentos e analisar grandes volumes de dados (*Big Data*). Atualmente, estima-se que 40% das empresas brasileiras já utilizem a tecnologia de forma sistemática (Amazon Web Services, 2025; Ilia Digital, 2025). Essa rápida adoção impulsiona a produtividade e a modernização administrativa (MGI, 2025a), mas levanta preocupações estruturais e urgentes sobre os impactos ético-sociais, o risco de automação da desigualdade e a efetiva proteção dos dados pessoais dos cidadãos.

Os modelos generativos contemporâneos distanciam-se dos sistemas computacionais tradicionais, que operavam sob lógicas dedutivas baseadas em regras de programação "se-então" (*if-then*). A atual arquitetura apoia-se em Modelos de Linguagem de Larga Escala (LLMs) e em Redes Adversariais Generativas (GANs) (ANPD, 2024; Silva; Domingues, 2024). Tais sistemas codificam bilhões de parâmetros em redes neurais profundas (*Deep Learning*), operando, na prática, como "caixas pretas" (*black boxes*). Essa característica impõe severos e intrincados desafios técnicos de transparência e interpretabilidade (Santos; Chagas; Bottentuit Junior, 2024; Moretti; Zuffo, 2025).

Essa opacidade estrutural tensiona diretamente o ordenamento jurídico brasileiro, entrando em colisão frontal com a Lei Geral de Proteção de Dados Pessoais (LGPD). O artigo 20 do referido diploma legal assegura ao titular o direito de solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados, impondo ao controlador o dever de fornecer informações claras e adequadas sobre os critérios e procedimentos utilizados

(Brasil, 2018). Contudo, a opacidade inerente ao *Machine Learning* torna essa explicação humana um obstáculo de complexa execução técnica e jurídica (Silva; Domingues, 2024).

A relevância prática e a urgência dessa tensão evidenciam-se pela atuação incisiva da Autoridade Nacional de Proteção de Dados (ANPD). Recentemente, o órgão publicou a Nota Técnica nº 12/2025, buscando consolidar os entendimentos e as preocupações da sociedade em torno da transparência algorítmica e da revisão de decisões automatizadas frente à IA Generativa (ANPD, 2025). Somam-se a este cenário as medidas cautelares e investigações, a exemplo da suspensão imposta à empresa Meta (devido à coleta não transparente de dados para treinamento do modelo Llama 3) (Lima; Augustini, 2024), e as restrições à ferramenta Grok, da plataforma X, por conta da geração não consensual de imagens íntimas sintéticas (*deepfakes*) (ANPD, 2026).

Diante da necessidade imperativa de harmonizar a regulação — garantindo a proteção dos direitos fundamentais sem sufocar a inovação —, emerge a seguinte questão norteadora da pesquisa: *Quais são os principais desafios técnicos e jurídicos para a garantia do direito à revisão de decisões automatizadas (Art. 20 da LGPD) e da interpretabilidade em sistemas de IA generativa que operam como "caixas pretas" no ecossistema público e privado do Brasil?*

Para responder a essa problemática, o presente artigo tem como objetivo geral analisar, de maneira aprofundada, os principais obstáculos técnicos (derivados da opacidade estrutural dos algoritmos) e os entraves jurídicos (referentes à aplicabilidade da LGPD) que dificultam a efetivação do direito à explicação. Como objetivos específicos, delineiam-se: (i) mapear o escopo normativo da LGPD frente à IA Generativa, as diretrizes da ANPD e o conceito de *Privacy by Design*; (ii) compreender a matriz computacional das *black boxes* e os limites da *Explainable AI* (XAI); (iii) avaliar criticamente o impacto da transposição de regulações internacionais (efeito *Brussels*) no Brasil; e (iv) analisar os impactos práticos e socioeconômicos da automação mediante estudos de caso múltiplos.

REFERENCIAL TEÓRICO

O referencial teórico organiza-se em cinco dimensões inter-relacionadas: a evolução arquitetural da Inteligência Artificial; os preceitos da LGPD e a metodologia de *Privacy by Design*; o panorama regulatório internacional comparado; e os impactos socioeconômicos advindos da exclusão digital e da automação assimétrica.

Evolução Arquitetural, Aprendizado de Máquina e a Lógica das Caixas Pretas (*Black Boxes*)

A trajetória da Inteligência Artificial perpassou diversas fases evolutivas. Nas décadas passadas, o foco residia na IA Simbólica ou nos Sistemas Especialistas, os quais operavam com base em regras lógicas inseridas manualmente por programadores, permitindo total rastreabilidade das decisões (Moretti; Zuffo, 2025). Todavia, a revolução do *Machine Learning* (Aprendizado de Máquina) ocorrida a partir dos anos 1990, e consolidada com o *Deep Learning* (Aprendizado Profundo) na última década, reconfigurou esse paradigma tecnológico.

No *Machine Learning*, o sistema não recebe instruções explícitas para realizar uma tarefa; em vez disso, ele é alimentado com grandes *datasets* e utiliza algoritmos estatísticos para reconhecer padrões e inferir regras por conta própria (Silva; Domingues, 2024). O *Deep Learning* aprofunda essa mecânica ao empregar redes neurais artificiais compostas por múltiplas camadas ocultas, inspirando-se no funcionamento do córtex cerebral humano para processar informações de extrema complexidade, como o reconhecimento de fala e a visão computacional (Silva; Domingues, 2024; Moretti; Zuffo, 2025).

