

2

Big Data e democracia no Brasil



**Tecnologias digitais,
dados e democracia**
Leonardo Paz Neves

**O uso de Big Data e da
Inteligência Artificial Generativa nas
campanhas eleitorais**
Joscimar Souza Silva
Isabela Rocha

Lógicas que excluem
Patrícia Rangel
Jéssica Melo Rivetti
Flávia Rios

**Big Data, desinformação
e governos**
Luciana Santana
Wendel Palhares

**Big Data: os desafios à
regulamentação do uso de dados**
Eleonora Mesquita Ceia
Pedro Teixeira Gueiros

**A transparência governamental
no uso de Big Data**
Maria Paula Almada
Isabele B. Mitozo

**Desafios e potencialidades da
transparência, do Big Data e da IA
no combate à corrupção**
Ciro Moraes dos Reis
Guilherme France

Segurança pública na era do Big Data
Daniel Edler

**Big Data, democracia participativa
e o papel da Inteligência Artificial**
Filipe Medon

Democracia movida a dados
João Victor Archegas
Diana de Souza Fernandes

Big Data e democracia no Brasil

Cadernos **2**

ANO XXVI
2025

Adenauer

Big Data e democracia no Brasil

EDITOR RESPONSÁVEL
Maximilian Hedrich

CONSELHO EDITORIAL
Antônio Jorge Ramalho
Estevão de Rezende Martins
Fátima Anastasia
Humberto Dantas
José Mario Brasiliense Carneiro
Leonardo Nemer Caldeira Brant
Lúcia Avelar
Mario Monzoni
Rodrigo Perpétuo
Silvana Krause

COORDENAÇÃO EDITORIAL
Reinaldo J. Themoteo

REVISÃO
Reinaldo J. Themoteo

CAPA, PROJETO GRÁFICO E DIAGRAMAÇÃO
Claudia Mendes

ISSN 1519-0951

Cadernos Adenauer xxvi (2025), nº2

Big Data e democracia no Brasil

Rio de Janeiro: Fundação Konrad Adenauer, outubro 2025.

ISBN 978-65-89432-60-9

As opiniões externadas nesta publicação são de exclusiva responsabilidade de seus autores e não necessariamente representam as opiniões da Fundação Konrad Adenauer.

Todos os direitos desta edição reservados à

FUNDAÇÃO KONRAD ADENAUER
Representação no Brasil: Rua Guilhermina Guinle, 163 · Botafogo
Rio de Janeiro · RJ · 22270-060
Tel.: 0055-21-2220-5441 · Telefax: 0055-21-2220-5448
adenauer-brasil@kas.de · www.kas.de/brasil

Sumário

7 Apresentação

11 **Tecnologias digitais, dados e democracia: um breve debate sobre o impacto das tecnologias manipulativas nos processos eleitorais**

Leonardo Paz Neves

31 **O uso de Big Data e da Inteligência Artificial Generativa nas campanhas eleitorais**

Joscimar Souza Silva

Isabela Rocha

49 **Lógicas que excluem: inteligência artificial, gênero e os limites da democracia**

Patricia Rangel

Jéssica Melo Rivetti

Flávia Rios

67 **Big Data, desinformação e governos: desafios para garantir a integridade da Informação**

Luciana Santana

Wendel Palhares

81 **Big Data: os desafios à regulamentação do uso de dados**

Eleonora Mesquita Ceia

Pedro Teixeira Gueiros

101 **A transparência governamental no uso de Big Data**

Maria Paula Almada
Isabele B. Mitozo

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

Ciro Moraes dos Reis
Guilherme France

137 **Segurança pública na era do Big Data: policiamento algorítmico, suas limitações operacionais e seus dilemas éticos**

Daniel Edler

157 **Big Data, democracia participativa e o papel da Inteligência Artificial**

Filipe Medon

179 **Democracia movida a dados: valores democráticos e regulação na era da inteligência artificial**

João Victor Archegas
Diana de Souza Fernandes

Apresentação

Os avanços da tecnologia de telecomunicações junto com os ocorridos na informática engendraram uma verdadeira revolução tecnológica que se deu sobretudo a partir da segunda metade do século passado. Essa revolução foi denominada por Manuel Castells como sociedade em rede, e vem mudando de forma profunda os mais variados setores da vida humana, incluindo lazer, mobilidade, trabalho, e entre outros contextos, destacamos aqui a vida política. Os referidos avanços tecnológicos possibilitaram lidarmos, auxiliados por ferramentas de inteligência artificial, com volumes assombrosos de dados, de forma que poucas décadas atrás seria impensável. Conhecido como *Big Data*, o processamento e análise de grandes volumes de dados tem moldado o mundo em que vivemos, presente em todas as áreas. Contudo, assim como toda nova tecnologia, esta traz oportunidades mas também apresenta diversos desafios que precisam ser adequadamente enfrentados, a fim de que sociedade civil, governos e empresas sejam capazes de aproveitar ao máximo o que tais recursos podem oferecer, sendo capazes também de lidar com os potenciais problemas que a tecnologia acarreta.

O objetivo desta edição da série Cadernos Adenauer, a segunda de 2025, é apresentar uma série de artigos que analisem alguns dos principais aspectos do Big Data, no que concerne aos seus impactos na demo-

cracia brasileira, ofertando aos leitores textos que ajudem a entender o que é o *Big Data*, seus impactos e desafios no contexto da democracia brasileira, bem como tal tecnologia pode contribuir para o fortalecimento da democracia.

Esta publicação propõe uma análise sobre o fenômeno do *Big Data* e seus múltiplos impactos na sociedade contemporânea, e é composta por dez capítulos. O primeiro capítulo introduz a revolução dos dados, explorando como o volume, a velocidade e a variedade de informações transformam setores como saúde, educação, segurança e economia. Em seguida, investiga-se o uso estratégico do *Big Data* em campanhas eleitorais, revelando como algoritmos e perfis digitais influenciam decisões políticas e moldam o comportamento do eleitor. O terceiro capítulo aborda os riscos de viés e discriminação algorítmica, especialmente em relação ao gênero, discutindo como dados enviesados podem perpetuar desigualdades sociais. A obra também dedica espaço à questão da desinformação, examinando como sistemas automatizados podem amplificar fake news e comprometer o debate público em seu quarto capítulo. No campo jurídico e ético, o quinto capítulo traz uma análise acerca dos desafios da regulamentação do uso de dados, destacando a necessidade do desenvolvimento de políticas públicas que conciliem inovação com proteção dos direitos individuais. A transparência governamental é discutida em dois capítulos, que tratam da importância de tornar acessíveis os dados utilizados pelo Estado e de garantir que seu uso seja legítimo e auditável. A privacidade e a vigilância são temas centrais, com reflexões sobre os riscos da coleta massiva de dados e os limites entre segurança e liberdade individual, constituem o tema do oitavo capítulo. No nono capítulo é discutido o papel do *Big Data* na democracia participativa, mostrando como plataformas digitais podem ampliar o engajamento cívico e a escuta ativa da população. Por fim, o conceito de democracia data-driven é explorado como uma interseção entre inovação tecnológica e valores democráticos, questionando até que ponto a automação pode coexistir com princípios como transparência, equidade e representatividade.

Agradecemos a cada autora e a cada autor que participam nesta publicação, cujos artigos contribuem para o debate democrático. Cada capítulo convida leitoras e leitores a refletir criticamente sobre o poder dos dados e a responsabilidade coletiva em sua gestão, propondo caminhos para uma sociedade mais informada, justa e consciente. Boa leitura!

MAXIMILIAN HEDRICH

Diretor da Fundação Konrad Adenauer no Brasil

Tecnologias digitais, dados e democracia: um breve debate sobre o impacto das tecnologias manipulativas nos processos eleitorais

Leonardo Paz Neves

Resumo

O presente texto busca analisar como a explosão de dados e a expansão das infraestruturas digitais, combinadas à tecnologias como IA e a algoritmos, reconfiguram a política contemporânea. Ancorado em conceitos como “capitalismo de vigilância” de Zuboff, o texto sustenta que dados comportamentais se convertem em ativos para perfilar e induzir condutas no ambiente democrático, deslocando a disputa eleitoral para ambientes opacos de microtargeting, desinformação e uso massivo de bots. Revisitando a teoria democrática, o texto contrapõe eleições à limitação do poder proposta por Zakaria e, com Dahl, enfatiza que a qualidade democrática depende de fontes alternativas de informação acessíveis, confiáveis e plurais. O texto ainda busca explicar o funcionamento da *big data* (volume, variedade, velocidade) e sua simbiose com algoritmos: benefícios existem, mas a opacidade amplia vieses e incentiva manipulação. Exemplos como Cambridge Analytica em 2016, o papel do Facebook na violência contra os Rohingya e casos no Brexit e na Espanha ilustram a distorção do debate público e a assimetria entre plataformas, campanhas e cidadãos. Ao fim, propõe restringir ou banir microtargeting político, exigir transparência em propaganda e algoritmos, responsabilizar plataformas e atores políticos, avançar em modelos

de correção e investir em alfabetização digital para proteger a integridade eleitoral e a democracia – assim estabelecendo marcos regulatórios robustos.

Abstract

This text seeks to analyze how the explosion of data and the expansion of digital infrastructures, combined with technologies such as AI and algorithms, are reshaping contemporary politics. Anchored in concepts such as Zuboff’s “surveillance capitalism,” the text argues that behavioral data becomes an asset for profiling and inducing behavior in a democratic environment, shifting the electoral contest to opaque environments of microtargeting, disinformation, and massive use of bots. Revisiting democratic theory, the text contrasts elections with the limitation of power proposed by Zakaria and, with Dahl, emphasizes that democratic quality depends on accessible, reliable, and pluralistic alternative sources of information. The text also seeks to explain how big data (volume, variety, velocity) works and its symbiosis with algorithms: there are benefits, but opacity amplifies biases and encourages manipulation. Examples such as Cambridge Analytica in 2016, Facebook’s role in the violence against the Rohingya, and cases in Brexit and Spain illustrate the distortion of public debate and the asymmetry between platforms, campaigns, and citizens. Ultimately, it proposes restricting or banning political microtargeting, demanding transparency in advertising and algorithms, holding platforms and political actors accountable, advancing co-regulation models, and investing in digital literacy to protect electoral integrity and democracy—thus establishing robust regulatory frameworks.

1. Introdução

Nas últimas duas décadas, assistimos a uma transformação sem precedentes na capacidade de produzir, coletar e utilizar dados em

escala global. Esse fenômeno, frequentemente denominado como a “explosão dos dados”, está intrinsecamente ligado ao desenvolvimento de infraestruturas digitais, à proliferação de dispositivos conectados e ao avanço de tecnologias da chamada Quarta Revolução Industrial, como a inteligência artificial, *machine learning*, *blockchain* e *cloud computing*. A partir da popularização de plataformas online, como as mídias sociais, a captura de informações deixou de ser uma atividade restrita a governos e grandes corporações, passando a ser uma prática cotidiana e contínua, alimentada por interações, não apenas nas redes sociais, mas também em transações econômicas, deslocamentos georreferenciados e até em padrões biométricos, como reconhecimento facial.

O crescimento exponencial da produção de dados não ocorre de forma isolada, mas sim articulado a dinâmicas econômicas, culturais e, especialmente, políticas. O ‘capitalismo de vigilância’, notório conceito desenvolvido por Shoshana Zuboff¹, sintetiza como as plataformas digitais converteram dados comportamentais em ativos econômicos. Nesse modelo, dados não são apenas subprodutos das nossas interações online, mas sim matéria-prima para a construção de perfis, modelos preditivos e mecanismos de indução de comportamentos, seja para consumo na área econômica, seja para distorcer preferências na esfera política.

No campo político, o impacto dessas tecnologias em processos eleitorais é tão revolucionário quanto perigoso. Por um lado, a apropriação massiva de dados redefine práticas tradicionais de comunicação, mobilização e persuasão. O uso de *microtargeting*, a personalização de mensagens e os sistemas algorítmicos de recomendação conferem às campanhas eleitorais e aos atores políticos uma capacidade inédita de modular discursos para públicos específicos, muitas vezes de maneira opaca e assimétrica. Pelo outro lado, a capacidade de forjar narrativas, disseminar desinformação, mascarar o autor das mensagens e simular interação online pelo uso massivo de *bots*, têm abalado profundamente a forma

1 ZUBOFF. Shoshana. A Era do Capitalismo de Vigilância: a luta por um futuro humano na nova fronteira do poder. Rio de Janeiro, Ed. Intríneca, 2021.

como cidadãos/eleitores produzem suas preferências e como a própria natureza do debate político contemporâneo se dá. Esse processo não apenas tensiona os fundamentos da esfera pública deliberativa, como também introduz novos desafios para a regulação democrática, a proteção de dados pessoais e a garantia de eleições livres e justas.

2. Compreendendo a Democracia: o que está em risco?

A democracia é, simultaneamente, um conceito normativo e empírico, cuja definição atravessa séculos de debates filosóficos, políticos e sociais. Desde suas raízes na Grécia Antiga até sua consolidação como modelo hegemônico de organização política no mundo contemporâneo, a democracia permanece como um ideal em constante disputa, tendo sido reinterpretada segundo contextos históricos específicos. Embora suas raízes remontem à antiguidade clássica, particularmente à experiência ateniense, a compreensão contemporânea da democracia consolidou-se a partir de transformações ocorridas nos séculos XVIII e XIX, no contexto do constitucionalismo liberal e da ampliação progressiva do sufrágio. Desde então, a democracia passou a ser associada fundamentalmente à soberania popular, à competição eleitoral, ao pluralismo político e à proteção de direitos individuais e coletivos.

Destarte, apesar da extensa evolução pela qual o conceito de democracia passou, ela ainda carece de uma definição concreta, detalhada e de largo consenso. Tal fluidez contribui para que a democracia seja mais uma ideia ou ideal do que um conceito objetivo, o qual podemos operacionalmente delimitar. Isso faz com que a democracia seja percebida como um conjunto de elementos pouco coeso, e que, por tal, alguns desses elementos acabem se sobressaindo sobre outros, dando a entender que a democracia estaria intrinsecamente ligada à existência de alguns elementos e enquanto ela poderia prescindir de outros. Dito de outra forma, a correlação de democracia com elementos objetivos como soberania popular e competição eleitoral é amplamente aceita e entendida enquanto indispensável, sendo por muitos considerados como o núcleo

duro da democracia. Entretanto, outros elementos, como proteção de direitos individuais e necessidade de um governo responsivo e transparente, são mais difíceis de serem completamente entendidos e, portanto, tendem a ser mais facilmente relegados.

Essa problemática dinâmica, que tem ganhado destaque nas últimas décadas, foi muito bem capturada por Fareed Zakaria em seu conceito ‘democracia iliberal’². A ideia de Zakaria era a de chamar a atenção para uma preocupante tendência, que apenas se acentuou nos últimos anos, na qual regimes chegam ao poder por meio de eleições democráticas (ou seja, respeitando o princípio de escolha popular dos governantes), mas, uma vez no poder, ignoram os limites constitucionais e violam as liberdades e direitos individuais de seus cidadãos.

Em essência, Zakaria argumenta que a democracia, em sua forma mais básica de eleições e governo da maioria, não é sinônimo de liberalismo constitucional. Enquanto a democracia foca na origem do poder (soberania popular), o liberalismo constitucional foca nos limites do poder (proteção de direitos individuais e freios e contrapesos). Nesse sentido, Zakaria chama a atenção para o fato de que a mera realização de eleições pode levar a uma parcial “democratização”, mas potencialmente ao custo da garantia da liberdade – isso resultaria em regimes (semi)autoritários legitimados pelo voto popular. É interessante notar que a ideia de Zakaria não é exatamente nova, mas uma espécie repaginação aplicada aos tempos atuais: olhando para passado, o nazismo e o fascismo foram regimes que se arvoraram na vontade popular, e surgiram dentro dos limites estritamente legais. Nesse contexto de grave violação dos direitos humanos, não é de se estranhar que tais regimes tenham sido considerados adversos à democracia, mesmo estando neles presentes os elementos considerados pela maioria como o seu núcleo duro. Assim, o próprio conceito de “constituição” usado por Zakaria supera os limites do positivismo jurídico, na medida em que seu valor não

2 Zakaria, Fareed. *The Rise of Illiberal Democracy*. *Foreign Affairs*, v. 76, n. 6, p. 22-43, 1997.

está mais prioritariamente impresso no papel, mas sim nos princípios e ideais que podem ser interpretados a partir do texto constitucional³.

Outra influente contribuição para o debate sobre a democracia é a de Robert Dahl. Em seu livro *Polyarchy: Participation and Opposition*⁴, Dahl buscou operacionalizar o conceito, distanciando-se tanto de definições excessivamente idealistas quanto de abordagens meramente procedimentais. Seu objetivo foi construir um modelo analítico capaz de avaliar o grau de democratização de sistemas políticos concretos, levando em conta tanto as instituições formais, quanto os padrões de participação e competição.

A definição desenvolvida por Dahl não se limita ao ato eleitoral em si, mas abrange um conjunto de condições que garantem a pluralidade, a transparência e a responsividade do sistema político. Dessa forma, além de elementos óbvios como direito ao voto e importância de eleições livres e justas, Dahl destaca ainda que uma democracia depende da presença de fontes alternativas de informação. E é centrada na importância das fontes alternativas de informação que este artigo destaca como um dos principais desafios da democracia contemporânea.

Para Dahl, um sistema verdadeiramente democrático exige que os cidadãos tenham o direito e a capacidade real de buscar e acessar informações diversas que não sejam controladas, distorcidas ou monopolizadas pelo governo ou por um único grupo político dominante. Uma democracia saudável dependeria, portanto, da capacidade dos cidadãos de formularem livremente suas preferências de forma a obter uma participação informada sobre o debate público do qual fazem parte.

Em seus 'primórdios', a internet surgiu como uma força democratizante, prometendo acesso sem precedentes a informações alternativas e livre dos tradicionais filtros (*gatekeepers*). Entretanto, em um passado mais recente, a ampla adesão às plataformas online somada à difusão

3 Levitsky, S., & Ziblatt, D. *How democracies die*. Penguin Books, 2019.

4 Dahl, Robert A. *Polyarchy; participation and opposition*. New Haven, Ed. Yale University, 1971.

de tecnologias como *big data* e algoritmo, nos fez testemunhar o surgimento de um conjunto de fenômenos, como desinformação e *fake news*, bolhas, câmaras de eco, *trolls*, *microtargetting* etc., que tem impactado profundamente na manipulação das informações e, portanto, na capacidade dos cidadãos de se informar e de desenvolver suas preferências.

Dessa forma, embora as novas tecnologias tenham oferecido uma promessa inicial de expandir as fontes alternativas de informação, elas introduziram simultaneamente desafios profundos que ameaçam a qualidade e a integridade do ambiente informacional. A mera disponibilidade de fontes alternativas é insuficiente; tão importante quanto é a acessibilidade efetiva de informações confiáveis, diversas e verificáveis que os cidadãos possam usar para informar genuinamente sua participação política.

3. Contextualização: o que É *Big Data* e como ela opera

Nas últimas décadas, tecnologias como *big data* tornaram-se elementos centrais na organização da vida social, econômica e política. Para Cukier and Mayer-Schönberger⁵, *big data* pode ser entendido enquanto uma “ferramenta que aumenta a capacidade da sociedade de aproveitar as informações que gera, obtendo *insights* úteis, bens e serviços de valor significativo”. Isso significaria que tal tecnologia teria a capacidade de coletar, armazenar e processar conjuntos massivos de dados, caracterizados por três dimensões fundamentais: volume, variedade e velocidade.

Esses dados são provenientes de inúmeras fontes, como: redes sociais, sensores, registros financeiros, dispositivos móveis, entre outros. A análise massiva desses dados permite a identificação de padrões, corre-

5 Mayer-Schönberger, V., & Cukier, K. *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt, 2013. (tradução livre)

lações e tendências que eram inacessíveis pelos métodos tradicionais até então utilizados.

Big data, entretanto, só adquire sentido prático quando associado aos algoritmos, que são essencialmente sequências de instruções matemáticas e lógicas projetadas para organizar, filtrar e extrair informações relevantes desses vastos bancos de dados acumulados pela *big data*. Essa relação simbiótica entre dados e algoritmos sustenta desde serviços triviais, como recomendações de filmes, até aplicações complexas na área da saúde, segurança, finanças e governança.

Conforme já mencionado, essas tecnologias oferecem impactos ambivalentes. Por um lado, seus benefícios são inegáveis: permitem diagnósticos médicos mais precisos, otimizam a gestão de recursos públicos, aumentam a eficiência logística e personalizam serviços. Por outro lado, sua utilização também gera externalidades negativas, como a intensificação da vigilância, a erosão da privacidade, a amplificação de vieses algorítmicos e a manipulação das preferências das pessoas ao induzir comportamento. Isso ocorre especialmente quando os critérios de decisão dos algoritmos são opacos e não auditáveis (o que é bastante frequente, principalmente no contexto de grandes corporações protegidas pelo sigilo comercial).

No contexto eleitoral, os riscos associados ao uso de *big data* e algoritmos são particularmente preocupantes. Essas tecnologias podem ser instrumentalizadas para práticas de *microtargeting* político, nas quais mensagens específicas são direcionadas a segmentos ‘hipersegmentados’ da população, com base em traços psicológicos, preferências de consumo e comportamento online. Tal prática não apenas enfraquece a esfera pública deliberativa, como também favorece a disseminação de desinformação personalizada e não identificável para o restante do eleitorado, dificultando o contraditório.

Exemplos recentes desse fenômeno têm colocado tal desafio no topo da agenda política contemporânea. Talvez o primeiro escândalo de maior notoriedade foi o caso da empresa Cambridge Analytica nas eleições de 2016 nos Estados Unidos, no qual dados de 87 milhões de usuá-

rios do Facebook foram utilizados para influenciar preferências eleitorais de forma clandestina⁶. Por meio do mapeamento das interações dos usuários, desde repostas de *quizzes* até simples *likes*, A Cambridge Analytica foi capaz de desenvolver um banco de dados sofisticado com perfis psicométricos de cerca de 2 milhões de eleitores norte-americanos em onze estados considerados chave para disputa. Uma vez identificados, esses eleitores foram bombardeados de propaganda direcionada e customizada para o seu perfil, de modo a influenciar suas preferências por um determinado candidato (ou rejeitar o rival).

Já o Caso de Mianmar, ainda que não relacionado com eleições especificamente, é emblemático no que tange à capacidade das redes sociais desenvolverem câmaras de eco que fomentam discurso de ódio, polarização e, por consequência, violência. Nesse contexto, entenderemos câmaras de eco como espaços informacionais, especialmente em redes sociais e plataformas online, nos quais indivíduos são sistematicamente expostos a conteúdos que confirmam suas crenças, valores e visões de mundo preexistentes, enquanto informações dissonantes ou contraditórias são filtradas, marginalizadas ou invisibilizadas. Assim, o resultado frequente dessa dinâmica é a promoção de polarização, radicalização de opiniões e a circulação contínua de desinformação ou discursos de ódio.

Em 2017,⁷ grupos radicais budistas e outros ligados às forças armadas de Mianmar iniciaram uma campanha online de discurso de ódio e de desinformação sobre supostos crimes cometidos por integrantes do grupo étnico majoritariamente mulçumano Rohingya. Tal campanha

6 Alex Hern. Cambridge Analytica: how did it turn clicks into votes?. The Guardian. 06 May 2018. Acessado em: 28 Maio de 2025. <https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie>

7 Amnesty International. Myanmar: Facebook's systems promoted violence against Rohingya; Meta owes reparations. Amnesty International. 29 Sep 2022. <https://www.amnesty.org/en/latest/news/2022/09/myanmar-facebook-systems-promoted-violence-against-rohingya-meta-owes-reparations-new-report/>

alavancada pelo algoritmo do Facebook deu ampla visibilidade para tal conteúdo no seu *news feed* e recomendações – efetivamente criando uma câmara de eco anti-Rohingya no país. O que se seguiu foi uma explosão de violência em que os Rohingya foram alvos de ataques, resultando na fuga de milhares pessoas que tiveram de buscar refúgio em países vizinhos.

Esses casos evidenciam que, embora *big data* e algoritmos ofereçam inovações valiosas, sua utilização no campo da política, quando desregulada, representa uma ameaça concreta não apenas à integridade democrática, mas à liberdade e aos direitos humanos. A coleta, a gestão e o uso dos dados têm se tornado uma questão política e social central para o debate público contemporâneo, na medida em que seu uso tem reconfigurado de forma significativa as nossas escolhas e relacionamentos. Além das Big Tech, os atores políticos têm sido um dos grupos que tem repetido logrado em extrair significativos benefícios do uso da *big data* e dos algoritmos. Tirando proveito da frouxa ou inexistente regulamentação sobre o uso de dados, atores políticos têm aproveitado brechas na legislação para usar tais ferramentas para ativar suas bases, melhor difundir suas mensagens e, ao mesmo tempo (em alguns casos), disseminar desinformação, difundir discurso de ódio contra grupos rivais e tentar induzir o comportamento dos cidadãos para atender sua agenda política⁸. Assim, com uma regulamentação relativamente frágil, o cenário político se beneficia dessa ambiguidade, fator que impede constantemente um entendimento mais claro da influência que esses dados podem exercer sobre os processos de decisão.

4. Eleições, Dados e Tecnologias Manipulativas

Conforme visto acima, o avanço acelerado das tecnologias digitais nas últimas décadas tem reconfigurado profundamente os proces-

8 Macnish, K. and Galliot, J. *Big data and democracy*. Edinburgh: Edinburgh University Press, 2020.

tos democráticos, particularmente no que se refere às dinâmicas eleitorais. Ferramentas de coleta massiva de dados, algoritmos e técnicas de inteligência artificial são hoje elementos centrais nas estratégias de comunicação dos atores políticos. Entretanto, apesar do uso dessas tecnologias no contexto eleitoral ter a capacidade de, positivamente, aprimorar campanhas ou facilitar o contato com eleitores, elas também têm sido instrumentalizadas, frequentemente, para práticas manipulativas, que incluem desde a disseminação de desinformação até *microtargeting*, uso massivo de *bots* e falta de transparência dos transmissores das informações. Essas práticas criam distorções que representam riscos significativos para a integridade das eleições, corroendo os fundamentos do livre debate público, a autonomia dos cidadãos em produzir suas preferências e, por consequência, a própria legitimidade democrática⁹.

Entre os mecanismos mais problemáticos estão os sistemas de *microtargeting*, que permitem que campanhas políticas ou atores externos segmentem eleitores com base em dados comportamentais, socioeconômicos e até psicológicos (dados esses coletados sem o consentimento expresso das pessoas). Essa técnica, além de violar a privacidade dos eleitores, fragmenta o espaço informacional, permitindo que mensagens diferentes, e muitas vezes contraditórias, sejam direcionadas a grupos distintos, explorando suas crenças, medos ou preconceitos. Ao operar dessa forma, o *microtargeting* enfraquece o debate público aberto e dificulta o escrutínio coletivo das mensagens políticas.

Paralelamente, o uso de *bots* (contas automatizadas, que nesse caso são programadas para simular interações humanas) amplia artificialmente o alcance de conteúdo específico, muitas vezes desinformativos, gerando percepções distorcidas sobre apoio popular, legitimidade de discursos ou relevância de determinados temas. O uso de *bots* está vinculado à dificuldade dos usuários em identificar a identidade dos emissores de informações. Na medida em que as plataformas expõem

9 Efthymiou–Eggleton, T. W., Sidiropoulos, S., Spanos, E. and Stougiannou, E. Big data and Democracy. HAPSc Policy Briefs Series, 1(2): 18-25, 2020.

os usuários a um significativo número de informações provenientes de grupos de interesse políticos e de diferentes mídias (desde meios de comunicação tradicionais até mídias alternativas com interesses muito particulares), os usuários frequentemente não apenas têm dificuldade de identificar a identidade do autor da peça de informação (ou até em saber se trata-se de um *bot* ou de uma pessoa de fato), como ainda com frequência tendem a confundir a procedência das informações que lhes são expostas – criando assim maiores barreiras para os usuários discernirem a fazer discernimento da fonte da notícia/informação e, por conseguinte aferirem sua credibilidade.

Para compreender esse fenômeno, reflita sobre a seguinte situação. Imagine um determinado ator político que publica em uma rede social um conteúdo com informações inverídicas. Na sequência, ele aciona uma empresa que opera milhares de *bots*. Esses *bots* irão comentar, compartilhar e dar *likes* na postagem original. Essa postagem rapidamente terá centenas de milhares, senão milhões, de interações online, capturando assim a atenção da mídia tradicional que irá repercutir o fato imaginando que se trata de um tema muito popular. Tendo essa postagem um altíssimo número de interações e cobertura da mídia tradicional, agentes políticos irão reagir a esse tema. Dessa forma, um tema irá, artificialmente, mobilizar e alterar o debate público nesse determinado contexto – reagindo a informações inverídicas e efetivamente sem real interesse dos cidadãos.

Em pesquisa recente¹⁰, Van Gils (et al, 2020) buscou modular matematicamente o impacto da manipulação de preferências de eleitores alvos de práticas de *microtargeting* e de compartilhamento de desinformação na qual o indivíduo não tem ciência da identidade e da posição política do autor da informação. Seu objetivo era verificar quais dessas práticas distorcia mais as preferências originais dos eleitores. Através de seus modelos, Van Gils e seus colegas identificaram que, apesar de

10 Van Gils, Freek; Muller, Wieland and Prufer, Jens. *Big data and Democracy*. TIL-LEC Discussion Paper No. 2020-003, 2020.

ambas as práticas terem potencial de manipular as preferências dos eleitores, apenas a prática de compartilhamento de informações em que os eleitores não têm clareza sobre identidade e posição do emissor da mensagem tem capacidade de concretamente modificar os resultados eleitorais. Nesse sentido, o incrível fluxo de mensagens a que as pessoas cotidianamente são expostas, tende a dificultar que um dado eleitor seja capaz de associar um conteúdo ao seu emissor original, as reais intenções do emissor e a credibilidade da informação.

Ambas as práticas já causaram impactos significativos recentes em processos eleitorais ao redor do mundo. Um caso emblemático foi o referendo do Brexit em 2016, no qual grupos favoráveis à saída do Reino Unido da União Europeia contrataram empresas como AggregateIQ, vinculada à Cambridge Analytica¹¹. Investigações conduzidas pelo Information Commissioner's Office do Reino Unido revelaram que dados pessoais de milhões de cidadãos britânicos foram coletados ilegalmente por meio de aplicativos e questionários diversos em redes sociais. Esses dados foram usados para alimentar modelos psicográficos sofisticados que orientaram campanhas de *microtargeting*, com mensagens voltadas a acentuar medos relacionados à imigração, à insegurança econômica e à perda de soberania, impactando diretamente o comportamento dos eleitores.

Já nas eleições gerais na Espanha em 2019, foi documentado o uso sistemático de *bots* e campanhas coordenadas de desinformação nas redes sociais, especialmente no Twitter e no Facebook¹². Essas campanhas tinham como objetivo amplificar discursos polarizadores relacionados

-
- 11 Carole Cadwalladr. AggregateIQ: the obscure Canadian tech firm and the Brexit data riddle. The Guardian. 31 Mar 2018. <https://www.theguardian.com/uk-news/2018/mar/31/aggregateiq-canadian-tech-brexit-data-riddle-cambridge-analytica>
 - 12 Paniagua Rojano, Francisco; Seoane Pérez, Francisco y Magallón-Rosa, Raúl. Anatomía del bulo electoral: la desinformación política durante la campaña del 28-A en España. Revista CIDOB d'Afers Internacionals, n.º 124, abril de 2020, p. 123-145.

à crise na Catalunha, à imigração e à desconfiança nas instituições europeias. Relatórios produzidos pela própria União Europeia e por organizações de verificação de fatos, como a EU DisinfoLab, identificaram redes de contas falsas que, além de disseminarem conteúdos manipulativos, também direcionaram anúncios baseados em dados pessoais obtidos sem consentimento. Essas práticas foram associadas tanto a atores domésticos quanto estrangeiros interessados em gerar instabilidade política na Espanha e, por extensão, no bloco europeu.

Ambos os episódios não são casos isolados, tampouco restritos a falhas pontuais de regulação. Eles revelam um problema estrutural na intersecção entre dados, tecnologia e democracia. A assimetria informacional gerada pela posse e manipulação de grandes volumes de dados cria um desequilíbrio profundo nas relações entre campanhas, plataformas digitais e cidadãos. Mais grave ainda, a opacidade dos algoritmos que regem a distribuição de conteúdo nas redes sociais impede que haja transparência sobre quem emite e quem recebe um dado tipo de mensagem, dificultando o escrutínio público e jurídico dessas práticas. Combinado ao uso de *bots* e redes de desinformação, esse cenário fragiliza a capacidade dos eleitores de formar suas preferências de forma livre, informada e racional – pilares centrais da democracia deliberativa.

5. Considerações Finais: desafios regulatórios

O avanço exponencial das tecnologias digitais como *big data*, sistemas algorítmicos, e, especialmente, das plataformas de redes sociais, introduziu desafios substanciais à integridade dos processos eleitorais contemporâneos. A capacidade dessas tecnologias em moldar percepções, influenciar comportamentos e distorcer preferências dos eleitores por meio de desinformação e manipulação não pode ser subestimada. A assimetria informacional criada pela opacidade dos algoritmos e pelo uso intensivo de dados pessoais coloca em risco princípios fundamentais da democracia, como a igualdade no acesso à informação,

a liberdade de escolha e a autonomia dos eleitores em construir suas preferências com base em informações verdadeiras e verificáveis.

Dentre os instrumentos mais preocupantes encontra-se o *microtargeting*. Esse tipo de prática não apenas fragmenta o debate coletivo e cria câmaras de eco, mas também facilita a propagação de conteúdos falsos, descontextualizados ou manipuladores, impossibilitando que o eleitorado tenha uma visão ampla e plural das propostas em disputa. Diante desse cenário, torna-se necessário debater seriamente políticas que visem o **banimento ou restrições ao microtargeting político**, especialmente quando baseado em dados sensíveis ou em informações comportamentais extraídas sem consentimento claro das pessoas.

Paralelamente, a opacidade que permeia tanto a origem quanto a difusão de conteúdos patrocinados configura outro vetor crítico de erosão da transparência democrática. Portanto, é imperativo avançar na construção de marcos regulatórios que estabeleçam a **transparência obrigatória na propaganda política online**, incluindo a exigência de que seja explicitada, de forma acessível e compreensível, a identidade dos responsáveis pela produção, financiamento e disseminação dos conteúdos. Tal medida, além de coibir práticas fraudulentas, fortalece a *accountability* dos agentes políticos perante o eleitorado e as instituições reguladoras.

Além disso, os modelos algorítmicos que determinam a circulação e a visibilidade de conteúdos nas plataformas devem ser objetos de transparência e *accountability* estruturada. A sociedade civil, as autoridades eleitorais e os órgãos reguladores precisam ter acesso aos critérios, parâmetros e efeitos desses algoritmos, especialmente quando impactam diretamente o debate político e a formação da opinião pública. A responsabilização das plataformas não pode se limitar a compromissos voluntários ou a práticas de moderação reativa. É essencial que haja **responsabilização jurídica efetiva das plataformas digitais pela promoção de conteúdos que comprometam a integridade eleitoral**, em especial no contexto de eventual resistência ou obstrução a ordens da justiça no combate à desinformação e aos ataques à lisura do processo democrático.

Ainda, em publicação recente no Cadernos Adenauer 3 2024, João Victor Archegas¹³ defende uma abordagem interessante e inovadora baseada no constitucionalismo digital, que implicaria na **corregulação**, que configuraria uma alternativa média entre a autorregulação pura das *big techs* e a regulamentação unilateral proposta por governos. Tal modelo, segundo Archegas, permitiria uma ação conjunta baseada em cooperação e complementariedade entre governos e corporações potencialmente apoiada por especialistas neutros.

Por fim, qualquer estratégia regulatória será insuficiente se não for acompanhada por um investimento robusto e contínuo em **alfabetização e educação digital da sociedade**. Capacitar os cidadãos para reconhecer práticas manipulativas, interpretar criticamente as informações e compreender o funcionamento básico dos sistemas digitais é uma condição básica e indispensável para o fortalecimento da resiliência democrática. A promoção da cultura digital crítica deve ser incorporada às políticas públicas educacionais e à própria atuação das instituições democráticas, como forma de criar anticorpos sociais contra a erosão da verdade e a manipulação informacional.

Não obstante, o caminho para a regulação digital, inclusive em contexto eleitoral, tem se mostrado uma árdua batalha. A começar pelo impressionante contraste na assimetria de velocidade entre a inovação e a regulação, o debate sobre os limites da inteiração digital e da responsabilidade das corporações de serviços digitais tem sido sobrecarregado por discursos que defendem uma percepção distorcida de liberdade, que defende um ideal de liberdade pré-social. Tais desafios têm sido ainda turbinados pelo poderoso *lobby* das *big techs* sobre os agentes políticos, que ainda rompem com a neutralidade de suas funcionalidades e buscam manipular o debate público. Caso emblemático que ilustra esse ponto pode ser aferido à plataforma de buscas Google, que na ocasião do debate sobre o Projeto de Lei 2630/2020, que buscava regular

13 Archegas, João Victor. Inovações Tecnológicas e Democracia. Cadernos Adenauer 3, 2024. pp. 9-30.

as plataformas digitais, combater a desinformação e as *fake news*, e estabelecer responsabilidades para às *big techs*. Na ocasião, a Google não apenas incluiu mensagem em sua página de entrada contra o projeto de lei, como ainda foi detectado que a plataforma usou de seu algoritmo para privilegiar posições contrárias à legislação nas buscas feitas em sua plataforma¹⁴.

Nesse sentido, parece haver uma percepção de que as corporações digitais são o principal desafio a uma regulação digital equilibrada e responsável. Como o próprio Archegas (2024) pontua em seu artigo, o poderio das *big techs* tem um impacto expressivo no debate público, geralmente rivalizando com o poder estatal. Essa percepção, no entanto, é uma meia verdade. De fato, o poder econômico e a capacidade de usar tecnologias manipulativas para influenciar a percepção das pessoas torna as *big techs* um obstáculo formidável à implementação de legislações que limitem suas ações. Não obstante, julgo que a literatura ainda não pesou adequadamente o desafio posto por um outro segmento que também tem agido em consonância com as *big techs* em posição contrária a uma regulação digital mais rígida: [a classe política](#).

Por um lado, é verdade que em diversos casos pudemos testemunhar que o debate sobre a legislação digital pôs em lados opostos corporações digitais e o poder estatal. No entanto, na maior parte desses casos, os representantes do poder estatal foram às cortes e às agências reguladoras. Atores da classe política, em especial representantes de alguns segmentos ideológicos mais extremos do espectro político, têm se posicionado com os verdadeiros obstáculos do debate. Conforme já mencionado, a classe política tem extraído muitos benefícios de tecnologias como *big data*, algoritmo e *bots* para sua ação política. Sua utilização tem permitido não apenas campanhas políticas mais eficientes e ações de ataque a rivais políticos mais eficazes, como a capacidade de

14 Rodriques, Alex. regular as plataformas digitais, combater a desinformação e fake news, e estabelecer responsabilidades para as big techs. Agência Brasil. 02 Mai 2023. <https://agenciabrasil.ebc.com.br/geral/noticia/2023-05/notificada-google-retira-link-para-texto-contr-pl-das-fake-news>

alavancar artificialmente narrativas que servem a seus interesses pessoais. Assim sendo, tais atores políticos não estariam apenas vulneráveis ao *lobby* das *big tech*, com boa parte da comunidade epistêmica acreditada¹⁵. Esses atores políticos passam a ter interesses próprios no *status quo* do ambiente digital, pois além de extraírem benefícios políticos, muito ainda conseguem obter significativos benefícios econômicos dada a monetização do conteúdo que eles geram¹⁶. Dessa forma, uma regulação digital que exija maior transparência, responsabilidade e *accountability*, impactaria de forma determinante em práticas predatórias e ilegais que tanto têm favorecido determinadas forças políticas.

Diante desse panorama, torna-se evidente que os riscos impostos pelo uso de dados e tecnologias manipulativas no contexto eleitoral não podem ser tratados como externalidades do progresso tecnológico, mas sim como desafios centrais para a governança democrática. A resposta a essas ameaças passa necessariamente pela adoção de marcos regulatórios robustos que combinem a proteção de dados pessoais, a transparência algorítmica, a responsabilização das plataformas digitais e de atores (políticos) que incorram em práticas predatórias. Além disso, é fundamental investir em estratégias de educação e de desenvolvimento de competências digitais que permitam aos cidadãos reconhecer práticas manipulativas e navegar de forma crítica no ecossistema digital. Somente assim será possível garantir que a inovação tecnológica tenha seu impacto positivo na democracia, e não seja instrumento de sua erosão.

15 Cowgill, Bo; Prat, Andrea e Valletti, Tommaso. Political Power and Market Power. General Economics. 2023.

16 Motoryn, Paulo. Nikolas Ferreira e Gustavo Gayer driblam proibição e usam mandatos para ganhar dinheiro no Instagram. Intercept. 29 Feb 2024. <https://www.intercept.com.br/2024/02/29/nikolas-ferreira-e-gustavo-gayer-driblam-proibicao-e-usam-mandatos-para-ganhar-dinheiro-no-instagram/>

Leonardo Paz Neves · Cientista político. É analista de inteligência internacional na Unidade de Inteligência Internacional da Fundação Getulio Vargas e professor do Departamento de Relações Internacionais da Faculdade Ibmecc. Além disso, atuou como Coordenador de Projetos e Coordenador de Estudos e Debates do Centro Brasileiro de Estudos e Debates Internacionais (CEBRI), colaborou com a Seção de Assuntos Cíveis do Centro Conjunto Brasileiro de Operações de Paz (CCOPAB), das Forças Armadas Brasileiras, e foi Coordenador Executivo do Grupo de Análise sobre Prevenção de Conflitos Internacionais (GAPCon/UCAM). É formado em Ciências Sociais pela Universidade Federal Fluminense, mestre em Ciência Política pela IUPERJ e doutor em Desenvolvimento pelo Instituto de Economia da Universidade Federal do Rio de Janeiro. Publicou livros como “Estados falidos e o eixo do mal na política de segurança norte-americana”; “Repensando a política externa brasileira: em busca de novos consensos” e “O CEBRI e as relações internacionais no Brasil”; tendo este último recebido o Troféu Cultura Econômica do Jornal do Comércio em 2014.

O uso de Big Data e da Inteligência Artificial Generativa nas campanhas eleitorais

Joscimar Souza Silva
Isabela Rocha

Resumo

Este artigo propõe diretrizes práticas para o uso estratégico de *Big Data* e Inteligência Artificial Generativa (IAG) em campanhas eleitorais, com foco no contexto democrático brasileiro. Partindo da teoria da poliarquia de Robert Dahl, discutimos os riscos da concentração informacional e da vigilância algorítmica, destacando a importância das fontes alternativas de informação para o equilíbrio democrático. Em seguida, exploramos os dilemas éticos do uso dessas tecnologias em sociedades marcadas por desigualdades estruturais e baixa alfabetização digital. Por fim, propomos quatro eixos práticos de atuação: (1) planejamento orientado por dados, (2) conexão entre identidade da candidatura e tática digital, (3) sustentação da grande narrativa e (4) construção de conteúdo orgânico com base em afetos positivos e capilaridade espontânea. A proposta se afasta da visão tecno determinista e busca articular inovação tecnológica com responsabilidade política, oferecendo um roteiro crítico e aplicado para campanhas que desejam disputar o debate público sem romper com os princípios do pluralismo democrático.

Abstract

This article proposes practical guidelines for the strategic use of Big Data and Generative Artificial Intelligence (GAI) in election campaigns, focusing on the Brazilian democratic context. Based on Robert Dahl's theory of polyarchy, we discuss the risks of information concentration and algorithmic surveillance, highlighting the importance of alternative sources of information for democratic balance. We then explore the ethical dilemmas of using these technologies in societies marked by structural inequalities and low digital literacy. Finally, we propose four practical lines of action: (1) data-driven planning, (2) connection between candidate identity and digital tactics, (3) support for the grand narrative, and (4) construction of organic content based on positive affect and spontaneous capillarity. The proposal departs from a techno-deterministic view and seeks to articulate technological innovation with political responsibility, offering a critical and applied roadmap for campaigns that wish to engage in public debate without breaking with the principles of democratic pluralism.

1. Introdução

Como incorporar *Big Data* e ferramentas de Inteligência Artificial Generativa (IAG) no planejamento e execução de campanhas eleitorais, de forma eficiente e ética, respeitando os princípios do pluralismo democrático e a legislação vigente? Aqui, não tratamos este questionamento como uma pergunta de pesquisa tradicional, e sim, como provocação prática e teórica que começamos a explorar neste artigo à luz das transformações já em curso no marketing político e nas dinâmicas de comunicação eleitoral.

A questão acima tem se tornado cada vez mais relevante pois as recentes campanhas eleitorais têm sido marcadas pelo uso de *Big Data*. No Brasil, da disputa para a Presidência da República à disputa eleito-

ral em Correntina, pequeno município do Oeste baiano¹, as candidaturas utilizam *Big Data*. Por *Big Data* compreendemos um conjunto de tecnologias e práticas que permitem o processamento e a análise de grandes volumes de dados, em alta velocidade e de diversas origens, em diferentes formatos, como estruturados, semiestruturados e não estruturados, que são coletados continuamente de fontes variadas e processados em ambientes digitais.