Subcategoria proeminente dessa evolução é a Inteligência Artificial Generativa. Segundo o Radar Tecnológico nº 3, publicado pela ANPD (2024), essas tecnologias fundamentam-se em arquiteturas como os *Transformers* (base dos LLMs, como o ChatGPT) e as Redes Adversariais Generativas (GANs). As GANs, por exemplo, utilizam duas redes neurais em competição contínua: uma rede "Geradora", que cria amostras sintéticas de dados (imagens, textos, vozes), e uma rede "Discriminadora", que tenta distinguir o conteúdo gerado do conteúdo real. Através de um processo iterativo e matemático de minimização de perdas (*loss function*), o sistema aprende a gerar resultados com alto grau de verossimilhança (ANPD, 2024; Costa *et al.*, 2024).

A inovação central desse ecossistema reside na otimização de parâmetros matemáticos (pesos sinápticos). Durante o treinamento, os dados pessoais que alimentam o modelo não são armazenados de forma bruta ou linear em um banco de dados relacional (SQL). Suas características, semânticas e relações probabilísticas são diluídas e codificadas em uma matriz multidimensional invisível (ANPD, 2024). É essa particularidade que dá origem ao conceito de "caixas pretas" (*black boxes*). Nem mesmo os engenheiros que conceberam a rede neural conseguem rastrear retrospectivamente qual dado exato determinou um *output* específico, ou como excluir cirurgicamente a influência de um único indivíduo após a consolidação do

treinamento da máquina (Moretti; Zuffo, 2025; Silva; Domingues, 2024). Essa barreira estrutural desafia diretamente os pilares clássicos do Direito Digital.

O Escopo Normativo da LGPD, o Direito à Explicação e o *Privacy by Design*

O arcabouço normativo brasileiro, capitaneado pela Lei Geral de Proteção de Dados Pessoais (Lei nº 13.709/2018), estabeleceu a autodeterminação informativa como direito fundamental. Um dos instrumentos mais sensíveis da LGPD para o controle algorítmico encontra-se no seu Artigo 20, que garante ao titular o direito de solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado que afetem seus interesses — a exemplo de perfilamento profissional, avaliações de crédito ou análises de risco de segurança pública (Brasil, 2018).

O § 1º do Artigo 20 materializa o chamado "direito à explicação", estipulando que o controlador deve fornecer informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada (Brasil, 2018). Trata-se do alicerce da transparência algorítmica no Brasil. Para garantir a propriedade intelectual, o § 2º ressalva os segredos comerciais, conferindo, no entanto, poder à ANPD para realizar auditorias caso o controlador negue as informações alegando sigilo industrial, com o fito de apurar práticas discriminatórias (Silva; Domingues, 2024).

A literatura especializada, entretanto, evidencia as limitações desse preceito diante das IAs generativas. Primeiramente, garantir a revisão não implica a reversão do resultado (Mulholland; Frajhof, 2019 *apud* Silva; Domingues, 2024). Em segundo lugar, a exigência de "critérios claros" choca-se frontalmente com a opacidade das *black boxes* descritas na seção anterior. A dificuldade técnica de aplicar a *Explainable AI* (XAI) — metodologias voltadas para traduzir o raciocínio da máquina para a compreensão humana — torna o cumprimento irrestrito do Artigo 20 quase inatingível *a posteriori* (Moretti; Zuffo, 2025).

Por essa razão, especialistas e a própria ANPD recomendam a adoção compulsória da metodologia Privacy by Design (PbD), criada nos anos 1990 por Ann Cavoukian. O PbD preceitua que a proteção de dados não deve ser um recurso corretivo, mas sim uma matriz incorporada desde a primeira linha de código (Silva; Domingues, 2024). A metodologia assenta-se em sete princípios basilares: (1) Proativo, não reativo; (2) Privacidade por padrão (*default*); (3) Privacidade embarcada no design; (4) Funcionalidade integral (positive-sum, evitando falsas

dicotomias entre segurança e usabilidade); (5) Segurança de ponta a ponta; (6) Visibilidade e transparência; e (7) Respeito à privacidade do usuário (Cavoukian, 2009 *apud* Silva; Domingues, 2024).

A adoção do PbD no desenvolvimento de IA exige práticas rigorosas de minimização e anonimização de dados. A LGPD, em seu art. 12, estipula que dados anonimizados perdem a natureza de dado pessoal, salvo se o processo puder ser revertido com esforços razoáveis (Brasil, 2018). Contudo, o avanço tecnológico constante e a proliferação de técnicas de reidentificação de Big Data ameaçam a irreversibilidade da anonimização, exigindo que os desenvolvedores atualizem continuamente seus protocolos de entropia de dados (Faleiros Júnior; Martins, 2021 *apud* Silva; Domingues, 2024). Uma solução tecnológica alinhada ao Privacy by Design é o Aprendizado Federado (Federated Learning), uma Técnica de Aprimoramento da Privacidade (PET) que treina algoritmos descentralizadamente nos dispositivos locais dos usuários, enviando ao servidor central apenas as atualizações matemáticas do modelo, sem a transferência dos dados em formato bruto (Open Loop, 2023; Silva; Domingues, 2024).