Ainda que sem as estruturas técnicas para manipular grandes bases de dados, *Big Techs* como Meta e Google oferecem às equipes de campanha *Big Data* para impulsionamento e segmentação de conteúdo através de serviços pagos (Bittencourt, GOLTZMAN, 2024). Se antes as campanhas eleitorais se baseavam em comícios, panfletos e propagandas em meios tradicionais como rádio e TV aberta, hoje a vitória eleitoral depende do uso estratégico de ferramentas digitais sofisticadas.

O uso criativo e orgânico de *Big Data* e, agora com a disponibilidade e intensificação do uso de IAG, ampliam drasticamente a capacidade de segmentar e personalizar mensagens, permitindo modelar, e, em certa medida, prever, o comportamento político do eleitorado, além de automatizar a produção de conteúdos que ressoam com públicos específicos (RICHARD, 2024). A IAG “se refere a sistemas ou modelos de inteligência artificial que têm a capacidade de gerar conteúdo original e criativo, como imagens, música, textos e inclusive vídeos” (Richard, 2024, p. 129). Além disso, as inteligências artificiais (IA) possuem capacidade de classificação de dados e predição de resultados, bem como de análise quantitativa ou qualitativa de dados, com nível intermediário de sofisticação analítica, mediante comandos (*prompts*) simples.

Essas transformações, no entanto, não dizem respeito apenas à comunicação política. Elas tocam diretamente a manipulação da opinião pública e os novos desafios da desinformação, cada vez mais difícil de

1 Fonte: Divulgação de Candidaturas e Contas Eleitorais. Item: Impulsionamento de Conteúdo. Disponível em: <https://divulgacandcontas.tse.jus.br/divulga/#/home>. Acesso em Junho de 2025.

rastrear e conter. Trata-se, portanto, de um novo ecossistema informacional, no qual o domínio técnico dessas tecnologias pode significar vantagem eleitoral – mas também risco à integridade do debate democrático (RICHARD, 2024). Por exemplo, o 2º Seminário Nacional de Comunicação do Partido Liberal, do ex-presidente Bolsonaro, realizado em Fortaleza-Ceará em 30 de maio de 2025, contou tanto com a presença dos gigantes da tecnologia – a *Big Tech* – quanto de plataformas de comunicação, onde oficinas capacitantes foram oferecidas aos participantes, incluindo treinamentos práticos sobre uso avançado de redes sociais, automação de conteúdo e mobilização via aplicativos de mensagens. Representantes da Meta e do Google apresentaram ferramentas de segmentação de audiência, criação de narrativas personalizadas com o uso de inteligência artificial, e estratégias para impulsionamento de engajamento político. A Meta, por exemplo, demonstrou como elaborar *propostas* que orientam a IA a gerar conteúdos mais eficientes para contextos eleitorais, inclusive sugerindo formas de introduzir o adversário político na mensagem. Já o Google exibiu aplicações da IA generativa em vídeos para o YouTube e explicou como produzir podcasts automatizados com o Notebook LLM, ferramenta gratuita voltada à produção de conteúdo escalável².

Neste sentido, e ainda que grande parte da literatura sobre o uso de *Big Data* e IAG nas campanhas eleitorais têm dado ênfase às capacidades técnicas e aos desafios éticos e regulatórios, e reconheçamos a importância destas questões, também é claro que o campo de estudos sobre campanhas eleitorais ainda fez poucas incursões sobre o tema, exceto sob a ótica regulatória (BITTENCOURT, GOLTZMAN, 2024; AMAYA LÓPEZ, CUEVA GAIBOR, 2025; DOURADO *et al*, 2025). Assim, nosso artigo busca contribuir com um aspecto mais prático e rotineiro do planejamento e execução de campanhas eleitorais, pensando como usar *Big Data* e recursos de IAG de forma aplicada e contextualizada ao ambiente da dis-

2 PODER360. “Mais preparados do que nunca”, diz PL sobre seminário de comunicação. Poder360, 30 maio 2025. Disponível em: <https://www.poder360.com.br/poder-partidos-politicos/mais-preparados-do-que-nunca-diz-pl-sobre-seminario-de-comunicacao/>. Acesso em: 2 jun. 2025.

puta eleitoral, com foco em planejamento de campanhas e marketing eleitoral, resguardando os princípios do pluralismo democrático.

Dessa forma, este artigo tem como objetivo apresentar orientações práticas para o uso estratégico e democrático de *Big Data* e IAG em campanhas eleitorais, com foco no planejamento, na comunicação e na mobilização digital em campanhas eleitorais. Para fins práticos, partimos do exemplo brasileiro, dada a necessidade de respeito à legislação eleitoral. Para isso, trataremos brevemente do aporte teórico recente sobre uso de *Big Data* e IAG nas campanhas eleitorais em nível geral. Em seguida, apresentamos algumas considerações sobre planejamento de campanhas eleitorais e propaganda eleitoral e, por fim, apresentamos algumas proposições de uso de *Big Data* e a IAG nas campanhas eleitorais, tendo como base o contexto brasileiro.

Além desta seção inicial, o artigo contará com mais três. Na seção seguinte, apresentamos uma reflexão conceitual sobre o pluralismo democrático, articulada a uma revisão crítica da literatura recente sobre o uso de *Big Data* e IAG em campanhas eleitorais como base para justificar a construção de um manual aplicado que respeite os princípios democráticos. A terceira seção tem caráter mais instrumental, discutindo aspectos centrais do planejamento de campanha e do marketing político digital, e apresentando instruções práticas para o uso democrático e estratégico de *Big Data* e IAG, considerando tanto os recursos disponíveis quanto os limites éticos e legais da sua aplicação. Finalmente, a última seção reúne as considerações finais, retomando os objetivos do artigo e reforçando a importância de aliar inovação tecnológica, responsabilidade política e compromisso com os valores democráticos nas disputas eleitorais contemporâneas.

2. Marketing, *Big Data* e IAG numa sociedade democrática plural

Articulamos nesta seção os fundamentos teóricos, dilemas éticos e usos estratégicos relacionados ao emprego de *Big Data* e IAG em

campanhas eleitorais. Partimos da teoria da poliarquia de Robert Dahl para discutir o papel das fontes alternativas de informação como condição para a democracia pluralista e os riscos da concentração informacional no contexto digital, e, adiante, abordamos os desafios éticos do uso dessas tecnologias, com ênfase nas desigualdades estruturais e no impacto da vigilância algorítmica sobre a autonomia política dos cidadãos. Por fim, analisamos como o *Big Data* e a IAG vêm sendo incorporados às estratégias de marketing político, especialmente por meio do *microtargeting*, ressaltando seus efeitos sobre o espaço público, a equidade democrática e a qualidade do debate eleitoral.

2.1. Pluralismo Democrático e as fontes alternativas de informação como fundamento de equilíbrio do jogo democrático

A teoria da poliarquia de Robert Dahl, desenvolvida principalmente em obras como “Poliarquia: Participação e Oposição” (1997 [1972]) e aprofundada em “Democracia e seus críticos” (1989) e “*On Democracy*” (2001), estabelece um conjunto de garantias institucionais mínimas necessárias para que um regime político possa ser considerado uma aproximação funcional da democracia em larga escala. Dentro dessas garantias, a existência de fontes alternativas de informação emerge como um fundamento indispensável, ligado à própria viabilidade da participação cidadã e da contestação política, elementos centrais da democracia pluralista.

O pluralismo, não parte de uma visão idealizada da sociedade, mas do reconhecimento de que ela é composta por diversos centros de poder como grupos de interesse, partidos e organizações com preferências e visões diferentes, que disputam influência e negociam entre si no processo político. Desta maneira, as fontes alternativas de informação são vitais para a própria dinâmica pluralista pois permitem que os diversos grupos sociais articulem seus interesses, mobilizem apoio para suas causas, fiscalizem as ações dos detentores do poder (governamental ou não) e

participem de forma mais equitativa do debate público. Ao impedir que um único grupo monopolize a esfera informacional, as fontes alternativas contribuem para a dispersão do controle sobre a informação, que, por sua vez, é um reflexo e uma condição necessária para a dispersão do poder político que caracteriza a poliarquia (DAHL, 1989, 1997, 2001).

A campanha eleitoral representa um momento crucial na dinâmica democrática, um período delimitado no tempo e regido por normas específicas, onde se desenrola um diálogo fundamental entre as candidaturas e o eleitorado. Nesse contexto, a disponibilidade de fontes alternativas de informação tornam-se cruciais. Campanhas políticas são capazes de construir “*mundos possíveis*” através de narrativas e retóricas específicas de interpretação enquadrada sobre o mundo atual, buscando persuadir o eleitorado ao interpretar a realidade e projetar futuros desejáveis ou indesejáveis (FIGUEIREDO et. al. 1997), porém, esta construção discursiva depende da seleção, enquadramento e disseminação de informações.

Se o acesso dos cidadãos à informação for dominado pelas narrativas das próprias campanhas ou por fontes alinhadas a elas, sem o contraponto de fontes alternativas e independentes, a capacidade do eleitorado de avaliar criticamente esses “mundos possíveis” e formar preferências autônomas fica severamente comprometida. As fontes alternativas de informação, no modelo de Dahl, funcionam assim como um antídoto necessário contra a potencial manipulação inerente às estratégias de persuasão eleitoral, garantindo que o “diálogo” entre candidatos e eleitores não se transforme em um monólogo controlado por quem detém mais recursos comunicacionais. A diversidade e a integridade da informação circulante são comprometidas quando o uso de *Big Data*, IAG e as dinâmicas das mídias sociais digitais ficam concentrados nas mãos de apenas um dos lados da disputa eleitoral (RICHARD, 2024).

Assim, resguardados os cuidados e considerando ambientes democráticos pluralistas, não basta diagnosticar os riscos do uso abusivo da informação: é fundamental, produzir informação resguardando os limites da disputa democrática de maneira capaz de enfrentar a desinformação e as estratégias de manipulação. O desafio não está em demonizar

os meios, mas em dominá-los³. É preciso conhecê-los por dentro, com consciência crítica e domínio técnico, para transformar suas lógicas em ferramentas de disputa.

Produzir informação alternativa, com os mesmos instrumentos que hoje concentram o poder informacional, é condição para reequilibrar o jogo democrático.

2.2 *Big Data* e IAG numa sociedade democrática plural

A dimensão ética do uso de *Big Data* e inteligência artificial em campanhas eleitorais tem mobilizado crescente atenção na literatura científica. Munger et al. (2024), ao analisarem usuários de WhatsApp no México e na Colômbia, demonstram que a escolaridade, mais do que a familiaridade digital, é o principal fator para o desenvolvimento de pensamento crítico e validação da informação, o que torna vulneráveis os ambientes digitais fechados baseados em confiança interpessoal. Bittencourt e Goltzman (2024), observam que as estratégias sofisticadas de campanha baseadas em grandes volumes de dados são “*capazes de manipular a vontade dos cidadãos e alterar cenários eleitorais*” – extrapolando uma noção de mera influência. Os autores argumentam que o uso de dados pessoais para *microtargeting* político frequentemente ocorre em uma zona cinzenta do ponto de vista ético, onde o consentimento é obtido de forma ambígua ou insuficiente.

Nessa lógica, o que está em jogo não é apenas a eficácia eleitoral, mas a corrosão das condições mínimas de uma democracia pluralista: a possibilidade de acesso equitativo à informação e de formação autôno-

3 Salientamos que não trabalhamos neste texto com o pressuposto que as mídias sociais digitais, as BigTech ou IAGs sejam neutras. Mas reconhecemos que, para além dos seus pressupostos algorítmicos e caixas-pretas do modelo de negócios, as ferramentas que disponibilizam devem ser apropriadas para uma disputa menos desequilibrada nos processos eleitorais democráticos. Para que essa disputa seja equilibrada, é indispensável dominar as técnicas e estruturar um planejamento sólido de campanha e marketing eleitoral, aplicando Big Data e recursos de IAG de forma estratégica, ética e ajustada às condições concretas da disputa.

ma de preferências, como compreendido por Dahl (1997). A literatura aponta para a necessidade de um equilíbrio que permita o desenvolvimento de campanhas eleitorais modernas e eficientes, sem comprometer valores democráticos fundamentais como a privacidade, a autonomia e a equidade no acesso à informação (BARREDO-IBÁÑEZ et al., 2021; BITTENCOURT e GOLTZMAN, 2024; RICHARD, 2024).

Diante deste cenário o desafio é político e estratégico: produzir informação alternativa, com as mesmas ferramentas que hoje concentram o poder informacional, torna-se fundamental para reduzir o desequilíbrio informacional e que a inovação tecnológica não seja sinônimo de assimetria e manipulação, já que a *Big Tec* e a IAG tornaram-se intermediadores políticos na arena pública (SUNSTEIN, 2017; JUNGHERR, SCHROEDER, 2023).

2.3 Marketing Político, *microtargeting* e o uso de *Big Data* e IA Generativa

A crescente sofisticação das ferramentas digitais vem transformando profundamente o marketing político, tornando campanhas mais segmentadas, aparentemente responsivas e moldadas pelo comportamento do público-alvo. Esse fenômeno tem impacto particular em locais onde a popularização das redes sociais consolidou plataformas digitais como arenas centrais da disputa política. Mesmo em contextos com infraestrutura limitada, o uso estratégico de recursos digitais, como memes e formatos otimizados para redes, tem gerado efeitos concretos sobre a percepção pública e decisões de voto (GODOY-GUEVARA et al., 2025).

Nesse cenário, o *microtargeting* político desponta como uma das aplicações mais poderosas do *Big Data*. Baseado em segmentação ultra específica por perfil comportamental e demográfico, ele permite ajustar mensagens em tempo real, com base em modelos preditivos de engajamento, consumo de informação e localização (Álvarez-García et al, 2024). Contudo, a eficácia técnica do *microtargeting* não pode ser dissociada de seus efeitos sobre a esfera pública. O risco é substituir o

espaço comum do debate por múltiplas realidades políticas isoladas, tendo o uso de *Big Data* e IAG como intermediadores na arena pública (JUNGHERR, SCHROEDER, 2023). Além disso, o *microtargeting* aprofunda desigualdades já existentes. Estudos comparativos conduzidos por Bittencourt e Goltzman (2024) mostram que, no Brasil, Chile e México, a adoção dessas técnicas ocorre de forma desigual e favorece campanhas com maior acesso a recursos técnicos e financeiros.

Assim, para enfrentar a desigualdade entre as alternativas políticas, que pode pôr em risco a existência de fontes alternativas de informação, é necessário o fortalecimento da equidade na disputa eleitoral, possibilitando, ao menos, alternativas informacionais. Para isso, um aprendizado prático do uso de *Big Data* e IAG são fundamentais para disputas mais equânimes.

3. *Big Data*, Inteligência Artificial Generativa e Campanhas: um enfoque prático

As seções anteriores demonstraram que, para que a disputa eleitoral ocorra em bases democráticas, é necessário garantir pluralidade informacional no espaço público. Diante desse cenário, esta seção propõe um conjunto de diretrizes práticas para o uso de *Big Data* e IAG no planejamento e execução de campanhas eleitorais. Ressaltamos que a incorporação eficaz de *Big Data* e IAG no planejamento de campanhas eleitorais exige mais do que entusiasmo tecnológico. Trata-se de um processo político e técnico que demanda domínio metodológico, leitura de contexto e capacidade de adaptação narrativa. A seguir, propomos quatro eixos fundamentais para orientar esse uso estratégico, com base em práticas emergentes e recursos já disponíveis.

3.1 Eixo 1: Planejamento e estratégias baseadas em *Big Data*

A geração e coleta de dados têm custo relativamente baixo em comparação ao impacto potencial que oferecem. Ainda assim, muitas

campanhas falham em investir antecipadamente em pesquisas quantitativas e qualitativas, desperdiçando a janela mais estratégica do ciclo eleitoral.

Ferramentas baseadas em modelos de linguagem generativos (LLMs), como o ChatGPT, o [Manus.ai](#) e o [Julius.ai](#), têm ampliado essa capacidade analítica que pode ser útil para análise de dados de pesquisas de opinião pública. Enquanto Manus apresenta melhor desempenho na leitura e cruzamento de dados estruturados, Julius se destaca na organização de dados qualitativos, como anotações de grupos focais e análises de discurso. Usadas estrategicamente, essas ferramentas podem não apenas acelerar diagnósticos eleitorais, mas orientar ajustes táticos em tempo real.

Mesmo que existam poucas plataformas com API (*Application Programming Interface*)⁴ abertas, ferramentas como o #TelegramScrap (SILVA, 2023) e o #YouTubeScrap (SILVA, ROCHA, 2024) são capazes de extrair e planilhar metadados de mensagens, vídeos e interações, oferecendo insumos para análise de engajamento, fluxos de disseminação e padrões temáticos nas redes. Diante das limitações impostas por políticas de acesso restritivo e modelos comerciais fechados, o uso combinado do #TelegramScrap e do #YouTubeScrap se torna especialmente estratégico. Essas duas plataformas oferecem janelas importantes para compreender a construção de autoridade, circulação de desinformação, timing de engajamento e, sobretudo, as formas pelas quais o eleitorado consome e replica conteúdo político.

Assim, a análise de dados extraídos dessas fontes permite traçar perfis de público, identificar nichos de radicalização, e mapear a centralidade de determinados atores ou mensagens: elementos decisivos para orientar tanto no campo da produção audiovisual, ferramentas como Pictory, Synthesia, HeyGen, Lumen5 e D-ID possibilitam a criação rá-

4 *Application Programming Interface* são os conjuntos de padrões e protocolos que permitem a comunicação entre sistemas distintos e são necessárias para acessar dados de plataformas.

pida de vídeos a partir de texto, com narrativas personalizadas e visuais adaptados para diferentes canais e públicos⁵. Não é recomendável o uso de personagens digitais, especialmente em contextos de eleições municipais, onde o contato direto com pessoas conhecidas tem um efeito positivo sobre as intenções de voto (AZEVEDO, SILVA, ALVES, 2020).

Assim, o uso combinado de extração de dados, visualização analítica e geração de informação pertinente permite às campanhas deixar de operar no improvisado ou na intuição. Com uma infraestrutura mínima e treinamento básico, já é possível estruturar centros de inteligência digital em campanhas de médio porte – com ganhos reais em eficiência, assertividade e capacidade de resposta.

3.2 Eixo 2: Conectar a estratégia digital ao perfil da candidatura

Um dos erros mais recorrentes nas campanhas contemporâneas é a reprodução mecânica de formatos digitais bem-sucedidos em outros contextos, como os modelos das campanhas de Jair Bolsonaro (2018) ou Pablo Marçal (2024). Estratégias digitais não são fórmulas replicáveis, e seus efeitos variam conforme o contexto, o posicionamento da candidatura e as demandas expressas nas pesquisas de opinião. Uma leitura orientada por dados, e não por modismos, é essencial para alinhar tática digital e identidade política.

A eficácia comunicacional, especialmente em ambientes digitais altamente voláteis, exige coerência entre linguagem, estética, valores simbólicos e o *ethos* da candidatura. Como adverte Shiller (2019), narrativas econômicas, e por extensão, políticas, se propagam como vírus, dependendo de sua capacidade de ressoar emocional e culturalmente com os públicos. O mesmo se aplica à comunicação eleitoral: uma mensagem só é eficaz se for percebida como autêntica, enraizada nas

5 O uso dessas ferramentas de geração de vídeos ainda carece de regulamentação pela Justiça Eleitoral brasileira.

identidades e expectativas do eleitorado. Essa coerência discursiva exige que as candidaturas compreendam não apenas o que comunicar, mas como e por que determinadas mensagens geram engajamento. A arquitetura algorítmica das plataformas digitais premia a consistência narrativa e simbólica, o que torna contraproducente importar estratégias que não dialogam com os repertórios afetivos e morais dos públicos locais.

3.3 Eixo 3: Sustentação da grande narrativa de campanha

Campanhas bem-sucedidas não se sustentam apenas com posts segmentados ou slogans virais: elas exigem uma grande narrativa central, coerente e emocionalmente potente. Como alertam Figueiredo et al. (1997), essa narrativa articula fatos, afetos e projeções. Ferramentas de IAG são hoje capazes de ajudar na formulação e manutenção dessa narrativa, conectando dados fragmentados à estratégia discursiva central da campanha. Isso se torna ainda mais crucial no atual ambiente de “campanha permanente”, em que lideranças políticas operam como influenciadores digitais (SILVA, 2021; 2023), e a coerência simbólica importa tanto quanto o conteúdo político.

O conceito de “grande narrativa” aqui não se confunde com *storytelling* publicitário ou slogans isolados, mas com uma matriz semântica estável que confere sentido e direção à campanha, permitindo integrar as múltiplas frentes comunicacionais em torno de um horizonte comum. Como já demonstrado por Rocha (2023), campanhas que constroem constelações narrativas consistentes em ambientes digitais – com nós simbólicos fortes, alta coesão e ressonância afetiva – são mais capazes de sustentar engajamento orgânico ao longo do tempo, mesmo frente a crises reputacionais ou tentativas de contranarrativa.

Além disso, em contextos marcados por arrefecimento da polarização ou desmobilização política, como os indícios identificados por Viscarra et al. (2024) para o Brasil pós-2022, a capacidade de manter uma narrativa propositiva, empática e mobilizadora torna-se diferencial

competitivo. A disputa deixa de ser apenas por atenção momentânea e passa a ser pela construção de sentidos duradouros que conectem o eleitorado ao projeto político.

Na comunicação de campanha, a IAG pode atuar como suporte analítico e criativo, identificando padrões discursivos, sugerindo formas de atualização temática sem ruptura de coerência, e testando variações de linguagem e enquadramento que ressoem com diferentes segmentos do eleitorado. Isso é particularmente relevante em contextos de volatilidade informacional, em que a narrativa de campanha precisa se adaptar rapidamente a eventos imprevistos sem perder seu eixo simbólico. Ferramentas como Descript e Runway ML também podem ser empregadas para editar vídeos com base em transcrições automáticas e acelerar a publicação de conteúdos adaptados a eventos em tempo real.

3.4 Eixo 4: Conquistar mentes e corações para construção de conteúdo orgânico

O comportamento do eleitor não é estático. Em 2018, o Brasil assistiu a uma campanha movida por repúdio à política tradicional, polarização e agressividade discursiva. Já em 2022, a disputa foi guiada pelo medo e pela economia (SILVA, TELLES, 2022; LAVAREDA, 2024) e arrefecimento da polarização nas disputas subnacionais (VISCARRA et al, 2024). Para 2026, indícios de arrefecimento nas redes e sinais de desgaste do eleitor frente à polarização indicam que afetos positivos e discursos menos radicais podem mobilizar com mais eficácia. Estratégias baseadas em *Big Data* e IAG devem, portanto, estar atentas às nuances da opinião pública digital e não repetir fórmulas exauridas.