Transposição Mimética, o Efeito *Brussels* e a Regulação Comparada

O avanço acelerado da IA impôs uma revisão profunda das estratégias regulatórias globais. A governança digital tem sido dominada por dois grandes paradigmas. De um lado, os Estados Unidos adotam uma postura pró-mercado e baseada em soft law, consubstanciada no AI Risk Management Framework do NIST e em ordens executivas que priorizam a liderança tecnológica norte-americana sem imposição imediata de penalidades horizontais (Holderegger; Duarte, 2025; Serra *et al.*, 2025). De outro lado, a União Europeia aprovou, em 2024, o Artificial Intelligence Act (AI Act), a primeira regulação vinculante e horizontal baseada em níveis de risco — variando de riscos inaceitáveis (como vigilância biométrica indiscriminada) a riscos mínimos —, que prevê multas severas e a exigência de governança estruturada (União Europeia, 2024 *apud* Serra *et al.*, 2025).

A literatura descreve o fenômeno no qual padrões regulatórios europeus se irradiam para outras jurisdições como o "Efeito Brussels" (Bradford, 2020 *apud* Serra *et al.*, 2025). No Brasil, esse fenômeno é visível no que se convencionou chamar de "transposição mimética". O Projeto de Lei nº 2338/2023, sob debate no Senado e na Câmara dos Deputados, reflete diretamente a influência do AI Act europeu. O PL brasileiro adota a classificação de sistemas de IA baseada

em riscos e exige a elaboração da Avaliação de Impacto da Inteligência Artificial (AIA) e auditorias independentes para tecnologias categorizadas como de alto risco, além de preservar os direitos dos titulares e prever responsabilidade civil objetiva para danos (Câmara dos Deputados, 2025 *apud* Holderegger; Duarte, 2025; Serra *et al.*, 2025).

Contudo, acadêmicos alertam para os perigos dessa cópia regulatória sem a devida observância das assimetrias locais. O Brasil carece da maturidade institucional e do volume de investimentos europeus (Serra *et al.*, 2025). O risco da transposição mimética é a criação de obrigações regulatórias inexecutáveis para startups nacionais, asfixiando a inovação local e consolidando um oligopólio de ferramentas fornecidas exclusivamente pelas grandes corporações tecnológicas estrangeiras (Big Techs), perpetuando um quadro de submissão e dependência informacional crônica (Cavalcante, 2023; Holderegger; Duarte, 2025).

Desigualdades Socioeconômicas, Viés Algorítmico e Impactos no Trabalho

A adoção acrítica de IA no Sul Global cruza-se com fraturas históricas da sociedade brasileira. A literatura ressalta que o treinamento de algoritmos com bases de dados desatualizadas, não representativas ou predominantemente extraídas do Norte Global, tende a propagar e a automatizar discriminações pré-existentes — o chamado viés algorítmico (Holderegger; Duarte, 2025; Chen, 2023 *apud* Holderegger; Duarte, 2025). Estudos internacionais exemplificam esse fenômeno, como o caso do algoritmo de recrutamento da Amazon, que penalizou currículos de mulheres devido à base histórica masculina da empresa (Dastin, 2018 *apud* Holderegger; Duarte, 2025), ou do sistema de justiça penal norte-americano COMPAS, que superestimou o risco de reincidência para réus negros (Angwin *et al.*, 2016 *apud* Holderegger; Duarte, 2025).

Em paralelo, observa-se o profundo impacto no mercado de trabalho provocado pela automação cognitivo-generativa. Enquanto a digitalização anterior afetava trabalhos braçais, as IAs baseadas em LLMs impactam diretamente funções administrativas, analíticas e de gestão processual. Estimativas indicam que até 30% dos postos de trabalho globais poderão ser significativamente modificados até 2030 (National University, 2025 *apud* Holderegger; Duarte, 2025). Relatórios da Organização Internacional do Trabalho apontam que esse choque recairá de forma desproporcional sobre as funções de escritório predominantemente ocupadas por mulheres, exigindo políticas ativas de requalificação profissional (*reskilling*) (ILO, 2023;

Holderegger; Duarte, 2025).

Nesse contexto, o desafio torna-se bifronte: o Estado deve garantir o letramento digital da população e desenvolver infraestruturas independentes (como Nuvens de Governo Soberanas) para não sucumbir à exclusão digital, que vitima milhões de brasileiros, tornando-os alvos de exclusão sistêmica pelos próprios algoritmos encarregados da governança moderna (MGI, 2025b; CGI.br, 2025; Holderegger; Duarte, 2025).

METODOLOGIA

O presente estudo estrutura-se sob uma abordagem de cunho qualitativo, descritivo e de natureza exploratória (Holderegger; Duarte, 2025). Em face da complexidade inerente à regulação de tecnologias emergentes e da necessidade de articular o "dever ser" da esfera jurídica com a arquitetura técnica da Ciência da Computação, adotou-se o método de triangulação de dados. Para tal, a pesquisa dividiu-se em três etapas complementares: a Revisão Integrativa da Literatura, a Análise Documental e o Estudo de Casos Múltiplos.

Revisão Integrativa e Análise Documental

A pesquisa iniciou-se pelo levantamento bibliográfico e documental para compor o estado da arte e delinear o mapeamento regulatório (Holderegger; Duarte, 2025; Serra *et al.*, 2025). Analisaram-se artigos científicos nacionais e internacionais publicados entre 2018 e 2026, com foco em transparência algorítmica, LGPD, Privacy by Design e transposição mimética.