Nesse cenário, a centralidade do conteúdo orgânico ganha nova relevância. A saturação do eleitor diante de campanhas hiper profissionais e excessivamente centradas em impulsionamento pago abre espaço para estratégias baseadas em engajamento genuíno, proximidade afetiva e participação ativa. Em vez de apenas receber mensagens, o eleitor precisa se ver como parte da campanha: e mais do que isso,

como seu porta-voz. Trata-se de uma mudança de chave: da campanha que fala para o eleitor, para a campanha que é falada pelo eleitor (MOSER, 2023).

Big Data e IAG podem ajudar a identificar perfis com alto potencial de influência em suas redes locais, comunidades e grupos segmentados, inclusive fora das redes sociais tradicionais, como em aplicativos de mensagens privadas. Além disso, essas tecnologias permitem adaptar conteúdos e linguagens específicas a microgrupos com maior probabilidade de engajamento ativo, ampliando o alcance da campanha de maneira mais legítima e menos intrusiva. Campanhas com forte aderência emocional, capazes de mobilizar orgulho, pertencimento e entusiasmo, geram conteúdo orgânico mais consistente e duradouro. Esse tipo de comunicação horizontalizada e espontânea tende a ser mais confiável e mais eficaz em contextos de desinformação, pois atua com base em laços de confiança interpessoal, o que é particularmente relevante em plataformas como WhatsApp ou Telegram.

Conquistar mentes e corações, portanto, não é uma metáfora, mas uma estratégia baseada em evidências: é ao ativar laços afetivos e permitir que o eleitor se torne produtor de sentido, e não apenas consumidor de mensagens, que campanhas ganham capilaridade, autenticidade e resistência às dinâmicas de crise.

Considerações finais

Este artigo partiu de uma provocação prática: como incorporar *Big Data* e ferramentas de IAG ao planejamento e execução de campanhas eleitorais de forma estratégica e ética, respeitando os princípios democráticos e a realidade latino-americana. Para responder a essa questão, articulamos uma base teórica centrada na teoria da poliarquia de Robert Dahl, destacando a importância das fontes alternativas de informação para a manutenção de um ecossistema democrático plural. Discutimos os dilemas éticos do uso de dados em sociedades marcadas por desigualdades estruturais, o impacto da hiper persona-

lização informacional e os riscos de novas formas de manipulação da vontade política.

Diante desse cenário, propusemos quatro eixos práticos para o uso estratégico de *Big Data* e IAG em campanhas eleitorais, com foco em: (1) planejamento orientado por dados, (2) alinhamento entre identidade da candidatura e tática digital, (3) sustentação da grande narrativa e (4) construção de conteúdo orgânico e afetivo.

Mais do que oferecer um roteiro técnico, o objetivo deste artigo foi apresentar uma leitura crítica e aplicada das transformações em curso, indicando que a disputa democrática não se limita ao conteúdo das propostas, mas também envolve as condições pelas quais a informação circula, se cristaliza e se converte em ação política. O que está em jogo é a própria estrutura do debate público e a capacidade das candidaturas de disputar sentidos num ambiente cada vez mais assimétrico, opaco e polarizado.

Big Data e IAG não são, por si só, ameaças ou soluções. São meios, e como tais, dependem da forma como são apropriados. Usá-los com consciência, clareza estratégica e compromisso democrático não é apenas uma opção técnica: é uma decisão política. É dessa escolha que dependerá a qualidade do debate público que seremos capazes de sustentar, e, em última instância, o tipo de democracia que construiremos.

Referências Bibliográficas

AMAYA LÓPEZ, Carlos Alberto; CUEVA GAIBOR, Diego Abraham. Inteligencia artificial y comunicación política en campañas electorales: mirada crítica, implicaciones y desafíos. **Revista Social Fronteriza**, v. 5, n. 2, p. 1-19, 2025.

BARREDO-IBÁÑEZ, Daniel; DE-LA-GARZA-MONTEMAYOR, Daniel J.; TORRES-TOUKOUMIDIS, Ángel; LÓPEZ-LÓPEZ, Paulo C. Artificial intelligence, communication, and democracy in Latin America: a review of the cases of Colombia, Ecuador, and Mexico. **El Profesional de la información**, v. 30, n. 6, 2021.

BITTENCOURT, Nathália Viviani; GOLTZMAN, Elder. Novas dinâmicas, novas regras: a propaganda eleitoral digital e o seu marco regulatório em perspectiva comparada. **Estudos Eleitorais**, Brasília, DF, v. 18, n. 1, p. 263-286, jan./jun. 2024.

DAHL, Robert A. **Democracy and its critics**. New Haven: Yale University Press, 1989.

DAHL, Robert A. **Poliarquia**: participação e oposição. São Paulo: Edusp, 1997.

DAHL, Robert A. **Sobre a democracia**. Brasília, DF: Editora Universidade de Brasília, 2001.

DOURADO, Tatiana; MARTINS, Bruna; SANTOS, Nina; CARREIRO, Rodrigo; VRECHE, Carla. Crise de equidade na governança das plataformas digitais no mega ciclo de 2024 e nas eleições municipais no Brasil. **Cadernos Adenauer XXV**, 2024, n. 1, Eleições Municipais 2024.

GODOY-GUEVARA, Gandhi; BASANTES-ANDRADE, Andrea; NARANJO-TORO, Miguel; ORTIZ-DÁVILA, David. The Voter and the Propagandistic Content on Facebook: Analysis of the Content and Perception of the Ecuadorian Voter. **Societies**, v. 15, n. 2, p. 41, 2025.

JUNGHER, Andreas; SCHROEDER, Ralph. Artificial intelligence and the public arena. **Communication Theory**, [s. l.], v. 00, p. 1-10, 2023.

LAVAREDA, Antonio. **De Bolsonaro a Lula 3**: pesquisas, eleição, democracia e governabilidade. Sagga, 2024.

MOSER, Fabricio. **Manual de Mobilização Política** (Para campanhas eleitorais e todas as outras também). Motive Editora, 2023.

MUNGER, Kevin et al. “Reenviado Muitas Vezes”: How Platform Warnings Affect WhatsApp Users in Mexico and Colombia. **Political Communication**, v. 41, n. 1, p. 1–20, 2024.

RICHARD, Eugénie. IAG (IAG). In: Crespo Martinez et al. **Diccionario Enciclopédico de Polarización y Emociones**. Madrid: CEPC/ALICE, 2024.

ROCHA, Isabela Silveira. Digital Identities and Constellation Narratives: A Topological Data Analysis model tested on Brazil’s 2022 electoral campaign. In: **WORLD CONGRESS OF POLITICAL SCIENCE, 27., 2023, Buenos Aires**. Anais [...]. Buenos Aires: IPSA, 2023.

SHILLER, Robert. **Narrative Economics: How Stories Go Viral & Drive Major Economic Events**. Princeton University Press, 2019.

SILVA, Joscimar S. **Surfando na crise de representação e nos valores**: lideranças políticas emergentes e mídias sociais digitais na América Latina. 2021. Tese (Doutorado em Ciência Política) – Programa de Pós-Graduação em Ciência Política, Universidade Federal de Minas Gerais, Belo Horizonte, 2021.

SILVA, Joscimar; TELLES, Helcimara. Una campaña de voto en contra del enemigo: esperanza y miedo en las elecciones de 2022 en Brasil. **La Revista de ACOP**, v. 75, p. 23-25, 2022.

SILVA, Joscimar S. Os mandatos da antipolítica: líderes políticos digitais e a representação política na América Latina. **Revista Compolítica**, v. 13, p. 171-196, 2023.

SILVA, Ergon. **TelegramScrap: A comprehensive tool for scraping Telegram data**. (feb) 2023. Available at: <https://doi.org/10.48550/arXiv.2412.16786>.

SILVA, Ergon; ROCHA, Isabela. **YouTubeScrap: A comprehensive tool for scraping YouTube data and transcript**. (Dez) 2024. Available at: <https://github.com/ergoncugler/web-scraping-youtube/>.

STATISTA. **Leading social media platforms in Brazil 2024, by reach. 2024**. Disponível em: <https://www.statista.com/statistics/1307747/social-networks-penetration-brazil/>. Acesso em: 2 jun. 2025.

SUNSTEIN, Cass R. **#Republic: Divided Democracy in the Age of Social Media**. Princeton: Princeton University Press, 2017.

VISCARRA, Simone; DE ALMEIDA, Helga;; SILAME, Thiago; SILVA, Joscimar. O arrefecimento da polarização afetiva: análise dos discursos dos governadores no Twitter na campanha eleitoral de 2022. **Caderno CRH** (Online), v. 37, p. 1-21, 2024.

Joscimar Souza Silva · Professor no Instituto de Ciência Política da Universidade de Brasília (IPOL – UnB). Doutor em Ciência Política pela Universidade Federal de Minas Gerais (UFMG). Compõe a diretoria da Associação Brasileira de Pesquisadores Eleitorais (ABRAPEL) e coordena o Grupo de Pesquisa Comunicação Política e Comportamento Eleitoral na América Latina da Associação Latinoamericana de Ciência Política (LATICOM – ALACIP) e o Grupo Informação Pública e Eleições (IPê) na UnB. E-mail: joscimar.silva@unb.br.

Isabela Rocha · Doutoranda em Ciência Política e pesquisadora do Grupo IPê na UnB. E-mail: isabelarocha.contato@gmail.com.

Lógicas que excluem: inteligência artificial, gênero e os limites da democracia

Patricia Rangel
Jéssica Melo Rivetti
Flávia Rios

Resumo

O avanço acelerado do uso de *Big Data* e da inteligência artificial (IA) tem transformado profundamente as democracias contemporâneas. Embora essas tecnologias ofereçam oportunidades inéditas para a personalização de serviços, análise de tendências e automação de processos, elas também trazem desafios éticos importantes. Um dos mais críticos é o viés algorítmico – uma reprodução automatizada de desigualdades sociais historicamente construídas. Este artigo discute como os sistemas baseados em *Big Data* podem reforçar desigualdades de gênero e comprometer direitos fundamentais, especialmente os das mulheres. A partir da análise de estudos recentes e de casos emblemáticos – como as *deepfakes* –, argumenta-se que o uso indiscriminado de dados “neutros” pode ter efeitos profundos na exclusão simbólica e material de grupos historicamente marginalizados. O artigo propõe uma abordagem crítica e interseccional para o desenvolvimento e a regulação dessas tecnologias, destacando o papel estratégico da participação feminina nos campos de ciência e tecnologia como condição para a justiça algorítmica e a equidade democrática.

Abstract

The rapid advancement of Big Data and artificial intelligence (AI) has profoundly transformed contemporary democracies. Although these technologies offer unprecedented opportunities for personalization of services, trend analysis, and process automation, they also pose significant ethical challenges. One of the most critical is algorithmic bias—an automated reproduction of historically constructed social inequalities. This article discusses how systems based on Big Data can reinforce gender inequalities and compromise fundamental rights, especially those of women. Based on an analysis of recent studies and emblematic cases – such as deepfakes – it argues that the indiscriminate use of “neutral” data can have profound effects on the symbolic and material exclusion of historically marginalized groups. The article proposes a critical and intersectional approach to the development and regulation of these technologies, highlighting the strategic role of female participation in the fields of science and technology as a condition for algorithmic justice and democratic equity.

1. Introdução

A crescente centralidade dos dados no funcionamento de governos, empresas e organizações sociais transforma a forma como decisões são tomadas e políticas públicas são elaboradas. A era do *Big Data* – caracterizada pelo uso massivo de dados estruturados e não estruturados, processados por tecnologias de inteligência artificial (IA) – tem sido apresentada como promessa de eficiência, precisão e inovação. No entanto, o entusiasmo com essas inovações muitas vezes negligencia uma questão fundamental: os dados não são neutros.

Sistemas automatizados de tomada de decisão, algoritmos de recomendação e ferramentas de tradução automática, entre outros, reproduzem – e, em alguns casos, amplificam – desigualdades históricas de gênero, raça, classe, orientação sexual e outras dimensões da vida social,

sobretudo, em regiões marcadas pelas assimetrias econômicas e de poder como a América Latina¹. A dependência de bases de dados enviesadas, associada à opacidade dos processos algorítmicos e à baixa diversidade das equipes técnicas, resulta em tecnologias que perpetuam estereótipos e comprometem os direitos de mulheres e meninas, especialmente aquelas em situação de interseccionalidade de vulnerabilidades.

Este artigo examina como os sistemas baseados em *Big Data* operam a partir de vieses de gênero, afetando diretamente os direitos das mulheres e, de forma mais ampla, os fundamentos democráticos das sociedades. A partir de uma perspectiva crítica e feminista, propõe-se uma reflexão sobre as condições de produção desses sistemas, seus efeitos na reprodução das representações sociais e suas implicações para a cidadania de mulheres em contextos altamente digitalizados.

2. *Big Data*, IA e viés de gênero

A promessa da IA de resolver problemas complexos, personalizar serviços e otimizar processos repousa, em grande parte, na análise de grandes volumes de dados – o chamado *Big Data*. Esse termo se refere não apenas ao volume massivo de informações disponíveis, mas também à velocidade com que são processadas e à variedade de fontes utilizadas, como redes sociais, sensores, textos, imagens e registros administrativos. No entanto, ao contrário do que muitas vezes se assume, os dados não são meramente técnicos ou objetivos: eles carregam as marcas do contexto social, histórico e político em que foram produzidos².

-
- 1 NERI, Hugo; DOMINGOS, Veridiana. Navigating the Double Divide: Generative AI and the Dynamics of Inequality in Latin America. **ProtoSociology**, v. 40, p. 342-374, 2023.
 - 2 DOMINGUES, Jonathan Machado. Identidade de gênero e Inteligência Artificial: desafios e preconceitos nos algoritmos. In: DOMINGUES, Jonathan Machado (Org.). **Ressonâncias de luta: músicas, gênero e direito no Brasil contemporâneo**. São Paulo: Editora Akademy, 2025. p. 23-31.

Os sistemas de IAs, em especial os baseados em aprendizado de máquina (*machine learning*) e em redes neurais artificiais, aprendem com exemplos anteriores. Isso significa que algoritmos são treinados com conjuntos de dados que refletem decisões humanas passadas – frequentemente marcadas por desigualdades estruturais. Tanto é que se os dados de treinamento contêm discriminações de gênero, raça ou classe, o sistema replicará e, em alguns casos, até amplificará esses padrões³.

Algoritmos não apenas ignoram a diversidade de gênero, mas, frequentemente, reforçam normas binárias e excludentes que marginalizam as identidades que não se encaixam nos moldes estabelecidos de masculinidade e feminilidade. Esse fenômeno, longe de ser uma falha técnica isolada, é resultado de um viés epistemológico que nasce da ontologia moderna, que vê o gênero como uma categoria fixa e reduzida⁴.

Outros estudos demonstram, por exemplo, que tradutores automáticos tendem a atribuir o gênero masculino a profissões consideradas de maior prestígio, como “engenheiro” ou “cientista”, mesmo quando o gênero do sujeito não está explícito no idioma original. Por outro lado, funções tradicionalmente femininas são associadas automaticamente a termos como “enfermeira” ou “secretária”. Esses vieses não são falhas aleatórias: são o reflexo de uma cultura estruturada por estereótipos, internalizados por sistemas que aprendem com o comportamento humano em larga escala. Esses estereótipos de gênero serão, analiticamente, referidos como imagens de controle (PATRÍCIA COLLINS, 2019), porque servem a um certo regime de poder. Não se trata de mera naturalização de certos perfis e padrões sociais, mas de enquadramentos interpretativos que fixam as mulheres em certos papéis sociais, num exercício de violência simbólica de gênero.

3 SILVA, Tarcízio. **Racismo Algorítmico**: inteligência artificial e discriminação nas redes digitais. São Paulo: Edições Sesc, 2022.

4 DOMINGUES, op. cit., p. 29.

Além da reprodução de estereótipos, o viés de gênero nos sistemas de IA tem implicações práticas diretas. Em plataformas de contratação automatizada, por exemplo, candidaturas femininas podem ser desvalorizadas por algoritmos treinados com históricos que privilegiam currículos masculinos. Em ferramentas de crédito, mulheres podem ser consideradas menos confiáveis devido a padrões históricos de acesso desigual ao sistema financeiro. O resultado é a manutenção de um ciclo de exclusão, agora travestido de neutralidade tecnológica.

Cabe destacar que o viés algorítmico não é apenas um problema técnico, mas político⁵. Como bem aponta Kate Crawford, “não se trata apenas de consertar os algoritmos, mas de compreender as estruturas sociais que os moldam”⁶. Portanto, qualquer esforço de correção desses sistemas deve ir além da revisão técnica e incorporar um olhar interseccional, que considere como gênero, raça, classe e outras categorias se entrecruzam na produção e uso dessas tecnologias.

3. Impactos para os direitos das mulheres e para a democracia

A presença de vieses de gênero nos sistemas de *Big Data* e IAs não se restringe ao campo simbólico. Suas consequências são concretas e afetam diretamente os direitos civis, políticos, econômicos e culturais das mulheres. À medida que algoritmos passam a mediar decisões em áreas estratégicas – como recrutamento, crédito, justiça criminal, saúde e comunicação pública —, torna-se urgente refletir sobre como esses sistemas contribuem para a (re)produção de desigualdades.

Nesse sentido, o artigo intitulado *Navigating the Double Divide: Generative AI and the Dynamics of Inequality in Latin America* contribui

5 DOMINGUES, op. cit.

6 CRAWFORD, Kate. **Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence**. New Heaven: Yale University Press, 2021.

com reflexões sobre as implicações sociais da IA na América Latina⁷. O argumento dos autores é que ao mesmo tempo que as ferramentas tecnológicas da IA podem abrir diversos caminhos em direção a uma educação mais inclusiva na região, pode também acentuar desigualdades já existentes pode marginalizar ainda mais grupos já desfavorecidos: “aqueles com melhor acesso à internet, maior alfabetização digital e meios financeiros para acessar tecnologias de IA podem obter vantagens substanciais em educação, emprego e empreendimentos criativos (KORINEK & STIGLITZ, 2021)”⁸.

Entre os principais riscos presentes na amplificação do uso dessas tecnologias está o que a literatura chama de danos representacionais e danos alocativos (CRAWFORD, 2017). Os primeiros dizem respeito à forma como determinados grupos são representados – ou não – nos sistemas digitais. No caso das mulheres, isso pode se traduzir na perpetuação de estereótipos, na masculinização de profissões valorizadas, na invisibilização de identidades não normativas e na exclusão de marcadores interseccionais, como raça, etnia ou identidade de gênero não binária⁹ (SILVA, 2025; DOMINGOS, 2025). Já os danos alocativos envolvem a distribuição desigual de recursos e oportunidades: sistemas enviesados podem negar um crédito, desclassificar uma candidatura ou invisibilizar denúncias de violência.

Como aponta o livro *Racismo Algorítmico* (2022), esses efeitos estruturam uma nova forma de opressão, agora mediada por tecnologias opacas e de difícil contestação¹⁰. Se antes era possível identificar um agente discriminador – um empregador, um juiz, um atendente —,

7 NERI, DOMINGOS, op. cit.

8 *Ibidem*, p. 343; tradução livre.

9 COSTA, Lucas Murakami Rocha. **Um framework para análise de discurso transfóbico a partir de técnicas de aprendizado de máquina**. 2022. Trabalho de conclusão de curso de graduação – Instituto de Computação da Universidade Federal do Rio de Janeiro. Disponível em: <http://hdl.handle.net/11422/20459>. Acesso em: 8 jun. 2025.

10 SILVA, op. cit.

hoje, a responsabilização se dilui entre dados históricos, decisões automatizadas e empresas que alegam neutralidade técnica. Como resultado, mulheres e meninas, especialmente as negras, indígenas, migrantes, LGBTQIA+ e com deficiência, enfrentam barreiras adicionais para o exercício de seus direitos.

Na esfera política, os impactos são igualmente preocupantes. O uso de dados e algoritmos para campanhas eleitorais personalizadas pode reforçar estereótipos de gênero¹¹, excluir candidaturas femininas de segmentos estratégicos do eleitorado ou até mesmo amplificar discursos de ódio¹². A violência política contra as mulheres também se intensifica no ambiente digital¹³, com o uso de tecnologias como *deepfakes* para desacreditizar mulheres públicas, conforme apontam estudos recentes da União Europeia e da ONU Mulheres. Contudo, como destaca Mariana Valente: “não existem estatísticas oficiais suficientes para compreender o tamanho do problema”¹⁴, uma vez que há muita subnotificação dos casos e que dados importantes como orientação sexual, raça e outros não são registrados ou registrados adequadamente.

A desinformação baseada em gênero opera como uma das faces contemporâneas da violência contra as mulheres, especialmente quando associada a tecnologias emergentes como algoritmos de ranqueamento e geração automatizada de conteúdo. O guia desenvolvido pelo InternetLab e STF e publicado em 2025, indica que essas práticas se in-

11 GATTO, Malu; THOMÉ, Débora. **Candidatas**: Os primeiros passos das mulheres na política no Brasil. Rio de Janeiro: FGV Editora, 2024.

12 CABRAL, Leandro Freire. **Detecção de discurso de ódio transfóbico na rede social X mediante técnicas de aprendizado de máquina**. 2024. Monografia apresentada à Universidade Federal Rural do Semi-Árido. Disponível em: <https://repositorio.ufersa.edu.br/handle/prefix/12401>. Acesso em: 08/06/2025.

13 KROOK, Mona Lena. 2022. Semiotic violence against women: Theorizing harms against female politicians. **Signs: Journal of Women in Culture and Society**, v. 47, n. 2, p. 371-397. Disponível em: <https://www.journals.uchicago.edu/doi/epdf/10.1086/716642>. Acesso em: 10 jun. 2025.

14 VALENTE, Mariana. **Misoginia na internet**. São Paulo: Editora Fósforo, 2023.

tensificam em períodos eleitorais e afetam diretamente a atuação política, a saúde mental e a segurança de mulheres públicas e ativistas¹⁵.

Mais ainda, a exclusão simbólica e material das mulheres dos campos de ciência, tecnologia, engenharia e matemática (STEM) limita sua capacidade de influenciar os rumos do desenvolvimento tecnológico. Como mostram os dados da UNESCO, apenas 22% dos profissionais de IA no mundo são mulheres¹⁶. No Brasil, a Pesquisa Nacional por Amostra de Domicílios (PNAD) divulgada em 2009 mostra que apenas 18,84% dos profissionais de TI eram mulheres. Mais de uma década depois, em 2022, esse número chega a 20%, ou seja, um crescimento pequeno de apenas 1,16%. Isso considerando que o perfil feminino corresponde a 52,52% da população e considerando que a área de tecnologia é uma área promissora no mercado de trabalho, por ter potencial de crescimento de vagas e valorização dos salários¹⁷. Sem participação feminina qualificada e significativa nessas áreas, a própria concepção dos sistemas tecnológicos tendem a ignorar ou amplificar as desigualdades de gênero, perpetuando uma lógica excludente sob o disfarce da inovação.

Esses efeitos combinados ameaçam o próprio ideal democrático de igualdade, participação e transparência. Uma democracia que delega decisões relevantes a sistemas opacos e enviesados enfraquece os mecanismos de controle social, a justiça distributiva e a confiança pública. Para que as tecnologias emergentes contribuam para o aprofundamento democrático, é imprescindível garantir sua construção e regulação a partir de uma perspectiva de direitos humanos, com atenção especial às desigualdades de gênero.

15 CINTRA, Anna Marta; VILELA, Catharina; BORGES, Ester; TAVARES, Clarice; MARTINS, Fernanda. **Desinformação: uma das dimensões da violência de gênero**. São Paulo, 2024.

16 UNESCO. **The Effects of AI on the Working Lives of Women**. Paris: UNESCO, 2021.

17 Dados da Associação Brasileira das Empresas de Tecnologia da Informação e Comunicação (Brasscom) indicou que até 2025 o setor vai demandar 797 mil profissionais e ressalta que pouco mais de 53 mil profissionais de tecnologia são formados por ano.

4. As *deepfakes* como violência política contra mulheres

Nos últimos anos, as tecnologias de manipulação audiovisual baseadas em IAs, conhecidas como *deepfakes*, passaram a ocupar lugar de destaque nos debates sobre segurança digital, desinformação e democracia. Trata-se de arquivos de imagem ou vídeo alterados por meio de redes neurais artificiais – capazes de simular com alto grau de verossimilhança expressões faciais, movimentos e falas —, gerando conteúdos falsos que são, à primeira vista, indistinguíveis da realidade.

Em junho de 2025, em um colégio em Belo Horizonte, Minas Gerais, 17 estudantes denunciaram que tiveram suas imagens adulteradas por IA, criando conteúdo pornográfico. Essas imagens foram amplamente compartilhadas e, inclusive, comercializadas em um grupo de mensagens do Telegram¹⁸. Este caso recente joga luz sobre como as novas tecnologias de informação e comunicação podem ser usadas como ferramentas para práticas de ciberviolência e cyberbullying¹⁹ – nova modalidade do bullying que ultrapassa as fronteiras das escolas e ganha terreno fértil no campo digital ao transmitir uma sensação de impunidade pelo “anonimato”. Nesse sentido, essa forma de bullying pode ocorrer por meio da “difusão de informações lesivas ou difamatórias em formato eletrônico [...] através de telefones ou dispositivos móveis ou a publicação de vídeos e fotografias em plataformas eletrônicas de difusão de conteúdo”²⁰.

18 ZUBA, Fernando, GURGEL, Luís. Estudantes denunciam que tiveram imagens íntimas criadas por IA e compartilhadas em colégio de BH. **G1**. 04 de junho de 2025. Disponível em: <https://g1.globo.com/mg/minas-gerais/noticia/2025/06/04/estudantes-manipulacao-vazamento-imagens-intimas-ia-colegio-bh.ghtml>. Acesso em: 05 jun. 2025.

19 Essa modalidade de violência também pode ser denominada de ciber assédio, assédio cibernético, eletrônico, digital etc.

20 GARAIGORDOBIL, Maite. Prevalencia y consecuencias del cyberbullying: una revisión. **International journal of psychology and psychological therapy**, v. 11, n. 2, 2011. p. 235.

O exemplo evidencia o risco das tecnologias serem utilizadas para atacar publicamente meninas e crianças menores de idade, pessoas vulneráveis mas também, mulheres em espaços de poder ou visibilidade. Nesses casos, as IAS são usadas para “a criação e disseminação de imagens e vídeos íntimos de mulheres, que constitui uma forma de violência de gênero, visto que são usadas para constranger, humilhar”²¹. Isto é, as ferramentas tecnológicas são usadas para reforçar a violência política de gênero por meio da difamação, erotização não consentida e destruição de reputações²².

A produção de *deepfakes* sexuais contra mulheres com atuação política, acadêmica ou jornalística tem se tornado uma tática recorrente de intimidação. O guia *Desinformação: uma das dimensões da violência de gênero* relata como imagens íntimas ou manipuladas, muitas vezes circuladas em massa durante períodos eleitorais, são usadas para associar candidatas a escândalos, “fraudes morais” ou acusações criminosas infundadas²³. Esses conteúdos operam com base em narrativas falsas, mas emocionalmente mobilizadoras, muitas vezes reforçadas por campanhas coordenadas que exploram o algoritmo das redes sociais para garantir ampla circulação.

O impacto dessas ações é profundo: a violência por *deepfake* combina os efeitos psicológicos da exposição e da humilhação com as consequências políticas da descredibilização e do silenciamento. Nesse sentido, estudos desenvolvidos pelo Instituto Marielle Franco mostram que 98,5% das candidatas negras nas eleições municipais de 2020 – que mesmo com a política de ações afirmativas, ainda são minoria disputando cargos eletivos de maior prestígio²⁴ – relataram ter sofrido algum tipo de

21 CINTRA et al., op. cit., p. 16.

22 EIGE – European Institute for Gender Equality. **Combating Cyber Violence Against Women and Girls: Developing an EU Measurement Framework**. Publications Office of the European Union, 2025.

23 CINTRA et al., op. cit.

24 GATTO, THOMÉ, op. cit. p. 30.

violência política²⁵. Entre as violências, os ataques virtuais tiveram destaque, para as denúncias de propagação de imagens falsas ou manipuladas. Estes ataques articulam racismo, etarismo, classicismo, sexismo e lesbofobia – como no caso citado no relatório do InternetLab, em que uma deputada negra e lésbica teve sua imagem associada a um suposto relacionamento com um chefe do tráfico, gerando ameaças de morte e difamação sistemática²⁶.

As *deepfakes* são instrumentalizadas como armas de violência política e podem ser entendidas como uma expressão da violência semiótica, cuja eficácia se deve à lógica algorítmica das plataformas digitais. Segundo Mona Lena Krook, essa violência “não é apenas parte de um *continuum* mais amplo de atos violentos, mas também a forma mais disseminada, oculta e trivializada, de violência contra as mulheres.”²⁷. Tanto é que conteúdos com alto potencial de engajamento – ainda que falsos ou ofensivos – são impulsionados automaticamente, ampliando estrategicamente o alcance da desinformação aos usuários e dificultando as tentativas de resposta²⁸.

A possibilidade dos ataques virtuais deixou as candidatas muito mais vulneráveis. Isso porque, com meia dúzia de palavras e um clique, alguém pode atacá-las diretamente, com ofensas e xingamentos chegando em suas caixas de mensagem pessoais. Isso aumenta o dano psicológico de forma exponencial [...] ²⁹.

25 INSTITUTO MARIELLE FRANCO. **Violência Política de Gênero e Raça no Brasil:** Eleitas ou não, mulheres negras seguem desprotegidas. Rio de Janeiro, 2021. 93p.

26 Escusado é dizer que Marielle Franco também viveu um processo de difamação virtual após seu assassinato. Na tentativa de torná-la descredibilizada, Marielle foi associada à imagem de uma mulher que tinha relações com um famoso traficante de drogas do Rio de Janeiro. O objetivo flagrante dessa estratégia política foi desqualificar o argumento de que sua morte se tratava de um feminicídio político.

27 KROOK, op. cit., p. 372; tradução livre.

28 SILVA, op. cit.

29 GATTO, THOMÉ, op. cit., p. 19.

Como aponta Kate Crawford, “a aparência de neutralidade algorítmica mascara uma arquitetura de poder” que reproduz assimetrias sociais³⁰. No caso da violência digital com perspectiva de gênero, essa arquitetura legitima ataques como se fossem entretenimento, opinião ou liberdade de expressão, esvaziando a gravidade dos crimes e isolando as vítimas. Ou seja, essas formas de violência, incluindo a semiótica, “servem como uma ferramenta para negar às mulheres o direito pleno e igualitário de participar da política, minando tanto a democracia quanto a igualdade de gênero”³¹.

A ausência de respostas eficazes por parte das plataformas e das instituições públicas agrava ainda mais o cenário. Embora existam legislações em construção – como a Diretiva (UE) 2024/1385 sobre violência contra as mulheres e violência doméstica, que menciona o uso de tecnologias digitais na perpetuação da violência —, ainda são raros os instrumentos normativos e jurídicos capazes de enfrentar o uso de *deepfakes* sob a perspectiva dos direitos humanos e da justiça de gênero. No Brasil, embora a Lei nº 14.192/2021 represente um avanço ao tipificar a violência política contra as mulheres, ela ainda carece de mecanismos de aplicação específicos para a esfera digital³².

Diante disso, é urgente reconhecer os *deepfakes* como uma nova forma de arma política e digitalizada utilizada por narrativas misóginas para práticas de violência on-line. Para enfrentá-la, é necessário combinar regulação tecnológica com estratégias de educação midiática e alfabetização digital crítica. Também é essencial apoiar iniciativas de monitoramento, denúncia e acolhimento das vítimas – garantindo não apenas respostas punitivas, mas reparação simbólica, reconhecimento político e reconstrução da confiança. Como destaca o guia do InternetLab, “a desinformação de gênero afeta negativamente o deba-

30 CRAWFORD, op. cit.

31 KROOK, op. cit., p. 372.

32 BRASIL, **Lei nº 14.192**, de 4 de agosto de 2021. Disponível em: https://www.planalto.gov.br/ccivil_03/ato2019-2022/2021/lei/l14192.htm. Acesso em: 10 jun. 2025.

te público digital, pois, em um ambiente hostil, as pessoas podem se sentir inseguras para compartilhar seus pensamentos e opiniões”³³. É necessário que o enfrentamento a essa violência seja, portanto, parte fundamental da agenda por uma democracia informacional feminista e plural³⁴, com discussões sobre governança digital e com a criação de um ecossistema digital (de aprendizagem das máquinas e IAs) mais crítica, logo, mais inclusiva para diferentes grupos sociais³⁵.

5. Possibilidades de transformação

Não é raro que sistemas tecnológicos como softwares e demais tecnologias emergentes sejam introduzidos ao mercado sem antes passar por avaliações adequadas, o que pode intensificar, muitas vezes, preconceitos estruturais³⁶ – como a associação de fenótipos negros com supostos comportamentos criminosos em IAs de reconhecimento facial ou então, o embranquecimento da pele em filtros de “beleza” em aplicativos como FaceApp e redes sociais como Snapchat, Tiktok e Instagram. O que reforça padrões discriminatórios de “racismo algorítmico”³⁷. Esta realidade denunciada indica que recursos mal desenvolvidos têm grande impacto social, servindo como uma espécie de ferramenta que atua no controle social de grupos historicamente marginalizados.

Diante desses múltiplos desafios éticos apresentados por esses sistemas de *Big Data* e IAs, é fundamental construir alternativas que promovam a equidade e a justiça, entendendo que é preciso “fazer política em cenários e esferas que se dão no *continuum* on-line e off-line”³⁸. Para isso, não basta apenas corrigir os algoritmos – é necessário transformar

33 CINTRA et al., 2024.

34 MARTINS, Fernanda K. Ciberfeminismo. In: GALETTI, Camila; RIVETTI, Jéssica. **Feminismos em Movimento**. Belo Horizonte: Editora Luas, 2023. p. 36-45.

35 VALENTE, op. cit.

36 SILVA, op. cit.

37 *Ibidem*.

38 MARTINS, op. cit., p. 43.

as estruturas que os sustentam. Essa transformação deve passar por três frentes principais: (i) regulação; (ii) diversidade nas equipes técnicas; (iii) participação social no desenvolvimento tecnológico.

A primeira delas envolve a regulação democrática das tecnologias emergentes. Não se trata de frear a inovação, mas de estabelecer limites éticos e jurídicos que impeçam abusos e garantam a proteção de direitos. A União Europeia tem dado passos importantes nesse sentido, com normas como o Digital Services Act e, mais recentemente, a Diretiva (UE) 2024/1385 sobre violência contra as mulheres e violência doméstica, que inclui obrigações específicas sobre a coleta de dados relativos à violência cibernética baseada em gênero³⁹. Contudo, essas iniciativas ainda precisam ser acompanhadas por medidas eficazes de implementação e fiscalização, especialmente nos países do Sul Global.

A segunda frente diz respeito à inclusão efetiva de mulheres – e, em especial, de mulheres negras, indígenas, LGBTQIA+ e com deficiência – nos campos de ciência, tecnologia, engenharia e matemática (STEM). Sem a presença ativa desses grupos na concepção, desenvolvimento e avaliação dos sistemas de IA, os riscos de exclusão e reprodução e amplificação de estereótipos permanecem altos⁴⁰. Programas de incentivo à formação e permanência de meninas em STEM, bem como ações afirmativas para a contratação em empresas de tecnologia, são passos fundamentais para reverter esse quadro⁴¹.

39 Diretiva (UE) 2024/1385 do Parlamento Europeu e do Conselho, de 14 de maio de 2024, relativa ao combate à violência contra as mulheres e à violência doméstica. Disponível em: <http://data.europa.eu/eli/dir/2024/1385/oj>. Acesso em: 10 jun. 2025.

40 VALENTE, op. cit.

41 Um exemplo expressivo de iniciativas nessa direção é o da Rede Nacional de Educação e Extensão Meninas Digitais (ReNacEE_MD) que visa despertar o interesse de meninas e mulheres para as áreas de Computação. Essa rede é importante porque trabalha desde a base, ou seja, desde a educação pública básica propondo metodologias de conscientização e de empoderamento das estudantes e professoras/as sobre os desafios relacionados ao mundo das tecnologias digitais, levando em conta o enfrentamento do viés de gênero e outras desigualdades interseccio-

A terceira frente é o fortalecimento de uma perspectiva feminista (ciberfeminista), interseccional e decolonial no debate público sobre tecnologia⁴². Organizações da sociedade civil, universidades, coletivos de mulheres e movimentos sociais têm desempenhado um papel crucial na denúncia dos abusos, na construção de indicadores de impacto e na proposição de alternativas. O reconhecimento desses atores como interlocutores legítimos nas arenas de regulação e inovação é essencial para democratizar os rumos da transformação digital.

Além disso, o fomento à alfabetização crítica em dados e tecnologia deve ser uma prioridade para os sistemas educacionais e de formação continuada.⁴³ Usuárias e usuários precisam compreender como os algoritmos funcionam, de onde vêm os dados e quais as consequências de seu uso, especialmente nos casos em que há impactos diretos sobre os direitos civis. Nesse sentido, a construção de uma cidadania digital plena depende da capacidade crítica da população em interagir com sistemas que operam, cada vez mais, como mediadores da vida social.

Por fim, é preciso reconhecer que a tecnologia é sempre política. Não existem soluções puramente técnicas para problemas profundamente enraizados nas desigualdades sociais. Portanto, qualquer projeto que pretenda construir uma IA mais justa e um uso mais ético dos dados deve estar comprometido com a transformação das relações de poder, com a reparação histórica e com a promoção ativa da igualdade de gênero e de justiça social.

nais desse campo ainda muito hostil para as mulheres e minorias étnicas e sociais.

42 MARTINS, op. cit.

43 Algumas organizações brasileiras como PretaLab, MariaLab, Odara – Instituto da Mulher Negra, Coding Rights, Minas Programam, Conexão Malunga e ITS Rio atuam na intersecção entre tecnologia, educação, gênero e justiça social, combatendo as desigualdades estruturais no acesso e uso da tecnologia. Enquanto a maioria das iniciativas tech focam na capacitação, essas organizações, em específico, se destacam por: (i) denunciar estruturas de poder por trás dos algoritmos; (ii) criar ferramentas alternativas centradas em justiça racial; (iii) exigir accountability de empresas e governos.

6. Conclusão

O uso crescente de *Big Data* e IAs nas sociedades contemporâneas tem implicações profundas para os direitos humanos e para a qualidade das democracias. Longe de serem neutros, os sistemas baseados em dados carregam os valores, preconceitos e hierarquias dos contextos em que são produzidos. Como demonstrado ao longo deste artigo, os vieses de gênero presentes nesses sistemas impactam de forma direta e estrutural a vida das mulheres, limitando suas oportunidades, distorcendo suas representações e comprometendo sua cidadania digital.

Diante da complexidade crescente das tecnologias digitais e de sua centralidade na vida social e política, torna-se urgente consolidar uma agenda de pesquisa e formulação de políticas públicas que enfrente, de forma estrutural, os vínculos entre IAs, as desinformações estratégicas e a manutenção das desigualdades de gênero e raça. As evidências apresentadas neste artigo demonstram que os sistemas de *Big Data* não apenas refletem vieses sociais preexistentes, mas também os amplificam e automatizam, afetando diretamente os direitos e a cidadania de mulheres e populações historicamente marginalizadas.

A desinformação, especialmente quando potencializada por algoritmos, emerge como uma das formas mais sofisticadas e perigosas da violência de gênero contemporânea, minando a credibilidade, a segurança e a participação política das mulheres. Como destaca o guia produzido pelo InternetLab e pelo Supremo Tribunal Federal, “a desinformação baseada em gênero afeta negativamente o debate público digital” e pode operar como instrumento de exclusão e silenciamento com graves impactos democráticos⁴⁴. Portanto, é imperativo que pesquisadores/as, instituições públicas, movimentos sociais e organismos internacionais atuem de forma coordenada para garantir a regulação ética dessas tecnologias, a inclusão de perspectivas críticas e interseccionais em sua concepção e o fortalecimento de espaços democráticos de produção e

44 CINTRA et al., 2024.

circulação de conhecimento. O futuro da justiça digital depende da capacidade de reconhecer que não há neutralidade tecnológica possível em um mundo profundamente desigual.

É crucial evitar tanto discursos demagógicos sobre tecnologia quanto visões extremamente negativas⁴⁵. Isso ocorre porque o verdadeiro desafio está em aliar o avanço tecnológico ao debate no campo da ética, sendo o caminho mais produtivo, o de fomentar práticas educacionais que integrem as tecnologias digitais de maneira pedagógica, promovendo relações construtivas e responsivas em seu uso. Para que as tecnologias digitais atuem como aliadas da democracia e da justiça de gênero, é urgente a adoção de uma abordagem ética, interseccional e regulatória – que deve incluir não apenas a revisão técnica dos sistemas, mas também a valorização da diversidade nas equipes de desenvolvimento, a regulação pública dos algoritmos, a escuta ativa de grupos historicamente marginalizados e a construção de uma cultura crítica sobre o papel da tecnologia em nossas sociedades.

Construir IAs feministas, que respeitem os direitos humanos e promovam a equidade, não é uma utopia: é uma necessidade diante dos riscos concretos que enfrentamos. E isso só será possível com mais mulheres produzindo ciência, formulando políticas e transformando, com saber e presença, os códigos que moldam o futuro.

45 GARAIGORDOBIL, op. cit.

Patricia Rangel · <https://orcid.org/0000-0002-3093-8011> Especialista em Diversidade e Inclusão com foco em equidade de gênero e tecnologia, é pós-doutora em Sociologia pela FFLCH/USP com estágio de pesquisa na Universidade Livre de Berlim e doutora em Ciência Política pelo Ipol/UnB com missão de estudos na Universidad Nacional de General Sarmiento (Argentina).

Jéssica Melo Rivetti · <https://orcid.org/0000-0003-4642-6878> Doutora com cotutela e dupla titulação pelo Programa de Pós-graduação em Sociologia pela Universidade de São Paulo (USP) e Filosofia I por la Universidad de Granada (Espanha). Pesquisadora da Cátedra Extraordinaria de Filosofía Moral de la Discriminación Corporal da Universidad de Granada e do Ministério da Igualdade da Espanha (InMujeres-UGR) e subcoordenadora de pesquisa do Núcleo de Estudos e Pesquisas sobre a Mulher (NEPEM-UFGM).

Flávia Rios · <https://orcid.org/0000-0001-9864-0644> Professora de Sociologia da Faculdade de Filosofia, Letras e Ciências Humanas (FFLCH) da Universidade de São Paulo (USP). É fundadora e ex-coordenadora do Grupo de Estudos e Pesquisa Guerreiro Ramos (NEGRA). Integrou o Comitê Científico do AFRO/CEBRAP, coordenou o projeto “Gestão municipal da igualdade racial e políticas inclusivas de educação e trabalho no município de Niterói: estudos e ações para sua implementação” e o projeto “Origens e destinos: uma avaliação da política de cotas universitárias e seus efeitos no mercado de trabalho”.

Big Data, desinformação e governos: desafios para garantir a integridade da Informação

Luciana Santana
Wendel Palhares

Resumo

A transformação digital e o avanço das tecnologias de informação e comunicação ampliaram a capacidade de coleta e análise de dados, redefinindo práticas sociais, políticas e governamentais. Nesse contexto, o *Big Data* e a inteligência artificial potencializam tanto a personalização de conteúdos quanto a disseminação de desinformação, ameaçando a confiança institucional e a qualidade democrática. Estudos indicam que, embora a desinformação raramente altere crenças centrais, seus efeitos indiretos fragilizam instituições, aumentam a desconfiança e podem afetar processos eleitorais. Surge, assim, a necessidade de fortalecer a integridade da informação, entendida como a construção de ecossistemas informacionais confiáveis e resilientes. O Núcleo de Integridade da Informação de Alagoas (NII/AL), criado em 2025, exemplifica inovação estatal ao integrar regulação, monitoramento, educação midiática e participação social. A iniciativa articula políticas de cidadania digital, combate a fake news e transparência pública, configurando-se como modelo replicável de governança colaborativa frente aos desafios da sociedade digital.

Abstract

Digital transformation and advances in information and communication technologies have expanded the capacity for data collection and analysis, redefining social, political, and governmental practices. In this context, *Big Data* and artificial intelligence enhance both the personalization of content and the dissemination of misinformation, threatening institutional trust and democratic quality. Studies indicate that, although misinformation rarely alters core beliefs, its indirect effects weaken institutions, increase mistrust, and can affect electoral processes. This gives rise to the need to strengthen information integrity, understood as the construction of reliable and resilient information ecosystems. The Alagoas Information Integrity Center (NII/AL), created in 2025, exemplifies state innovation by integrating regulation, monitoring, media education, and social participation. The initiative articulates digital citizenship policies, combats fake news, and promotes public transparency, establishing itself as a replicable model of collaborative governance in the face of the challenges of the digital society.