Concomitantemente, procedeu-se à Análise Documental minuciosa de normativas e orientações estatais. O corpus documental abrangeu a Constituição Federal (Brasil, 1988); a Lei Geral de Proteção de Dados (Brasil, 2018); as fundamentais Notas Técnicas da Autoridade Nacional de Proteção de Dados, em especial a Nota Técnica nº 12/2025 (sobre IA e revisão de decisões) e a Nota Técnica nº 1/2026 (sobre a IA Grok); o PL 2338/2023 (Câmara dos Deputados, 2025); relatórios metodológicos sobre Sandboxes Regulatórios expedidos pela AGU (2024); bem como relatórios demográficos e estatísticos, a exemplo da pesquisa TIC Domicílios (CGI.br, 2025).

CrITÉRIOS de Seleção e Caracterização do Estudo de Casos MÚltiplos

A etapa empírica consistiu na aplicação do Estudo de Casos Múltiplos (Holderegger; Duarte, 2025). O objetivo dessa fase foi investigar como os obstáculos da opacidade técnica identificados na teoria operam no mundo fático. A amostragem foi intencional (não probabilística), sendo selecionados eventos devidamente documentados por decisões da ANPD e pela literatura científica, divididos em dois grandes eixos: Setor Privado e Setor Público.

Eixo 1 (Setor Privado e Atuação da ANPD)

- **Caso Meta Platforms (Llama 3):** Investigou-se o bloqueio cautelar da coleta massiva de informações e a inadequação da base legal do "legítimo interesse" perante os direitos dos titulares e a proteção infantojuvenil (Lima; Augustini, 2024; Badillo, 2024).
- **Caso Plataforma X (Sistema Grok):** Analisou-se a Nota Técnica nº 1/2026 da ANPD, com foco no fenômeno dos *deepfakes* pornográficos não consensuais, na ineficácia do *Privacy by Design* por parte da plataforma e na definição regulatória de "conteúdo sintético" (ANPD, 2026).

Eixo 2 (Setor Público e a Automação do Estado)

- **Caso Smart Sampa (Prefeitura de SP):** Avaliou-se o uso de IA preditiva e reconhecimento facial pelas forças de segurança urbana, sob a ótica dos riscos de vigilância em massa, viés algorítmico e proteção constitucional (Paula, 2025).
- **Caso das Ferramentas da Administração Federal (ChatTCU, IARA, Victor):** Estudou-se a incorporação de LLMs em processos do Tribunal de Contas da União, do Conselho Administrativo de Recursos Fiscais e do Supremo Tribunal Federal, sopesando o ganho de eficiência técnica contra o imperativo do direito constitucional à motivação dos atos administrativos (Holderegger; Duarte, 2025; Moretti; Zuffo, 2025; Bittencourt, 2024; Serpro, 2024).
- **Caso da Educação Pública e Plataformização:** Abordou-se a integração de IAGen no ensino básico, focando no extrativismo de dados de menores e na infraestrutura

educacional dependente das corporações tecnológicas (Duque *et al.*, 2026; Aguirre, 2024).

Os dados advindos da teoria e dos casos foram processados mediante Análise de Conteúdo, promovendo um cruzamento interpretativo visando diagnosticar a governança atual da IA no Brasil.

RESULTADOS E DISCUSSÃO

A análise combinada do referencial teórico e dos casos selecionados corrobora a existência de um grave descompasso estrutural entre o ordenamento jurídico analógico brasileiro — que demanda transparência processual —, e a opacidade matemática arquitetural das "caixas pretas" (black boxes) inerentes aos LLMs. Os resultados foram agrupados em subseções para facilitar a compreensão dos impactos da IA Generativa na sociedade e na governança.

Oposição entre as Exigências da LGPD e a Postura das *Big Techs*: Os Casos Meta e Grok

A literatura técnica atesta a impossibilidade de apagar cirurgicamente um dado específico codificado nas matrizes da rede neural após o seu treinamento (ANPD, 2024). Esta limitação teórica materializou-se no contencioso envolvendo a Meta. O escrutínio documental revelou que a corporação modificou seus termos de privacidade tentando utilizar a base legal do "legítimo interesse" para realizar raspagem massiva (web scraping) de postagens de usuários visando ao treinamento da IA Llama 3 (Lima; Augustini, 2024; Badillo, 2024).

A intervenção cautelar da ANPD consolidou uma interpretação protetiva fundamental. Verificou-se que o "legítimo interesse" não constitui "cheque em branco" e é incabível para justificar tratamentos de dados sensíveis ou informações de crianças e adolescentes coletados indiscriminadamente da rede (Lima; Augustini, 2024; Reuso..., 2026). A ANPD utilizou o critério das "legítimas expectativas" para fundamentar a restrição: o usuário que postou conteúdo anos antes, visando a interação social, não possuía a expectativa razoável de que tal conteúdo retroativo alimentaria uma máquina comercial (Lima; Augustini, 2024). Além disso, a empresa havia criado barreiras obscuras de design (exigindo até oito etapas de cliques complexos) para dificultar o exercício do direito de oposição (opt-out), o que fere os pilares do Privacy by Design (Lima; Augustini, 2024; Cavoukian, 2009 *apud* Silva; Domingues, 2024).