1. Introdução

A transformação digital, impulsionada pelo avanço das tecnologias de informação e comunicação (TIC), têm reconfigurado profundamente o ecossistema informacional contemporâneo. O surgimento e a expansão do *Big Data* (KITCHIN, 2014) – caracterizado pelo processamento e análise de grandes volumes de dados provenientes de fontes diversas e em tempo real – possibilitaram novas formas de coleta, armazenamento e interpretação de informações, alterando tanto a maneira como indivíduos consomem conteúdos quanto a forma como instituições públicas e privadas tomam decisões estratégicas.

Ou seja, tem transformado a política contemporânea, a governança estatal e as dinâmicas de comunicação. Plataformas digitais consolidaram-se como espaços centrais de circulação informacional, enquanto

governos e corporações expandiram suas capacidades de coleta, armazenamento e processamento de dados pessoais (BOYD; CRAWFORD, 2012). Esse processo criou um ambiente altamente personalizado, no qual a desinformação – entendida como conteúdos falsos ou enganosos, disseminados com finalidades políticas ou econômicas – pode ser direcionada a grupos segmentados com precisão (WARDLE; DERAKHSHAN, 2017), impondo desafios inéditos às democracias, aos sistemas eleitorais e à proteção de direitos civis (SUNSTEIN, 2017).

Este artigo analisa a interseção entre *Big Data*, desinformação e governos, com ênfase nos impactos do uso massivo de dados e inteligência artificial (IA) de forma responsável para o combate à desinformação. Uma experiência ilustrativa de como governos podem ser protagonistas na adoção de medidas responsáveis é a criação do Núcleo de integridade da informação no Governo do Estado de Alagoas (2025).

Argumentamos que a assimetria de poder informacional entre Estados, políticas públicas e cidadãos exige novas abordagens capazes de articular inovação tecnológica, proteção de dados, direitos civis e integridade democrática. A criação de núcleos governamentais voltados à integridade da informação, como o Núcleo de Integridade da Informação do Estado de Alagoas (NII/AL), representa uma resposta inovadora e potencialmente eficaz para reduzir os impactos da desinformação em ecossistemas digitais altamente personalizados, pois articula uso responsável do *Big Data* e da IA com políticas públicas de proteção de dados, cidadania digital e transparência, fortalecendo a resiliência democrática frente às assimetrias de poder informacional.

2. A desinformação, integridade da informação e impactos na democracia

Castells (2003) argumenta que vivemos em uma “sociedade em rede”, na qual a informação circula de forma contínua e global, reconfigurando relações de poder e influência. Nesse cenário, a desinformação emerge como um desafio central para a manutenção da qualidade de-

mocrática e para o funcionamento das instituições políticas contemporâneas (GARROSSINI, 2024).

E o que seria essa desinformação? Para Grangeiro (2025), é preciso:

...fazer a separação de desinformação da má-informação. Enquanto a primeira refere-se a um processo de produção de conteúdos falsos com o dolo específico de gerar caos ou desmobilizar grupos sociais em determinados contextos, a má-informação está relacionada ao compartilhamento de conteúdos equivocados ou descontextualizados, sem dolo específico de gerar caos, muitas vezes na forma de conteúdos legítimos circulados fora de contexto ou para legitimar um conjunto de crenças pessoais. Essas diferenças são delineadas pela combinação de veracidade, intencionalidade e contexto de uso, categorias relevantes para uma análise jurídica do fenômeno.

Ou seja, essa distinção é crucial para a análise dos mecanismos de circulação de informações no ambiente digital e para a compreensão de seus impactos políticos e sociais.

Segundo Pavão (2024), os estudos sobre desinformação se dividem em dois campos principais: de um lado, aqueles que apontam a ausência de efeitos persuasivos significativos da desinformação sobre crenças e preferências dos cidadãos; de outro, pesquisas que, indo além da dimensão da persuasão, alertam para os efeitos indiretos sobre atitudes e comportamentos políticos relevantes.

De fato, alguns estudos indicam que a desinformação possui um efeito persuasivo limitado. Notícias falsas raramente convencem eleitores ou alteram substancialmente resultados eleitorais, já que sua exposição está fortemente vinculada às identidades políticas e ao raciocínio motivado, o que tende a reforçar crenças já existentes em vez de modificar opiniões centrais.

Contudo, a literatura demonstra que a desinformação pode produzir efeitos políticos indiretos de grande relevância que podem impactar a confiança em instituições e governos, ampliando a exposição frequen-

te a desinformação pode reduzir a confiança em instituições democráticas e diminuir a disposição dos cidadãos para engajamento político (BATISTA PEREIRA et al., 2023); pode influenciar o apoio à democracia e à integridade eleitoral, com a circulação de notícias falsas e teorias conspiratórias tendem a abalar a percepção de imparcialidade e legitimidade dos processos eleitorais (AMARAL et al., 2023; NORRIS et al., 2020) .

Outro efeito indireto muito mencionado pela opinião pública é a polarização política, entretanto, embora as evidências causais sejam ainda limitadas, estudos recentes no Brasil não encontraram efeitos consistentes da desinformação sobre o aprofundamento da polarização (AMARAL et al., 2023; BATISTA PEREIRA et al., 2023). Assim, mesmo quando seu impacto direto sobre as crenças centrais é restrito, a desinformação se revela um fenômeno politicamente relevante por fragilizar a confiança institucional, corroer o apoio à democracia e estimular comportamentos antidemocráticos, exigindo respostas coordenadas de governos, sociedade civil e plataformas digitais.

E o que podemos compreender como integridade da informação? Podemos considerar como um elemento complementar e, ao mesmo tempo, indispensável ao combate à desinformação. Enquanto esforços concentram-se na prevenção, identificação e correção de narrativas falsas ou manipuladas, a integridade da informação visa assegurar que os ecossistemas informacionais sejam confiáveis, auditáveis e resilientes.

Em um contexto de hiperconectividade e circulação instantânea de dados, não basta apenas reagir a conteúdos falsos; é necessário construir estruturas informacionais robustas que promovam a confiança pública. Essa integração envolve múltiplos elementos interdependentes que envolvem uma combinação de infraestruturas confiáveis, políticas públicas claras, educação midiática e cooperação entre governos, sociedade civil e plataformas digitais constituem um marco essencial para enfrentar os desafios da desinformação contemporânea.

A integridade da informação pode ser definida como a capacidade de garantir precisão, consistência, confiabilidade e proteção contra adulteração de dados ao longo de todo o seu ciclo de vida. Nos gover-

nos, está ligada à transparência e confiança pública, à gestão de dados sensíveis e à integridade na comunicação oficial sobre políticas. A manutenção da integridade é essencial para que cidadãos confiem nos dados divulgados por órgãos públicos, evitando manipulações, omissões ou falta de padronização que impactem a credibilidade institucional, prejudicando políticas públicas e a própria democracia.

Isso reforça os argumentos de Garrossini (2024) de que as plataformas digitais e os governos desempenham papéis centrais no combate à desinformação, mas sua atuação envolve desafios de regulação e riscos à liberdade de expressão. O enfrentamento desse fenômeno demanda uma abordagem holística, envolvendo governos, empresas de tecnologia e sociedade civil, para preservar a estabilidade e a legitimidade das democracias.

3. O Núcleo de Integridade da Informação de Alagoas: inovação institucional na proteção do ecossistema informacional

Diante da intensificação de ataques coordenados ao ecossistema informacional e do uso crescente de tecnologias digitais para fins de manipulação, o Governo do Estado de Alagoas criou, em março de 2025, o Núcleo de Integridade da Informação (NII/AL), vinculado à Secretaria de Estado da Comunicação (SECOM/AL). Trata-se de uma iniciativa inédita na América Latina, voltada à formulação de políticas públicas permanentes de proteção do espaço informacional, com ênfase na promoção de cidadania digital, no enfrentamento à desinformação e na construção de uma comunicação pública íntegra, transparente e baseada em evidências.

O NII/AL atua como um centro de inteligência, pesquisa, regulação e mobilização social, com o objetivo de garantir que as informações produzidas e circuladas no Estado – tanto por entes governamentais quanto por atores sociais – sejam confiáveis, auditáveis e alinhadas ao interesse público. Sua estrutura está organizada em cinco eixos estratégicos de atuação:

- Cidadania digital: promoção de programas de educação midiática e alfabetização digital, com foco na proteção de dados pessoais, no combate à desinformação e no fortalecimento da autonomia informacional dos cidadãos;
- Regulação e transparência da comunicação governamental: elaboração e revisão normativa de práticas comunicacionais estatais, assegurando legalidade, eficiência e integridade na veiculação de informações públicas;
- Vigilância e monitoramento do ecossistema informacional: utilização de inteligência artificial e análise de dados para identificação de fluxos de desinformação, discursos de ódio e campanhas coordenadas contra políticas públicas;
- Pesquisa e produção de evidências: incentivo a pesquisas aplicadas, em parceria com instituições acadêmicas, para fundamentar ações estatais de comunicação e fortalecer a base empírica das políticas públicas de integridade da informação;
- Mobilização e participação social: articulação com comunidades, jovens e veículos populares por meio de programas como o *Digaê* e as *Incubadoras de Comunicação Comunitária*, ampliando a presença de vozes locais e fortalecendo a democracia comunicacional.

Além desses eixos, o NII/AL é composto por cinco subnúcleos temáticos, voltados a áreas sensíveis de circulação de desinformação: Saúde e Consciência; Clima e Meio Ambiente; Democracia e Eleições; Justiça e Segurança Pública; e Cidadania e Direitos Humanos. Em cada subnúcleo, articulam-se ações específicas de monitoramento, produção de conteúdo e resposta institucional, alinhadas às peculiaridades de cada campo temático.

O Núcleo também abriga programas emblemáticos, como o *Saúde com Ciência* – voltado à identificação e mitigação de fake news em saúde, especialmente em campanhas de vacinação – e o *Observatório da Desinformação Eleitoral*, em cooperação com o TRE, MPE, MPF e universidades, que monitora a circulação de conteúdos enganosos durante o período eleitoral. A iniciativa de bolsas de pesquisa para mestrado e

doutorado fortalece o vínculo entre academia e governo, subsidiando políticas públicas baseadas em dados e orientadas por impacto social.

A atuação do NII/AL é interinstitucional e multiescalar, envolvendo cooperação com a Secretaria de Comunicação da Presidência da República, ministérios, órgãos de controle, sistema de justiça, universidades, organismos internacionais e representantes da sociedade civil. O núcleo representa, portanto, uma inovação política e administrativa com alto potencial de replicabilidade, combinando regulação, participação social, produção de conhecimento e uso estratégico de tecnologias para fortalecer a integridade da informação como bem público.

Ao propor uma abordagem integrada – que vai da normatização de contratos publicitários à formação de jovens comunicadores – o modelo alagoano reposiciona o papel do Estado na era da informação: não apenas como emissor de dados oficiais, mas como curador ético e responsável do ecossistema informacional, apto a promover confiança pública, combater a manipulação informacional e garantir a resiliência democrática frente aos desafios da sociedade digital.

4. A mídia intermediária como vetor informacional e desafio ético

Um dos desafios frequentemente negligenciados no campo da governança informacional refere-se ao relacionamento dos governos com veículos noticiosos e de comunicação digital, especialmente portais locais, blogs e mídias de pequeno porte. Esses agentes não apenas produzem conteúdo, mas atuam como vetores de distribuição de mensagens estatais e privadas em ambientes como WhatsApp, Telegram, Instagram, TikTok, Facebook e X, influenciando diretamente a circulação de informações e desinformações nos ecossistemas digitais.

No contexto de *Big Data*, esses veículos funcionam como nós intermediários na arquitetura da desinformação ou da integridade, dependendo das práticas editoriais, do financiamento e da regulação pública. O caso de Alagoas exemplifica essa preocupação: o Governo do Estado,

reconhecendo o papel estratégico dos veículos locais na produção e difusão de informação, enfrentou o desafio de criar critérios públicos para o financiamento desses agentes com recursos estatais, de forma a preservar o valor público da comunicação.

Para isso, foi realizada uma consulta pública, seguida da publicação da Instrução Normativa SECOM nº 001/2025, que estabeleceu critérios técnicos, jurídicos e éticos para a contratação de mídia digital institucional. O normativo prevê, por exemplo: *“O veículo não deve promover conteúdo ofensivo, discriminatório ou notícias falsas, mantendo o compromisso público com a verificação dos fatos, a responsabilidade editorial e as disposições da Lei Geral de Proteção de Dados (LGPD).”*

A Instrução Normativa também prevê auditorias regulares e a exigência de métricas auditáveis (como Google Analytics 4), além de determinar que: *“As denúncias relativas à desinformação deverão ser encaminhadas ao e-mail consecomal25@gmail.com e, posteriormente, analisadas pelo Conselho Estadual de Comunicação – Consecom, que emitirá parecer recomendando, ou não, o descredenciamento do veículo.”*

Com isso, observa-se um esforço governamental não apenas de ordenar a aplicação de recursos públicos, mas de institucionalizar a integridade da informação como critério de elegibilidade para o financiamento público de mídia, promovendo o alinhamento entre valores democráticos, pluralidade midiática e responsabilidade editorial.

5. Considerações finais

Big Data e o uso massivo de inteligência artificial têm provocado transformações profundas em diversas esferas da sociedade, incluindo o funcionamento dos governos nas democracias contemporâneas. A capacidade de coletar, armazenar e processar grandes volumes de dados pessoais e comportamentais expande significativamente o alcance de Estados e corporações, criando estruturas de monitoramento mais sofisticadas e possibilitando o microdirecionamento de informações com precisão sem precedentes.

Ao mesmo tempo, essas tecnologias oferecem ferramentas avançadas para detecção e mitigação da desinformação, embora também apresentem desafios éticos e desigualdades de acesso. A promoção da equidade, a proteção de direitos civis e a preservação da integridade democrática dependem da formulação de políticas públicas e regulações tecnológicas capazes de equilibrar inovação, segurança e liberdade, fomentando um ecossistema informacional transparente e responsável.

Nesse contexto, a desinformação se consolida como um fenômeno estratégico, alimentado por algoritmos e plataformas digitais que segmentam públicos e reforçam bolhas informacionais, tornando o ecossistema comunicacional mais complexo e vulnerável. A atuação governamental diante dessas transformações não pode ser encarada de forma isolada. Casos emblemáticos que movimentaram a opinião pública como o escândalo Cambridge Analytica e o vazamento de dados em Xangai, evidenciam que a propagação de conteúdos falsos ou manipulativos gera impactos transnacionais, exigindo mecanismos coordenados de regulação, transparência e responsabilização de atores públicos e privados.

Iniciativas como a do NII/AL sinalizam para uma atuação estatal ativa, porém estruturada em modelo cooperativo com setores relevantes da sociedade civil e instituições públicas diversas. Essa lógica de governança colaborativa contribui para evitar a concentração de poder regulatório, ao mesmo tempo em que promove a construção coletiva de conhecimento, práticas de verificação e padrões mínimos de integridade editorial e informacional – inclusive no uso de *Big Data*.

Ao fomentar um ecossistema informacional mais saudável, o Estado não abdica de seu papel institucional, mas o exerce por meio de escuta ativa, articulação interinstitucional e valorização da participação social. Tal abordagem tem o potencial de se consolidar como ferramenta eficaz contra a desinformação, ao fortalecer a cidadania digital, ampliar a qualidade do debate público e aprimorar a prestação de serviços de comunicação pública e privada. Isso é especialmente relevante em contextos marcados pela “enxurrada informacional”, que tende a obscurecer a compreensão crítica e enfraquecer a confiança cidadã nas instituições democráticas.

Podemos considerar que, apesar de casos concretos de respostas institucionais governamentais, a resposta institucional à desinformação ainda é fragmentada e a ausência de consenso internacional sobre regulação de plataformas digitais é algo preocupante.

Grangeiro (2025) reforça que a regulação das plataformas é necessária, mas insuficiente, sendo preciso articular políticas de educação midiática, comunicação pública baseada em IA, checagem de fatos e abordagens empáticas com os usuários, compondo uma estratégia multidisciplinar de mitigação da desinformação. Não há tanta clareza sobre priorização de conteúdos criando assimetrias de poder nos diferentes governos, sejam eles locais ou nacionais. Torna-se importante que os princípios de justiça, não discriminação e respeito aos direitos humanos na utilização da IA sejam devidamente seguidos (UNESCO, 2021), assim como com o uso massivo de dados.

Nesta direção, a luta contra a desinformação tem se mostrado multifacetada: governos adotam políticas que vão da regulação rigorosa a estratégias de desregulação seletiva, enquanto organismos internacionais e a sociedade civil fortalecem padrões éticos, promovem alfabetização midiática e fiscalizam práticas de coleta e uso de dados. Uma diretriz central consiste em garantir a integridade da informação e implementar medidas capazes de reduzir impactos sobre as políticas públicas. Enfrentar os riscos da desinformação na era do *Big Data* demanda, portanto, uma abordagem integrada que articule governança digital, cooperação internacional, responsabilidade corporativa e participação cidadã ativa, reconhecendo que a produção, disseminação e controle da informação são centrais para a resiliência e sustentabilidade das democracias contemporâneas.

Referências

ALAGOAS. Secretaria de Estado da Comunicação. Instrução Normativa SECOM nº 001/2025, de 23 de maio de 2025. Estabelece critérios técnicos para a contratação de espaços publicitários na internet e medidas de integridade informacional. Diário Oficial

do Estado de Alagoas, Maceió, 26 maio de 2025. Disponível em: <https://diario.impressaooficial.al.gov.br/apinova/api/editions/viewPdf/50761>. Acesso em: 08 set. 2025.

AMARAL, M.; BORBA, L.; LESSA, E.; LIRA, V.; PAVÃO, N. **Do conspiracy theories undermine support for democracy? Experimental evidence from Brazil**. 2023.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD)**. Diário Oficial da União, Brasília, DF, 15 ago. 2018.

BOYD, D.; CRAWFORD, K. Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. **Information, Communication & Society**, v. 15, n. 5, p. 662-679, 2012.

GARROSSINI, Daniela Fávaro. Desinformação e o Desafio das Instituições Democráticas. In: Daroit, Dorian; Cruz, Fernanda Natasha Bravo (Orgs.). **Diretrizes Nacionais para Ações de Cidadania das EJE**s: Fundamentos, Caminhos e Desafios. 1. ed. Brasília: Tribunal Superior Eleitoral, 2024. p. 65-68. Disponível em: https://www.tse.jus.br/institucional/catalogo-de-publicacoes/arquivos/diretrizes-nacionais-para-acoes-de-cidadania-das-ejes-fundamentos-caminhos-e-desafios/@@display-file/file/Acao_formativa_Web_1.pdf. Acesso em 08 de setembro de 2025.

GRANGEIRO, J. J. M. Por um conceito jurídico de desinformação que enxergue além do Direito. **JOTA**, 16 jan. 2025. Disponível em: <https://www.jota.info/opiniao-e-analise/artigos/por-um-conceito-juridico-de-desinformacao-que-enxergue-alem-do-direito>. Acesso em: 4 set. 2025.

KITCHIN, R. **The data revolution**: big data, open data, data infrastructures and their consequences. London: Sage, 2014.

PEREIRA, F. B.; BUENO, N. S.; NUNES, F.; PAVÃO, N. and WIRTSCHAFTER, V. The (de) mobilizing effects of misinformation: Evidence from Brazil. 2023.

NORRIS, Pippa. **Why Electoral Integrity Matters**. Cambridge: Cambridge University Press, 2014.

NORRIS, P. GARNETT, H. A. and GROMPING, M. The paranoid style of American elections: explaining perceptions of electoral integrity in an age of populism. *Journal of elections, public opinion and parties*, 30(1), p.105–125, 2020.

PAVÃO, N. Os efeitos da desinformação. In: BRASIL. Tribunal Superior Eleitoral. **Diretrizes nacionais para ações de cidadania das EJE**s: fundamentos, caminhos e desafios. Brasília: Tribunal Superior Eleitoral, 2024. p. 72-77. Disponível em: https://www.tse.jus.br/institucional/catalogo-de-publicacoes/arquivos/diretrizes-nacionais-para-acoes-de-cidadania-das-ejes-fundamentos-caminhos-e-desafios/@@display-file/file/Acao_formativa_Web_1.pdf. Acesso em 08 de setembro de 2025.

SECRETARIA DE COMUNICAÇÃO SOCIAL DA PRESIDÊNCIA DA REPÚBLICA; SECRETARIA DE ESTADO DA COMUNICAÇÃO DE ALAGOAS. Secom/PR e Secom de Alagoas firmam parceria para promoção da integridade da informação, com enfrentamento às fake news sobre vacinas. **GazetaWeb**, Maceió, 04 ago. 2025. Disponível em: <https://www.gazetaweb.com/noticias/politica/secompr-e-secom-de-alagoas-firmam-parceria-para-promocao-da-integridade-da-informacao-com-enfrentamento-as-fake-news-sobre-vacinas-829414>. Acesso em: 08 de setembro de 2025.

SUNSTEIN, C. R. **#Republic**: divided democracy in the age of social media. Princeton: Princeton University Press, 2017.

WARDLE, C.; DERAKHSHAN, H. **Information disorder**: toward an interdisciplinary framework for research and policy making. Strasbourg: Council of Europe, 2017.

UNESCO. **Recommendation on the Ethics of Artificial Intelligence**. Paris: UNESCO, 2021.

CASTELLS, Manuel. **A sociedade em rede**. 11. ed. São Paulo: Paz e Terra, 2003.

Luciana Santana · Doutora em Ciência Política (UFMG) e Pesquisadora no PPGCP/UFAL e PPGCP/UFPI. E-mail: lucianasantana@ics.ufal.br

Wendel Palhares · Mestrando em Comunicação Digital pelo IDP e pós-graduando em Comunicação Pública pela PUC Minas. E-mail: wendelpalhares@gmail.com

Big Data: os desafios à regulamentação do uso de dados

Eleonora Mesquita Ceia
Pedro Teixeira Gueiros

Resumo

O artigo examina os desafios da regulamentação do uso de dados em um cenário global marcado pelo avanço exponencial das tecnologias digitais e pela centralidade informacional do *Big Data*. Parte-se da constatação de que os dados se tornaram o principal ativo econômico da era digital, impulsionando modelos de negócio baseados em vigilância algorítmica e na extração massiva de informações pessoais. O estudo estrutura-se em três partes: inicialmente, parte-se de uma análise dos riscos e potencialidades associados ao *Big Data*, destacando a natureza não neutra das decisões algorítmicas. Em seguida, passa-se a uma abordagem comparativa do panorama regulatório, com ênfase no *AI Act* da União Europeia, no Projeto de Lei nº 2.338/2023 no Brasil e na Resolução do Conselho Nacional de Justiça nº 615/2025. A seguir, faz-se uma proposição de caminhos normativos para o fortalecimento da governança democrática dos dados, com base em quatro eixos: regulação coordenada e multissetorial, transparência algorítmica, promoção da justiça de dados e educação digital. Por fim, sustenta-se que regular o *Big Data* apresenta-se, antes de tudo, um enfrentamento de uma questão política e moral.