O caso do sistema Grok (Plataforma X) agudiza essa discussão no campo da ética e das violações de direitos de personalidade. O modelo gerou, sem consentimento, material sintético erotizado a partir de fotos comuns, ilustrando uma falência abissal da proteção preventiva (ANPD, 2026). A Nota Técnica nº 1/2026 expedida pela ANPD representa um marco jurisprudencial: determinou-se que a geração de conteúdo sintético simulando pessoa natural configura "tratamento de dados" e o output em si é um "dado pessoal" — e, a depender do envolvimento biométrico, dado "sensível" (ANPD, 2026; Costa *et al.*, 2024). Esse entendimento afasta, definitivamente, a chamada "imunidade algorítmica" pleiteada pelo setor tecnológico, sujeitando a IAGen às sanções administrativas da Lei 13.709/2018. Restou provado que, diante da arquitetura da caixa preta que falhou em limitar "alucinações" e violações (Moretti; Zuffo, 2025), o regulador precisou usar a norma cogente para impor limitações técnicas forçadas.

A Automação das Desigualdades Estruturais: O Caso *Smart Sampa* e o Viés Algorítmico

O cenário aponta que a adoção precipitada de inteligência artificial em um país continental marcado por iniquidades crônicas pode funcionar como vetor de amplificação de discriminações (Holderegger; Duarte, 2025). O estudo do caso *Smart Sampa*, focado no uso de câmeras de segurança dotadas de IA e reconhecimento facial no município de São Paulo, demonstrou os riscos colossais associados à vigilância em massa.

Na perspectiva metodológica e social, as bases de dados oficiais carregam o histórico sociológico de superencarceramento da população negra e periférica. Ao se alimentar uma rede neural opaca com esses dados corrompidos, automatiza-se e legitima-se o preconceito racial — o viés algorítmico. Situações concretas, como a identificação errônea e o constrangimento de uma mulher negra em Sergipe, presa por uma correlação falha do sistema, atestam a materialidade desse perigo (Dataprivacy BR Research, 2023 *apud* Holderegger; Duarte, 2025). A adoção dessas ferramentas de segurança sem a existência prévia de metodologias rigorosas de Explainable AI (XAI) contraria o direito de defesa do cidadão perante eventuais "falsos positivos", demonstrando que a eficiência probabilística da máquina não pode se sobrepor às garantias constitucionais da dignidade da pessoa humana e do devido processo legal (Paula, 2025; Moretti; Zuffo, 2025).

O Paradoxo da Eficiência no Judiciário, no Executivo e na Educação

No setor público, a busca pela eficiência impulsionou o desenvolvimento de robôs baseados em LLMs, a exemplo do Projeto Victor no STF, do ChatTCU para processos de controle externo, e do sistema IARA do CARF para análises tributárias (Moretti; Zuffo, 2025; Bittencourt, 2024; Holderegger; Duarte, 2025). Embora tais ferramentas sejam projetadas para reduzir o tempo de tramitação e desburocratizar as instâncias estatais (Serpro, 2024), elas levantam um grave alerta à luz do Art. 20 da LGPD e do preceito de motivação da administração pública. Se uma decisão tributária ou punitiva for inferida por um algoritmo opaco, o Estado perde a capacidade de explicar, motivar e fundamentar a sua atuação para o administrado. Isso reforça a obrigatoriedade da arquitetura Human-in-the-loop, onde a IA deve atuar exclusivamente como co-piloto, jamais desprovida da curadoria e da supervisão final do agente estatal competente (Moretti; Zuffo, 2025; Holderegger; Duarte, 2025).

Um impacto análogo verifica-se na área social, com a crescente plataformização da educação pública, impulsionada pelas medidas emergenciais adotadas no contexto pós-pandêmico. A inserção de ferramentas de aprendizagem mediadas por IAGen e administradas por Big Techs estrangeiras coloca dados sensíveis de menores de idade em ambientes de assimetria contratual. Pesquisas apontam que o consentimento obtido em ambiente escolar frequentemente carece de validade substancial e os contratos de infraestrutura geram opacidade informacional crônica, exigindo uma direção jurídico-administrativa vigorosa para evitar o extrativismo de dados de crianças e adolescentes (Duque *et al.*, 2026; Aguirre, 2024). Em face das realidades pedagógicas, urge promover processos que incluam "Insubordinação Criativa", preparando criticamente os professores e os discentes não apenas como consumidores, mas como pensadores críticos em face das alucinações e da ética das IA (Aguirre, 2024; Santos *et al.*, 2024).

Exclusão Digital, Mercado de Trabalho e o Déficit de Infraestrutura

Os resultados empíricos da pesquisa explicitam que a aplicação universal das exigências da LGPD esbarra na realidade nua e crua da exclusão digital brasileira. O mapeamento demográfico confirma que, conquanto exista uma rápida penetração de smartphones, apenas 22% dos brasileiros possuem "condições satisfatórias de conectividade", restando 29 milhões de indivíduos absolutamente "não usuários" da internet (CGI.br, 2025 *apud* Holderegger; Duarte,

2025). Esse contingente, majoritariamente composto por cidadãos negros e integrantes das classes de menor renda, sofre de dupla exclusão: ao não gerar dados no ambiente digital, tornam-se silenciados nos treinamentos dos algoritmos ou são parametrizados apenas via estereótipos estatísticos limitados (Holderegger; Duarte, 2025; Chen, 2023 *apud* Holderegger; Duarte, 2025). Em consequência, a deficiência de conectividade impõe obstáculos severos à universalização de inovações benéficas, como a telemedicina para populações remotas ou a adoção plena da indústria 4.0 (Cardoso *et al.*, 2017; Kubota; Rosa, 2024; IBGE, 2025).

Simultaneamente, constata-se a assimetria internacional. Embora concentre alto volume de consumo corporativo de IAGen, o Brasil amarga enorme defasagem em investimento no setor e em soberania de infraestrutura (Penso, 2025; Holderegger; Duarte, 2025). Essa dependência expõe os dados sensíveis dos cidadãos aos modelos comerciais do Norte Global e a normativas de jurisdição externa (como o CLOUD Act norte-americano). Ademais, com previsões de que grande parcela dos empregos administrativos sofrerão drástica metamorfose nos próximos anos, a política pública deve antever impactos severos e oferecer reskilling aos profissionais afetados (National University, 2025; Pwc Brasil, 2025; ILO, 2023).

Recomendações e Diretrizes de Governança para o Brasil

O conjunto de dados coletados e analisados conduz à ineficácia do emprego passivo da regulação em países emergentes. Para transpor o choque entre o Direito Analógico e a Caixa Preta Matemática, propõe-se, à luz da literatura investigada (AGU, 2024; Oliveira *et al.*, 2025; Zanatta; Rielli, 2024; Holderegger; Duarte, 2025), a adoção de um quadro sistêmico estruturado nos seguintes eixos para a governança brasileira:

- **Governança Experimental e Sandboxes Regulatórios:** Em vez de sufocar o desenvolvimento pré-concebendo exigências punitivas, a ANPD deve alargar a implantação de Sandboxes Regulatórios. Tais ambientes de testes supervisionados permitem que startups e agentes públicos experimentem a implementação prática do Privacy by Design e de Tecnologias de Aprimoramento da Privacidade (PETs) no ambiente real, consolidando regulações factíveis baseadas em dados empíricos antes da comercialização em larga escala do produto (Oliveira *et al.*, 2025; AGU, 2024; Polonetsky *et al.*, 2024).

- **Aplicação Mandatória de Avaliação de Impacto Algorítmico (AIA):** Especialmente para sistemas classificados como de alto risco no vindouro arcabouço normativo (PL 2338/2023), deve-se exigir a realização da AIA e de auditorias independentes antes e durante o período de funcionamento dos modelos (Holderegger; Duarte, 2025).
- **Fortalecimento da Infraestrutura Soberana de Nuvem:** Incentivo massivo e direcionamento orçamentário à consolidação da "Nuvem de Governo" e ao fomento a LLMs nativos, baseados na cultura linguística e social local, resguardando o controle jurisdicional, a privacidade informacional de setores críticos (educação e judiciário) e estimulando o talento tecnológico interno (MGI, 2025b).
- **Mitigação dos Impactos Laborais e Inclusão Digital:** Políticas transversais voltadas à alfabetização algorítmica da sociedade civil, garantindo que o cidadão possua meios de identificar *deepfakes* e decisões abusivas, e programas robustos de requalificação voltados para os profissionais ameaçados pela substituição laboral decorrente da inteligência artificial (Holderegger; Duarte, 2025).

CONCLUSÃO

O objetivo diretor deste artigo científico consistiu em analisar os obstáculos técnicos e jurídicos que inviabilizam a plena garantia do direito à revisão e explicação de decisões automatizadas, previstas no Artigo 20 da LGPD, ante a adoção disseminada de sistemas de Inteligência Artificial Generativa calcados em arquiteturas de "caixa preta" (*black box*). Demonstrou-se que os propósitos alinhavados foram plenamente atingidos por intermédio do exame metódico das fontes e dos casos empíricos expostos.

Identificou-se que a atual modelagem computacional — consubstanciada em Redes Neurais Profundas, LLMs e GANs —, oblitera a transparência algorítmica ao fragmentar a influência dos dados em parâmetros matemáticos multidimensionais irreversíveis, fato que desafia visceralmente os pressupostos analógicos da prestação de contas exigida pela lei civil. O estudo dos casos Meta Platforms e X (Grok) comprovou que o setor corporativo global tenta explorar essas "caixas pretas" para neutralizar a autodeterminação informativa dos titulares, valendo-se abusivamente da base de "legítimo interesse" ou desconsiderando preceitos primários da engenharia protetiva. A célere intervenção cautelar da ANPD (nas Notas Técnicas nº 12/2025 e 1/2026), expandindo o escopo sancionador sobre o conteúdo sintético e exigindo barreiras reais

de *opt-out*, confirmou que a "imunidade algorítmica" não encontra esteio no marco legal pátrio.

Simultaneamente, os casos envolvendo a máquina estatal — o sistema de vigilância *Smart Sampa*, as ferramentas preditivas do Executivo e Judiciário (IARA, ChatTCU) e a plataforma digital da educação básica —, evidenciaram o agudo paradoxo das economias emergentes. Embora propulsora de ganhos de eficiência administrativa e operacionais inegáveis, a importação cega da IA para o cenário brasileiro encontra um tecido social fraturado. Ao retroalimentar a máquina com bases não curadas de um país onde cerca de 29 milhões de cidadãos sofrem com a exclusão digital, a automação tende a atuar como vetor profilático da discriminação, solidificando o viés algorítmico, inviabilizando a defesa humana (*human-in-the-loop*) e precarizando direitos.