Abstract

This article examines the challenges of regulating data use in a global scenario characterized by the exponential advancement of digital technologies and the informational centrality of Big Data. It begins with the observation that data has become the primary economic asset of the digital age, driving business models based on algorithmic surveillance and the extensive extraction of personal information. The study is structured in three parts: first, an analysis of the risks and potential associated with Big Data, highlighting the non-neutral nature of algorithmic decisions. This is followed by a comparative approach to the regulatory landscape, with emphasis on the European Union's AI Act, Brazilian Bill of Law No. 2.338/2023, and the Resolution No. 615/2025 from the Brazilian National Council of Justice. Next, a proposal is made for regulatory paths to strengthen democratic data governance, based on four axes: coordinated and multisectoral regulation, algorithmic transparency, promotion of data justice, and digital education. Finally, it is argued that regulating Big Data is first and foremost, a political and moral issue.

1. Introdução

Em uma breve pesquisa sobre as marcas mais valiosas do mundo, é possível observar que empresas de tecnologia se mantêm invariavelmente hegemônicas em termos de liderança na atuação do mercado. *Apple, Microsoft, Google, Amazon, Samsung, TikTok e Facebook*, são alguns exemplos que encabeçam as dez primeiras posições do ranking atualizado em 2025.¹ Ainda assim, é curioso constatar como uma ampla série de serviços disponibilizados por essas empresas são efetivamente

1 PIO, Juliana. *As marcas mais valiosas do mundo em 2025 (duas delas são brasileiras)*. Disponível em: <https://exame.com/marketing/saiba-quais-sao-as-marcas-mais-valiosas-do-mundo-em-2025-duas-delas-sao-brasileiras/>. Acesso em: 10.05.2025.

gratuitos, como redes sociais, serviços de mensageria ou apenas provedores de buscas. Nesse cenário, passa a ser questionável como tais companhias podem ser até mesmo mais valiosas que o PIB de países inteiros,² quando não há produtos efetivamente custosos, a rigor, sendo comercializados.³

Nesse sentido, a escalada vultosa de grandes corporações tecnológicas, também conhecidas como *Big Techs*, está indiscutivelmente imbricada no uso massivo de um ativo que parece ter se tornado uma verdadeira *commodity*: os dados pessoais. Sob esse fenômeno diversas expressões foram cunhadas por estudiosos com o intuito de compreender a nova ordem econômica mundial, que se edificou de forma tão recente na história humana contemporânea: *capitalismo de vigilância*,⁴

2 A título de exemplo, o valor de mercado da *Apple* alcançou o valor de US\$ 3 trilhões em 2023, situação em que apenas seis países no mundo à época tinham o Produto Interno Produto superior a este valor: Estados Unidos, China, Índia, Japão, Alemanha e Reino Unido. GUILHERME, Guilherme. *Apenas 6 países têm PIB maior que valor de mercado da Apple*. Disponível em: <https://exame.com/in-vest/mercados/apenas-6-paises-tem-pib-maior-que-valor-de-mercado-da-apple/amp/>. Acesso em: 11.05.2025.

3 Nesse sentido, Kris Shaffer observa: “Algorithmic recommendation engines and social media feeds have been created to help users find the most relevant content and to help media producers find the most appropriate audiences. But the ways in which media producers compete for our attention, the amount of personal data mined to make the algorithms work, and the natural way our cognitive systems function all combine to make the modern media landscape ripe for propaganda”. SHAFFER, Kris. *Data versus Democracy: How Big Data Algorithms Shape Opinions and Alter the Course of History*. Nova York: Apress, 2019, p. 17-18.

4 Acerca do capitalismo de vigilância, Shoshana Zuboff, pesquisadora de psicologia social, descreve: “Os capitalistas de vigilância fazem um grande esforço para camuflar seu propósito enquanto dominam o uso do poder instrumentário para moldar nosso comportamento e, ao mesmo tempo, escapar da nossa consciência. É por isso que o Google esconde as operações que nos transformam em objetos de busca e o Facebook nos distrai do fato de que nossas tão adoradas conexões são essenciais para o lucro e poder que fluem da ubiquidade da rede e de seu conhecimento totalista”. ZUBOFF, Shoshana. *A era do capitalismo de vigilância: a luta por um futuro humano na nova fronteira do poder*. Rio de Janeiro: Intrínseca, 2020, p. 500.

dataísmo,⁵ *armas de destruição de massa*,⁶ *infocracia*,⁷ dentre outras nomenclaturas. Em comum, observa-se que as mais variadas atividades humanas dependem progressivamente do processamento de máquinas inteligentes que se desenvolvem a partir de um número quase incalculável de dados e informações. Sob essa lógica, a confiança tradicionalmente atribuída a algoritmos humanos, naturalmente limitados, dá lugar à confiança a algoritmos efetivamente artificiais, cuja potencialidade é exponencial e de certa forma, irrestrita, como verificado pela *Lei de Moore*.⁸

-
- 5 O historiador Yuval Noah Harari observa que: “O dataísmo adota uma abordagem estritamente funcional em relação à humanidade, avaliando o valor das experiências humanas de acordo com sua função nos mecanismos de processamento de dados. Se desenvolvermos um algoritmo que realize melhor a mesma função, as experiências humanas perderão seu valor”. HARARI, Yuval Noah. *Homo Deus: uma breve história do amanhã*. São Paulo: Companhia das Letras, 2016, p. 391.
- 6 Como pondera a matemática Cathy O’Neil acerca das denominadas *Armas de Destruição em Massa* (ADMs): “Dados não vão desaparecer. Nem computadores – e muito menos a matemática. Modelos de previsão são, cada vez mais, as ferramentas com as quais contaremos para administrar nossas instituições, aplicar nossos recursos e gerenciar nossas vidas. Mas, como tentei mostrar ao longo deste livro, esses modelos são construídos não apenas de dados, mas das escolhas que fazemos sobre em quais dados prestar atenção – e quais deixar de fora. Essas escolhas não tratam apenas de logísticas, lucros e eficiência. Elas são fundamentalmente morais”. O’NEIL, Cathy. *Armas de Destruição em Massa: como o big data aumenta a desigualdade e ameaça a democracia*. São Paulo: Rua do Sabão, 2020, p. 337.
- 7 Nesse sentido, o filósofo Byung-Chul Han, “A digitalização do mundo da vida avança, implacável. Submete a uma mudança radical nossa percepção, nossa relação com o mundo, nossa convivência. Ficamos atordoados pela embriaguez de comunicação e informação. O tsunami de informação desencadeia forças destrutivas. Abrange também, nesse meio-tempo, âmbitos políticos e leva a fraturas e rupturas massivas no processo democrático. A democracia degenera em *infocracia*”. HAN, Byung-Chul. *Infocracia: digitalização e a crise da democracia*. Petrópolis: Vozes, 2022, p. 25.
- 8 Conjecturada em 1965, por Gordon Moore, cofundador da Intel, a “Lei de Moore” ficou conhecida como sendo a observação empírica que prevê que o número de transistores de um *chip* dobra aproximadamente a cada 18 meses, mantendo-se o mesmo custo. A rigor, esta tendência impulsionou o avanço exponencial da capa-

As atuais e populares Inteligências Artificiais Generativas, ilustrativamente, capazes de gerar conteúdos inéditos altamente criativos, como textos, imagens e músicas, dependem de uma quantidade colossal de dados para que os grandes modelos de linguagem sejam instrumentalizados, sob os chamados *data lakes*.⁹ Em 2023, por exemplo, o célebre *ChatGPT* da OpenAI, em seu modelo GPT-3, utilizava entre 570 *gigabytes* e 45 *terabytes* de informações coletadas na internet.¹⁰ Nesse cenário, o *Big Data*, enquanto fenômeno transversal às diversas ciências do conhecimento, apresenta uma miríade de desafios e potencialidades. Sob a perspectiva jurídica, uma das principais questões a ser endereçada diz respeito à regulação das atividades que se desenvolvem à luz do grande processamento de dados.

O presente artigo estrutura-se, nesse sentido, na compreensão inicial dos riscos e das potencialidades associadas ao *Big Data*. Em seguida, passa-se ao exame do atual panorama regulatório, por meio de uma análise comparativa entre diferentes jurisdições. Por fim, propõe-se alguns apontamentos para conferir maior legitimidade às atividades desenvolvidas no contexto do *Big Data*.

cidade computacional em paralelo à diminuição de dispositivos eletrônicos. JORDÃO, Fabio. *Afinal, o que é a Lei de Moore?* Disponível em: <https://www.tecmundo.com.br/curiosidade/701-o-que-e-a-lei-de-moore-.htm>. Acesso em: 12.05.2025.

- 9 Imprescindível ao funcionamento de Inteligências Artificiais, como um todo, “Um *data lake* é um repositório centralizado que ingere e armazena grandes volumes de dados em sua forma original. Os dados podem ser processados e usados como base para uma variedade de necessidades analíticas.” MICROSOFT. *O que é um data lake*. Disponível em: <https://azure.microsoft.com/pt-br/resources/cloud-computing-dictionary/what-is-a-data-lake>. Acesso em: 12.05.2025
- 10 OPENAI. *What is the size of the training set for GPT-3*. Disponível em: <https://community.openai.com/t/what-is-the-size-of-the-training-set-for-gpt-3/360896>. Acesso em: 11.05.2025.

2. *Big Data*: entre riscos e potencialidades

O *Big Data*, enquanto uma expressão em certa medida amorfa voltada à compreensão de atividades que envolvem um grande volume de dados complexos e diversificados, impõe nuances concretas para que seus riscos e potencialidades sejam examinados. Justamente em razão disso, para colocar o pingô nos is, é necessário responder a seguinte pergunta: de que tipo de “dados” está se falando?

No Brasil, dados e informações de interesse público, por exemplo, estão sujeitas à regulação da Lei de Acesso à Informação (Lei nº 12.527/2011),¹¹ o que impõe ao Poder Público o dever de publicizar, como regra, as atividades desenvolvidas. Informações podem ainda ser de caráter pessoal, o que a rigor, obriga ao cumprimento das diretrizes estabelecidas pela Lei Geral de Proteção de Dados (Lei nº 13.709/2018),¹² sem prejuízo de outras previsões específicas, como relações de consumo e as determinações do Código de Defesa do Consumidor (Lei nº 8.078/1990)¹³ e da Lei do Cadastro Positivo (Lei

11 Como determina as diretrizes da LAI: “Art. 3º Os procedimentos previstos nesta Lei destinam-se a assegurar o direito fundamental de acesso à informação e devem ser executados em conformidade com os princípios básicos da administração pública e com as seguintes diretrizes: I – observância da publicidade como preceito geral e do sigilo como exceção; II – divulgação de informações de interesse público, independentemente de solicitações; III – utilização de meios de comunicação viabilizados pela tecnologia da informação; IV – fomento ao desenvolvimento da cultura de transparência na administração pública; V – desenvolvimento do controle social da administração pública”.

12 Como se extrai dos objetivos da LGPD: “Art. 1º Esta Lei dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural”.

13 Conforme expressa determinação do CDC: “Art. 43. O consumidor, sem prejuízo do disposto no art. 86, terá acesso às informações existentes em cadastros, fichas, registros e dados pessoais e de consumo arquivados sobre ele, bem como sobre as suas respectivas fontes”.

nº 12.414/2011),¹⁴ ou ainda no bojo de relações virtuais, às regras do Marco Civil da Internet (Lei nº 12.965/1014).¹⁵ Outras informações podem envolver também tutelas mais específicas, e nem sempre tão óbvias, a exemplo de obras criativas, sinais distintivos, patentes, segredos comerciais, que estão sujeitos à regulação em matéria de Propriedade Intelectual, notadamente por meio da Lei de Direitos Autorais (Lei nº 9.610/1998)¹⁶ e da Lei de Propriedade Industrial (Lei nº 9.279/1996).¹⁷ No que se refere ao *Big Data*, a questão se coloca em termos da escala com que tantas informações, de natureza jurídica distintas, são processadas, o que impõe verdadeira governança holística em atenção aos diferentes microssistemas.

A economia do *Big Data* trouxe consigo a promessa de notáveis ganhos e oportunidades, mediante a análise de imensos volumes de dados diversos, de forma veloz, imparcial e objetiva. Porém, tal análise parte de modelos matemáticos e estatísticos desenvolvidos por seres

14 Como se depreende dentre os direitos facultados à pessoa cadastrada em bancos de adimplemento na LCP: “Art. 5º São direitos do cadastrado: II – acessar gratuitamente, independentemente de justificativa, as informações sobre ele existentes no banco de dados, inclusive seu histórico e sua nota ou pontuação de crédito, cabendo ao gestor manter sistemas seguros, por telefone ou por meio eletrônico, de consulta às informações pelo cadastrado”.

15 Como se observa dos deveres incumbidos aos provedores de conexão de internet e provedores de aplicações de internet no MCI: “Art. 10. A guarda e a disponibilização dos registros de conexão e de acesso a aplicações de internet de que trata esta Lei, bem como de dados pessoais e do conteúdo de comunicações privadas, devem atender à preservação da intimidade, da vida privada, da honra e da imagem das partes direta ou indiretamente envolvidas”.

16 Nos termos do art. 7º, da LDA, ao apresentar as obras protegidas por direitos autorais, para além textos, fotografias, pinturas, esculturas, filmes, músicas, o inciso XIII, menciona expressamente que também são protegidas “as coletâneas ou compilações, antologias, enciclopédias, dicionários, bases de dados e outras obras, que, por sua seleção, organização ou disposição de seu conteúdo, constituam uma criação intelectual”.

17 Paralelamente, entende-se que são protegidos mediante registro invenções relacionadas a patentes, modelos de utilidade, marcas, indicações geográficas, além de jogos eletrônicos, conforme art. 2º da LPI.

humanos imbuídos de preconceitos e vieses. O problema é agravado quando se constata que esses modelos são opacos, isto é, não compreensíveis às pessoas em geral, salvo aquelas que dominam a complexidade algorítmica neles envolvida. Além disso, os modelos são protegidos por regras legais de segredos comercial e industrial¹⁸. Isso dificulta ou até mesmo impede questionamentos contra decisões tomadas no âmbito do *Big Data*. Em síntese, o *Big Data* abarca operações não transparentes, dificilmente contestáveis e, em muitos casos, ainda não passíveis de regulamentação¹⁹.

A maior preocupação, no que diz respeito aos riscos do *Big Data*, está associada, não apenas ao risco da mera inobservância do uso de dados em sentido *lato*, mas também, às consequências negativas diante do risco de violação a direitos fundamentais. Assim sendo, importa reconhecer que o *Big Data*, enquanto expressão inerente ao atual cenário de novas tecnologias, não se traduz em um fenômeno neutro *per se*. Trata-se de uma escolha tecnológica ligada a relações e estruturas de poder, que beneficiam uns e excluem outros. Vale dizer, muito embora se apresente conjecturado à luz de critérios estáticos e analíticos baseados nas ciências exatas, sua aplicação não deixa de ser efetivamente social e co-

18 Frank Pasquale trata dessa questão alertando para a presença massiva de ferramentas tecnológicas em domínios vitais da nossa vida civil, como saúde, educação e trabalho, que funcionam como verdadeiras “caixas pretas”. Isso porque não são transparentes com relação a suas funções, finalidades e uso de dados, devido à alta complexidade tecnológica que apenas suas proprietárias privadas dominam, cujo sigilo é protegido pelo Estado por meio de legislação específica. PASQUALE, Frank. *The Black Box Society: The Secret Algorithms that Control Money and Information*. Cambridge/London: Harvard University Press, 2015.

19 “As aplicações matemáticas fomentando a economia dos dados eram baseadas em escolhas feitas por seres humanos falíveis. Algumas dessas escolhas sem dúvida eram feitas com as melhores intenções. Mesmo assim, muitos desses modelos programavam preconceitos, equívocos e vieses humanos nos sistemas de software que cada vez mais geriam nossas vidas. Como deuses, esses modelos matemáticos eram opacos, seus mecanismos invisíveis a todos exceto aos altos sacerdotes de seus domínios: os matemáticos e cientistas da computação. Suas decisões, mesmo quando erradas ou danosas, estavam para além de qualquer contestação”. O’NEIL, *op. cit.*, p. 12.

letiva, o que leva a que quaisquer tomadas de decisões impliquem em consequências de ordem política ou moral.

Por exemplo, no caso de uma máquina inteligente ser empregada para designar as chances de uma pessoa ser melhor qualificada para uma vaga de trabalho, ou ainda, de obter a concessão de um empréstimo bancário ou de reincidir em algum crime anteriormente praticado, implica invariavelmente em uma escolha e decisão humana dos desenvolvedores do sistema algorítmico em autorizar o seu uso dentro daquele determinado contexto.

Com efeito, a publicidade com uso intensivo de dados se tornou muito lucrativa. As empresas do ramo formam perfis a partir dos dados coletados de atividades empreendidas pelos indivíduos nas redes sociais. Trata-se de vantagem informacional que possibilita a identificação de consumidores ou eleitores ideais. Os dados e perfis são vendidos ou compartilhados ao Estado²⁰, empresas e empregadores. Essa atividade de perfilamento acaba construindo uma reputação digital dos indivíduos, por meio de informações do que elas consomem, suas preferências, interesses, hábitos, saúde, medicamentos que usam etc. O tratamento dos dados é realizado sem qualquer transparência, devido ao problema da opacidade, anteriormente explicado. Isso levanta preocupações em termos de tratamento justo de dados, pois o perfilamento é realizado sob vieses humanos, com a possibilidade de reproduzir padrões de desigualdade e discriminação, além da violação de direitos da personalidade²¹.

20 “The mountains of data collected by private corporations make them valuable partners in ‘information sharing’. There’s plenty of room for dealing on both sides. Government agencies want data that they can’t legally or constitutionally collect for themselves; data brokers have it and want to sell it”. PASQUALE, *op. cit.*, p. 48-49.

21 “Reputation systems are creating new (and largely invisible) minorities, disfavored due to error or unfairness. Algorithms are not immune from the fundamental problem of discrimination, in which negative and baseless assumptions congeal into prejudice”. PASQUALE, *op. cit.*, p. 38.

Com relação à democracia, as ferramentas tecnológicas, como a IA, os algoritmos e o *Big Data*, operam em termos contraditórios. As pessoas utilizam as redes sociais para exercerem liberdades fundamentais, como a informação, a comunicação e a manifestação do pensamento. Assim, nossa comunicação e o debate público ocorrem cada vez mais em espaços digitais privados, que são geridos por uma tecnologia complexa e não transparente. Isso resulta no monopólio e na concentração da informação pelas grandes empresas de Internet. As redes sociais, por meio de seus algoritmos, passam a deter o poder de influenciar o processo político e a opinião pública, por meio do direcionamento de notícias e posts, incluindo desinformação, ou seja, para fins não democráticos²².

Igualmente, o *Big Data* possui uma relação ambígua com o meio ambiente. De um lado, em fóruns internacionais, parte-se da premissa de que a análise de um grande volume de informações, coletadas automaticamente, formam subsídios fundamentais para a formulação de políticas de desenvolvimento sustentável, o monitoramento das mudanças climáticas e o combate ao desmatamento. De outro, neste debate foi negligenciada a preocupante “pegada ambiental” do *Big Data*²³. Tecnologias da escala do *Big Data* demandam um alto consumo de energia nas instalações de *data center*. Além disso, nessas instalações “precisam operar geradores a diesel em caso de falta de energia para manter os servidores funcionando, o que acarreta a emissão de

22 SANTOS, Andréia. O Impacto do Big Data e dos Algoritmos nas Campanhas Eleitorais. In: BRANCO, Sérgio; DE TEFFÉ, Chiara (orgs.). *Privacidade em perspectivas*. Rio de Janeiro: Lumen Juris, 2018, p. 1-24.

23 “High data volume requires extensive facilities for storage using natural resources such as water and non-renewable energy that require maintenance and accessibility. Furthermore, the manufacturing and disposal of devices to collect and process data emits polluting substances. The data revolution, advocated as a vehicle to achieve sustainable development, is supported by technologies that endanger sustainability and the environment by other means”. LUCIVERO, Federica. *Big Data, Big Waste? A Reflection on the Environmental Sustainability of Big Data Initiatives*. Science and Engineering Ethics, v. 26, 2020, p. 1013.

gases de efeito estufa com impactos significativos sobre a mudança climática”²⁴.

Outro problema ambiental ligado ao *Big Data* são as emissões nocivas provenientes do descarte do hardware de computação. Os centros de descarte e reciclagem estão localizados, em geral, em países menos desenvolvidos, o que suscita o debate sobre a justa distribuição dos benefícios e danos da revolução digital. de modo a evitar que “as populações mais vulneráveis sofram os efeitos ambientais da crescente demanda digital dos países mais desenvolvidos”^{25, 26}.

Diante da presença do *Big Data* em variados domínios da vida dos indivíduos e das relações entre os países, junto com os significativos impactos sobre a democracia e a justiça social e ambiental, justifica-se a emergência de iniciativas de regulação, com o propósito de garantir as devidas responsabilidades no caso de abusos e violações a direitos fundamentais.

3. Panorama regulatório nacional e internacional

Diante da ubiquidade do *Big Data* em sua progressiva confluência das atividades desenvolvidas por agentes públicos e privados, diversas jurisdições têm buscado estabelecer marcos normativos voltados a mitigar os riscos e promover o uso ético e responsável dos dados.

24 LUCIVERO, *op. cit.*, p. 1015 [tradução livre dos autores].

25 LUCIVERO, *op. cit.*, p. 1022 [tradução livre dos autores].

26 Interessante notar que usualmente o discurso adotado referente às tecnologias digitais utiliza termos (imateriais, abundantes e nuvens, por exemplo), que ocultam e minimizam o real impacto da infraestrutura física do *Big Data*. Nesse sentido, Lucivero adverte: “The digital has material implications for the physical “real” world, but the language used obscures this. Places, infrastructures, and buildings all play a role in constituting the online world, and vice versa: online behaviour also has material implications in the physical “real” world. False assumptions, created by misleading language, influence people’s understanding of the data universe and economy, and shapes their attitudes towards it”. LUCIVERO, *op. cit.*, p. 1019.

Sob a realidade hiperconectada, marcada pelo uso massivo de dados instrumentalizados por ferramentas de IA, a regulação contemporânea enfrenta o desafio de acompanhar a velocidade exponencial do desenvolvimento tecnológico, em atenção à preservação de direitos fundamentais e à governança democrática da informação.