A síntese conclusiva ressalta, por fim, que o Brasil situa-se em uma encruzilhada regulatória. A mera transposição mimética do *AI Act* europeu — em um cenário sem idêntica força orçamentária ou capacidade de conformidade — pode afogar a inovação das pequenas empresas locais, sedimentando uma dependência externa insustentável. As respostas para pacificar o conflito entre o sigilo da matemática generativa e a obrigação de clareza do Art. 20 da LGPD não emergem do passivismo reativo. Assevera-se imprescindível a adoção contundente da *Governança Experimental* ancorada em amplos *Sandboxes Regulatórios*, aliada ao caráter compulsório do *Privacy by Design* formulado em cada arquitetura digital antes mesmo de seu lançamento de mercado, somados ao desenvolvimento maciço de infraestrutura de nuvem nacional. Apenas por essa via institucional proativa, multifacetada e ancorada nos direitos humanos o país alcançará a soberania tecnológica indispensável e protegerá o cidadão frente aos imensos potenciais e riscos inerentes ao futuro da Inteligência Artificial.

REFERÊNCIAS

A GOVERNANÇA da Inteligência Artificial Generativa no Brasil: Análise Sistemática da Literatura e Atuação da ANPD (2023-2026). [S. l.: s. n.], 2026.

AGU - ADVOCACIA-GERAL DA UNIÃO. **Regulatory Sandbox Reference Guide**. Brasília: AGU, 2024.

AGUIRRE, U. J. C. Possibilidades entre a Educação Matemática e Inteligência Artificial Generativa (IAG) em sala de aula. **Seminário Internacional de Pesquisa em Educação Matemática**, p. 1-12, 2024.

AMAZON WEB SERVICES. **Pesquisa sobre adoção de IA na América Latina**. AWS, 2025.

ANPD - AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Nota Técnica nº 1/2026/FIS/CGF/ANPD**. Processo nº 00261.000178/2026-27. Brasília: ANPD, 2026.

ANPD - AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Nota Técnica nº 12/2025/CON1/CGN/ANPD**: Consolidação das contribuições recebidas na Tomada de Subsídios - Inteligência Artificial e Revisão de Decisões Automatizadas. Brasília: ANPD, 2025.

BADILLO, M. **Processing of Personal Data for AI Training in Brazil**: Takeaways from ANPD's preliminary decisions in the Meta Case. Future of Privacy Forum (FPF), 2024.

BITTENCOURT, R. O. **Aplicações e limitações de inteligência artificial generativa no direito**: um estudo de caso do Tribunal de Contas da União. 2024. Monografia Acadêmica – Universidade de Brasília, Brasília, 2024.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, 1988.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 2018.

BRASIL. Ministério da Fazenda. **Receita desenvolve ferramenta inovadora capaz de ampliar detecção de fraudes tributárias e aduaneiras**. Brasília: Ministério da Fazenda, 2024a.

BRASIL. Tribunal Superior Eleitoral. **Resolução nº 23.732, de 27 de fevereiro de 2024**. Brasília: TSE, 2024b.

CARDOSO, W. *et al.* Digital manufacturing, industry 4.0, cloud computing and thing internet: Brazilian contextualization and reality. **Independent Journal of Management & Production**, v. 8, n. 2, p. 459, 2017.

CAVALCANTE, A. P. Regulación brasileña de la inteligencia artificial. **Revista de la Facultad de Derecho de México**, v. 73, n. 287, p. 5-28, 2023.

CGI.BR. **Pesquisa TIC Domicílios 2024**. Comitê Gestor da Internet no Brasil, 2025.

CHEN, Z. Ethics and discrimination in artificial intelligence-enabled recruitment practices. **Humanities and Social Sciences Communications**, v. 10, n. 1, p. 1-16, 2023.

CNTE - CONFEDERAÇÃO NACIONAL DOS TRABALHADORES EM EDUCAÇÃO. **Desigualdade no acesso à internet impacta qualidade da educação**. CNTE Notícias, 2023.

COSTA, A. F. J. *et al.* **Radar Tecnológico nº 3: Inteligência Artificial Generativa**. Brasília: Autoridade Nacional de Proteção de Dados (ANPD), 2024.

DATAPRIVACY BR RESEARCH. **Confundida com criminoso por “Inteligência Artificial” denuncia racismo**: “Discriminada por ser pobre e preta”. DataPrivacy BR, 2023.

DEMAREST. **Artificial intelligence reignites debates in the Brazilian House of Representatives**. Demarest Advogados, 2024.

DFRLAB; NETLAB UFRJ. **A regulação do uso de IA para as eleições brasileiras. O que está em jogo?** NetLab UFRJ, 2024.

DUQUE, R. de C. S. *et al.* Governança algorítmica na escola pública: desafios da LGPD frente à IA generativa e plataformas de aprendizagem. **Revista DCS**, v. 23, n. 88, e4896, 2026.

ECONOMIST IMPACT; JPMORGAN. **The great leap**: Harnessing genAI to revolutionize Latin America's service economy. J.P. Morgan Private Bank, 2024.

EGOV-DF. **Inteligência Artificial no serviço público**. Brasília: Escola de Governo do Distrito Federal, 2025.

ESGINSIDE. **Desigualdade digital e o desafio da conectividade no Brasil**. ESG Inside, 2025.

HOLDEREGGER, R.; DUARTE, L. F. de A. Artificial Intelligence and Socioeconomic Inequalities: Algorithmic bias, labor market impacts, and governance agendas for emerging economies. **New Science** (Impacto Científico Editora), 2025.

IBGE - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Censo Demográfico 2022: Quilombolas** — Primeiros resultados. Rio de Janeiro: IBGE, 2025.

IBGE - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Pesquisa Nacional por Amostra de Domicílios Contínua — PNAD Contínua**. Rio de Janeiro: IBGE, 2023.

ILIA DIGITAL. **Inteligência Artificial no Brasil: 9 Milhões de Empresas já Adotaram IA em 2025**. Ilia, 2025.

ILO - INTERNATIONAL LABOUR ORGANIZATION. **Generative AI and Jobs**: A global analysis of potential effects on the quantity and quality of jobs. Genebra: ILO, 2023.

KASZNAR LEONARDOS. **ANPD Suspends the Processing of Data for AI Training by Meta in Brazil**. Kasznar Leonardos, 2024.

KUBOTA, L. C.; ROSA, M. B. Inteligência artificial no brasil: Adoção, produção científica e regulamentação. In: KUBOTA, L. C. (Ed.). **Digitalização e tecnologias da informação e comunicação**: Oportunidades e desafios para o Brasil. Brasília: Ipea, 2024. p. 9-32.

LIMA, C. R. P. de; AUGUSTINI, T. O caso ANPD v. Meta: análise do voto que suspendeu o tratamento de dados dos usuários para treinamento de IA generativa. **Migalhas**, 2024.

MACHER TECNOLOGIA. **ANPD, Inteligência Artificial e LGPD**: O que a Nota Técnica nº 12/2025 revela para as empresas. Macher Tecnologia, 2025.

MATTOS FILHO. **Data Privacy and Protection Day**: an overview of the Brazilian scenario. Mattos Filho Advogados, 2025.

MGI - MINISTÉRIO DA GESTÃO E DA INOVAÇÃO EM SERVIÇOS PÚBLICOS. **Estratégia de Inteligência Artificial para o Setor Público Federal**. Brasília: MGI, 2025a.

MGI - MINISTÉRIO DA GESTÃO E DA INOVAÇÃO EM SERVIÇOS PÚBLICOS. **Na Soberania Digital: Brasil consolida a Nuvem de Governo como passo para a independência e a proteção de dados**. Brasília: MGI, 2025b.

MORETTI, J. L.; ZUFFO, M. M. Artificial Intelligence in Law: Utilisation by Brazilian Legal Practitioners and Regulatory Challenges. **Beijing Law Review**, v. 16, n. 1, p. 331-352, 2025.

NATIONAL UNIVERSITY. **59 AI Job Statistics: Future of U.S. Jobs**. National University, 2025.

OBSERVATÓRIO DA IMPRENSA. **Acesso à internet no Brasil reproduz desigualdades e exige políticas públicas**. Observatório da Imprensa, 2024.

OLIVEIRA, C. G. *et al.* Brazil's First AI Regulatory Sandbox: Towards Responsible Innovation. In: **Information Society 2025**, Ljubljana, Slovenia, 2025.

OPEN LOOP. **Privacy Enhancing Technologies (PETs)**. Open Loop, 2023.

PAULA, L. C. de. Smart Sampa: A Importante Ferramenta de Tecnologia e Inteligência Artificial na Modernização da Segurança Pública. **Revista Científica da Polícia Militar de Mato Grosso do Sul - RevPMMS**, v. 2, n. 2, p. 188–203, 2025.

PENSO. **Soberania de Dados, CLOUD Act e Riscos Jurídicos: O alerta que o Brasil precisa ouvir**. Penso Consultoria, 2025.

POLONETSKY, J. *et al.* **Balancing Innovation and Oversight: Regulatory Sandboxes as a Tool for AI Governance**. Future of Privacy Forum (FPF), 2024.

PWC BRASIL. **Casos públicos de uso de IA no setor de saúde**. PwC, 2025.

REUSO de Dados e Direito Digital: Limites Regulatórios da Raspagem de Dados e da Inteligência Artificial Generativa no Brasil. [S. l.: s. n.], 2026.

SANTOS, E.; CHAGAS, A.; BOTTENTUIT JUNIOR, J. (Org.). **ChatGPT e educação na cibercultura: Fundamentos e primeiras aproximações com inteligência artificial**. v. 1. São Luís: EDUFMA, 2024.

SERPRO - SERVIÇO FEDERAL DE PROCESSAMENTO DE DADOS. **Inteligência Artificial desenvolvida pelo Serpro vai transformar o Carf**. Brasília: Serpro, 2024.

SERRA, F. A. R.; SCAFUTO, I. C.; VIEIRA, P. S. Regulation of Artificial Intelligence: Between the switch and innovation - Is Brazil prepared?. **BAR - Brazilian Administration Review**, v. 22, n. 4, e250015, 2025.

SILVA, T. D. da; DOMINGUES, P. M. Inteligência Artificial e Privacidade: Os desafios do

Privacy by Design. **Fronteiras da Representação do Conhecimento (ARC)**, v. 4, n. 2, 2024.

ZANATTA, R.; RIELLI, M. **Mapping the Future of AI Regulation in Latin America: Civil Society Perspectives on Brazil's Pilot AI Regulatory Sandbox**. ResearchGate, 2024.