No plano internacional, destaca-se a recente promulgação do *AI Act* pela União Europeia em meados de 2024. Trata-se do primeiro regulamento geral voltado à disciplina da IA de forma integral, estabelecendo diretrizes para seu uso em uma abordagem baseada em riscos. Sob esta perspectiva, os sistemas de IA são classificados conforme o grau de potenciais impactos a direitos e liberdades dos cidadãos, visando promover a segurança jurídica e equidade no acesso e disponibilização de soluções movidas à IA. A depender dessa categorização, o regulamento impõe obrigações mais ou menos rigorosas quanto à transparência, mitigação de vieses, supervisão humana e robustez dos dados utilizados na modelagem.²⁷ Ainda que o *AI Act* não se debruce diretamente sobre o fenômeno do *Big Data*, esta regulação apresenta-se harmonizada frente a toda árvore regulatória da UE, que sistematiza importantes diretivas na “Era dos Dados”, a exemplo do Regulamento Geral de Proteção de Dados (GDPR), Regulamento dos Serviços Digitais (DSA), Regulamento dos Mercados Digitais (DMA), dentre outras.

Em perspectiva comparativa, o Brasil tem buscado estruturar sua abordagem regulatória relativamente quanto a um “Marco Geral de

27 A título ilustrativo, IAs de risco inaceitável são aquelas expressamente proibidas por violar direitos fundamentais, a exemplo de técnicas subliminares que alterem a consciência humana ou que explorem a vulnerabilidade das pessoas em razão de idade, incapacidade ou situação socioeconômica (art. 5º). Por sua vez, IAs de risco elevado são soluções que, em razão da função que desempenham, podem oferecer riscos significativos, a exemplo sistemas de identificação biométrica ou ainda sistemas utilizados na educação e na formação profissional, devendo, portanto, cumprir uma série de requisitos para que sejam disponibilizadas no mercado (art. 6º e Anexo III). Tem-se ainda, IAs de risco sistêmico, isto é, IAs de propósito geral que, devido à sua escala, poder computacional e capacidade de influência, representam riscos significativos, devendo, nesse sentido, cumprir critérios específicos para que sejam igualmente disponibilizados (art. 51).

IA” por meio do Projeto de Lei nº 2.338/2023. Aprovado em texto substitutivo no Senado Federal, o PL tramita atualmente na Câmara dos Deputados e visa estabelecer, de forma muito semelhante ao modelo europeu, uma regulação de IA baseada em riscos, trazendo princípios e diretrizes voltadas à responsabilidade de agentes e fornecedores de IA e à proteção dos direitos fundamentais.²⁸ O projeto também prevê a criação de um Sistema Nacional de Governança de Inteligência Artificial capaz de integrar órgãos reguladores setoriais como a Autoridade Nacional de Proteção de Dados (ANPD).²⁹ A proposta se torna especialmente relevante em um país que enfrenta desafios históricos de desigualdade estrutural e de discriminação algorítmica, sendo essencial garantir que o uso massivo de dados não reproduza, de forma automatizada, preconceitos e exclusões já presentes no tecido social.

Observa-se, ainda, de forma setorial, a atuação do Conselho Nacional de Justiça (CNJ) ao aprovar a recente Resolução n.º 615/2025, que introduz um importante avanço ao regular diretamente o uso de sistemas de IA no âmbito das atividades judiciais. O texto normativo determina que os sistemas utilizados por tribunais devem observar os princípios da supervisão humana, da transparência, da imparcialidade,

28 Conforme se destaca do art. 1º: “1º Esta Lei estabelece normas gerais de caráter nacional para a governança responsável de sistemas de inteligência artificial (IA) no Brasil, com o objetivo de proteger os direitos fundamentais, estimular a inovação responsável e a competitividade e garantir a implementação de sistemas seguros e confiáveis, em benefício da pessoa humana, do regime democrático e do desenvolvimento social, científico, tecnológico e econômico”.

29 Como se extrai do § 2º do art. 1º: “A fim de promover o desenvolvimento tecnológico nacional, o Sistema Nacional de Regulação e Governança de Inteligência Artificial (SIA) regulamentará regimes simplificados, envolvendo flexibilização de obrigações regulatórias previstas nesta Lei, nos seguintes casos: I – padrões e formatos abertos e livres, com exceção daqueles considerados de alto risco; II – incentivo à inovação e à pesquisa científica e tecnológica no ambiente produtivo, com vistas à capacitação tecnológica, ao alcance da autonomia tecnológica e ao desenvolvimento do sistema produtivo nacional e regional do País; III – projetos de interesse público, os que atendam as prioridades das políticas industrial e de ciência, tecnologia e inovação e os que visem à solução dos problemas brasileiros”.

da rastreabilidade e da segurança da informação.³⁰ Nesse sentido, o CNJ reconhece que os dados, especialmente quando processados em larga escala, possuem centralidade na produção de decisões judiciais automatizadas e devem ser auditáveis, éticos e justos. A resolução também estabelece que nenhum sistema pode substituir a autonomia da magistratura ou comprometer o devido processo legal, em atenção aos riscos que o *Big Data*, operado por IA, representa quando não controlado adequadamente.³¹ Ainda que limitada à esfera judicial, essa norma representa um esforço institucional importante de internalização de boas práticas de governança algorítmica no setor público.

Dessa forma, tanto no plano internacional como no interno, observa-se uma movimentação normativa que reconhece a complexidade do ecossistema de dados e a necessidade de abordagens regulatórias transversais. O *Big Data*, por operar sobre diversas naturezas de informações e ser insumo essencial para a modelagem de sistemas preditivos, não pode ser compreendido de forma fragmentada. A regulação eficaz exige um olhar sistêmico, que articule proteção de dados, segurança informacional, responsabilidade algorítmica e participação democrática. Os marcos apresentados, ainda que incipientes e em constante construção,

30 Como se destaca do art. 1º: “A presente Resolução estabelece normas para o desenvolvimento, a governança, a auditoria, o monitoramento e o uso responsável de soluções que adotam técnicas de inteligência artificial (IA) no âmbito do Poder Judiciário, com o objetivo de promover a inovação tecnológica e a eficiência dos serviços judiciais de modo seguro, transparente, isonômico e ético, em benefício dos jurisdicionados e com estrita observância de seus direitos fundamentais”.

31 Nesse sentido: “Art. 32. O sistema inteligente deverá assegurar a autonomia dos usuários internos, com o uso de modelos que: I – promovam o incremento da eficiência, precisão e qualidade das atividades, sem limitar a capacidade de atuação dos usuários; II – possibilitem a revisão detalhada do conteúdo gerado e dos dados utilizados para sua elaboração, assegurando que os usuários tenham acesso às premissas e ao método empregado pela inteligência artificial na sua formulação, sem que haja qualquer espécie de vinculação à solução apresentada pela inteligência artificial e garantindo-se a possibilidade de correções ou ajustes. Parágrafo único. Em nenhum momento o sistema de IA poderá restringir ou substituir a autoridade final dos usuários internos”.

apontam para uma tendência de normatização orientada não apenas à eficiência tecnológica, mas à proteção da pessoa humana em uma sociedade global profundamente orientada por dados. Como observa Dora Kaufman, “[o] desafio é evitar que a legislação chegue tarde demais, quando a tecnologia já estiver incorporada na sociedade”.³²

4. Caminhos possíveis

Diante da complexidade estrutural e dos riscos sistêmicos associados ao *Big Data*, é possível inferir que a existência de normas isoladas ou setoriais não são suficientes para endereçar a questão. A construção de uma regulação eficaz requer um arranjo normativo integrado e dinâmico, capaz de lidar com o caráter complexo das tecnologias emergentes e com a velocidade de sua evolução.³³ Nesse sentido, ao menos quatro caminhos podem ser delineados como vetores para o aprimoramento da governança dos dados em larga escala.

Um primeiro caminho diz respeito à consolidação de uma abordagem regulatória coordenada e multissetorial, que articule diferentes microssistemas jurídicos, como a proteção de dados, defesa do consumidor, propriedade intelectual e livre concorrência, sob uma lógica de complementaridade. Esse alinhamento requer não apenas diálogo entre órgãos reguladores, mas também a adoção de estruturas de interoperabilidade normativa, a exemplo da proposta brasileira de um Sistema Nacional de Governança de IA.

32 KAUFMAN, Dora. *Desmistificando a inteligência artificial*. Belo Horizonte: Autêntica, 2022, p. 163.

33 Como observam Mustafa Suleyman e Michael Bhaskar, acerca dos desafios da IA: “A regulamentação não depende somente da aprovação de uma nova lei. Ela depende também de normas, estruturas de propriedade, códigos não escritos de conformidade e honestidade, procedimentos de arbitragem, cumprimento de contratos, mecanismos de supervisão. Tudo isso precisa ser integrado, e o público precisa participar”. SULEYMAN, Mustafa; BHASKAR, Michael. *A próxima onda: inteligência artificial, poder e o maior dilema do século XXI*. Rio de Janeiro: Record, 2023, p. 287

Um segundo caminho envolve a valorização do princípio da transparência algorítmica e da auditabilidade dos sistemas que operam por meio do *Big Data*. Trata-se de assegurar que as decisões automatizadas, em especial as que afetam notadamente os direitos fundamentais, possam ser compreendidas, questionadas e eventualmente revertidas por seus titulares, a exemplo do que se verifica do art. 20 da LGPD.³⁴ Isso implica estabelecer deveres de explicabilidade com nível técnico proporcional ao grau de impacto e risco social da tecnologia, além da disponibilização de canais eficazes para a prestação de contas.

Um terceiro eixo consiste na promoção da justiça de dados (*data justice*),³⁵ conceito que reconhece a centralidade dos dados na reprodução de assimetrias históricas e busca integrar a dimensão distributiva e participativa à regulação. Para tanto, inclui-se a avaliação de impactos algorítmicos com foco em grupos vulneráveis, o incentivo à representatividade dos dados utilizados em treinamentos de IA e a criação de mecanismos de escuta social na formulação de políticas públicas orientadas por dados.³⁶

34 Nos termos do referido dispositivo, assegura-se que o direito ao titular de dados de solicitar a revisão das decisões automatizadas envolvendo o tratamento de dados que afetem notadamente os interesses do titular, incluindo aquelas voltadas para a construção de perfil pessoal, profissional, de consumo e de crédito e ainda aspectos da personalidade.

35 De acordo com a doutrina, a “justiça de dados” pode ser conceituada a partir de três pilares nevrálgicos: (i) visibilidade, no sentido dos dados serem utilizados de forma representativa e em respeito à privacidade; (ii) engajamento com as tecnologias, relativamente, quanto à autonomia individual frente às tecnologias e sistemas de dados e; (iii) antidiscriminação, de modo a prevenir usos injustos ou discriminatórios dos dados. DENCİK, Lina et al. *Towards data justice? The ambiguity of anti-surveillance resistance in political activism*. *Big Data & Society*, v. 3, n. 2, nov. 2016. Disponível em: <https://journals.sagepub.com/doi/10.1177/2053951716679678>. Acesso em: 06.07.2025.

36 A exemplo do que se observa no bojo da proposta do PL 2.338/2023, que inclui o dever de realização da Avaliação de Impacto Algorítmico: “Art. 25. A avaliação de impacto algorítmico de sistemas de IA é obrigação do desenvolvedor ou do aplicador que introduzir ou colocar sistema de IA em circulação no mercado, sempre

Por fim, um quarto caminho passa pela educação digital e pelo fortalecimento da cultura de dados, tanto entre usuários quanto entre tomadores de decisão. A compreensão crítica sobre como dados são coletados, tratados e utilizados é condição essencial para a construção de uma cidadania informada e para o exercício efetivo de direitos em uma sociedade movida por informações. Essa capacitação deve envolver não apenas a alfabetização de aspectos técnicos, mas também ético e jurídico, em um esforço transversal de políticas públicas educacionais.³⁷

A conjugação desses caminhos não se traduz em uma eliminação dos desafios, mas contribui para uma abordagem normativa mais eficaz e democrática. Regular o *Big Data* revela-se, em última instância, instrumentalizar os meios de controle sobre o poder informacional em uma sociedade global progressivamente digital. A escolha entre uma potencial opacidade de ordem tecnocrática e uma governança multissetorial participativa dependerá do enfrentamento ao tema como uma questão multifacetada, isto é, de ordem técnica e científica, mas igualmente política e moral.

Referências bibliográficas

DENCIK, Lina et al. Towards data justice? The ambiguity of anti-surveillance resistance in political activism. **Big Data & Society**, v. 3, n. 2, nov. 2016. Disponível em: <<https://journals.sagepub.com/doi/10.1177/2053951716679678>>. Acesso em: 06.07.2025.

ECO, Umberto. **Pape Satàn aleppe**: crônicas de uma sociedade líquida. Rio de Janeiro: Record, 2019.

que o sistema ou o seu uso forem de alto risco, considerando o papel e a participação do agente na cadeia”.

- 37 Como reflete Umberto Eco: “a internet oferece um repertório fantástico de informações, mas não os meios para selecioná-las, e a educação não consiste apenas em transmitir informação, mas também em ensinar critérios de seleção”. ECO, Umberto. *Pape Satàn aleppe: crônicas de uma sociedade líquida*. Rio de Janeiro: Record, 2019, p. 80.

GUILHERME, Guilherme. **Apenas 6 países têm PIB maior que valor de mercado da Apple**. Disponível em: <<https://exame.com/invest/mercados/apenas-6-paises-tem-pib-maior-que-valor-de-mercado-da-apple/amp/>>. Acesso em: 11.05.2025.

HAN, Byung-Chul. **Infocracia: digitalização e a crise da democracia**. Petrópolis: Vozes, 2022.

HARARI, Yuval Noah. **Homo Deus: uma breve história do amanhã**. São Paulo: Companhia das Letras, 2016.

JORDÃO, Fabio. **Afinal, o que é a Lei de Moore?** Disponível em: <<https://www.tecmundo.com.br/curiosidade/701-o-que-e-a-lei-de-moore-.htm>>. Acesso em: 12.05.2025.

KAUFMAN, Dora. **Desmistificando a inteligência artificial**. Belo Horizonte: Autêntica, 2022.

LUCIVERO, Federica. **Big Data, Big Waste? A Reflection on the Environmental Sustainability of Big Data Initiatives**. Science and Engineering Ethics, v. 26, 2020, p. 1009-1030. Disponível em: <<https://doi.org/10.1007/s11948-019-00171-7>>. Acesso em: 10.07.2025.

MICROSOFT. **O que é um data lake**. Disponível em: <<https://azure.microsoft.com/pt-br/resources/cloud-computing-dictionary/what-is-a-data-lake>>. Acesso em: 12.05.2025

O'NEIL, Cathy. **Armas de Destruição em Massa: como o big data aumenta a desigualdade e ameaça a democracia**. São Paulo: Rua do Sabão, 2020.

OPENAI. **What is the size of the training set for GPT-3**. Disponível em: <<https://community.openai.com/t/what-is-the-size-of-the-training-set-for-gpt-3/360896>>. Acesso em: 11.05.2025.

PASQUALE, Frank. **The Black Box Society: The Secret Algorithms that Control Money and Information**. Cambridge/London: Harvard University Press, 2015.

PIO, Juliana. **As marcas mais valiosas do mundo em 2025 (duas delas são brasileiras)**. Disponível em: <<https://exame.com/marketing/saiba-quais-sao-as-marcas-mais-valiosas-do-mundo-em-2025-duas-delas-sao-brasileiras/>>. Acesso em: 10.05.2025.

SANTOS, Andréia. O Impacto do Big Data e dos Algoritmos nas Campanhas Eleitorais. In: BRANCO, Sérgio; DE TEFFÉ, Chiara (orgs.). **Privacidade em perspectivas**. Rio de Janeiro: Lumen Juris, 2018, p. 1-24.

SHAFFER, Kris. **Data versus Democracy: How Big Data Algorithms Shape Opinions and Alter the Course of History**. Nova York: Apress, 2019.

SULEYMAN, Mustafa; BHASKAR, Michael. **A próxima onda: inteligência artificial, poder e o maior dilema do século XXI**. Rio de Janeiro: Record, 2023.

ZUBOFF, Shoshana. **A era do capitalismo de vigilância: a luta por um futuro humano na nova fronteira do poder**. Rio de Janeiro: Intrínseca, 2020.

Eleonora Mesquita Ceia · Doutora em Direito pela Universidade de Saarbrücken. Professora Titular de Direito Constitucional do Ibmec-RJ e Professora Adjunta de Teoria do Estado da Faculdade Nacional de Direito da Universidade Federal do Rio de Janeiro (UFRJ). Professora Colaboradora do Programa de Pós-Graduação em Direito da UFRJ.

Pedro Teixeira Gueiros · Mestre em Direito Civil pela PUC-Rio. Ex-bolsista da Fundação Konrad Adenauer. Professor Substituto de Direito Civil e Internacional Privado da Faculdade Nacional de Direito, UFRJ (2024-2025). Professor Assistente e Advogado Orientador do Núcleo de Prática Jurídica do Ibmec-RJ. Professor convidado dos cursos de pós-graduação da PUC-Rio, CEPED/UERJ e EMERJ. Integrante do Núcleo Legalite da PUC-Rio. Membro Suplente do Conselho Municipal de Proteção de Dados Pessoais e da Privacidade do Rio de Janeiro (CMPDPP). Graduado em Direito pelo Ibmec-RJ.

A transparência governamental no uso de Big Data

Maria Paula Almada
Isabele B. Mitozo

Resumo

A transparência governamental, entendida aqui como a disponibilidade de dados e informações governamentais públicas de qualidade, é peça essencial para o controle social e a participação cidadã. O advento dos chamados *big data* não apenas têm transformado a transparência governamental, como lhe têm imposto desafios democráticos. No Brasil, marcos como a Lei de Acesso à Informação e a LGPD regulam a abertura de dados e a proteção da privacidade, mas enfrentam algumas tensões específicas. O uso massivo de dados, muitas vezes cruzados e coletados automaticamente, amplia riscos de desinformação, vigilância e micro-targeting. Ao passo que tecnologias como mineração de dados e web scraping possibilitam maior fiscalização, otimização na análise de políticas e prestação de serviços públicos. Neste trabalho, debatemos sobre esta temática abordando também a necessidade de se atualizar legislações e promover uma cultura de dados abertos visando um equilíbrio entre transparência, proteção de dados e fortalecimento democrático.

Abstract

Government transparency, understood here as the availability of quality public government data and information, is essential for social control

and citizen participation. The advent of big data has not only transformed government transparency, but has also imposed democratic challenges. In Brazil, milestones such as the Access to Information Act and the LGPD regulate data openness and privacy protection, but face some specific tensions. The massive use of data, often cross-referenced and collected automatically, increases the risks of misinformation, surveillance, and microtargeting. At the same time, technologies such as data mining and web scraping enable greater oversight, optimization of policy analysis, and delivery of public services. In this paper, we discuss this topic, also addressing the need to update legislation and promote a culture of open data, aiming for a balance between transparency, data protection, and democratic strengthening.

O que a era das tecnologias digitais tem fornecido à humanidade? O que os chamados *big data* nos dizem e como podem ser utilizados pelo Poder Público, mantendo garantias democráticas de governança e transparência, levando em conta que eles não consistem apenas em um grande volume de dados, mas também em como são coletados, sua forma, suas fontes, sua granularidade (MERGEL *et al.*, 2016) e armazenamento? Este ensaio pretende trazer reflexões acerca dos novos moldes pelos quais os dados têm impostos novos desafios às democracias, especialmente em relação a um de seus mais caros valores: a transparência pública.

A transparência é um valor intrínseco às democracias representativas e consiste no provimento de informações de interesse público, obedecendo a critérios de atualização, completude, utilidade e relevância, o que caracteriza a informação pública como um conteúdo de boa qualidade. As informações governamentais precisam ser, nesses moldes, tornadas públicas para que seja possível ao cidadão acompanhar e supervisionar as ações governamentais ou participar da tomada de decisão pública, uma vez que é a divulgação de informações precisas, relevantes e atuais sobre as contas, as políticas e os recursos públicos que permite que os cidadãos fiscalizem e monitorem ações governamentais,

formulem opiniões e atuem de maneira mais engajada na esfera pública (GRIMMELIKHUIJSEN & MEIJER, 2012; RUIJER *et al.*, 2020). É, em suma, a transparência pública o elemento que torna governos, Estados e atores políticos mais susceptíveis ao escrutínio público e auxilia os cidadãos no momento de tomada de decisão política, quando o poder de decisão é retornado para a esfera da cidadania por meio do voto.

O primeiro passo para a transparência governamental é o acesso às informações do governo, que consiste em um direito básico dos cidadãos (BOVENS, 2002). O direito de saber (*public's right to know*), previsto nas Leis de Acesso à Informação, deve ser balanceado pelo direito de entender (*public's need to understand*). Não basta ter acesso às informações abertas previstas legalmente, é preciso compreendê-las. Esta é a diferença entre a simples oferta de informação e a transparência (NAURIN, 2007). Além disso, o aumento da visibilidade e da eficiência, a prestação de contas e até a participação, envolvimento, capacitação e aumento da confiança do cidadão são efeitos esperados (HANSEN; CHRISTENSEN; FLYVERBOM, 2015), não garantidos. Tudo, no entanto, depende do modo como esta informação lhes é oferecida.

Processos e decisões governamentais abertos são parte essencial da governança pública. Quando dados integrais sobre processos e decisões governamentais são colocados à disposição, é possível que, a partir deles, seja produzida transparência. Por outro lado, quando a integralidade dos dados não é assegurada, a simples disposição não é garantia de que, a partir deste material, seja possível produzir transparência.

Quando falamos, aqui, em disposição de dados (*openness*), referimo-nos basicamente à exposição de conteúdos em estado bruto: dados técnicos, números ou documentos que têm o propósito de expor o Estado ao máximo possível. É o Estado operando sem paredes (FENSTER, 2015). O fato de os dados serem íntegros e abertos pode exercer a importante função de instrumentalizar especialistas, jornalistas, partidos políticos, lobistas, ONGs (AMORIM, 2012; HEALD, 2006; SILVA, 2009) e até empresas privadas, que deles lançarão mão para traçar estratégias de ação e comportamento (BRIGHT *et al.*, 2015). Em suma, os

dados podem ser utilizados pelos mais diversos atores e para os mais variados propósitos, muitas vezes ignorando os direitos dos titulares destes dados, o que confirma a importância em se assegurar a proteção dos dados pessoais.

Para oferecer dados abertos, um governo precisa ter como política a exposição de informação em forma integral, sem filtros ou análises prévias. Esta é uma das etapas a serem percorridas para permitir que os cidadãos acompanhem as operações e as atividades da administração pública durante a sua execução. A disponibilidade de informações governamentais representadas em formato aberto e acessível de tal modo que possam ser reutilizadas denomina-se dados abertos governamentais (COSTA PINHO, ALMADA, 2019), que podem ser somados a outras fontes de informações, gerando novos significados. O ambiente *online* é o *locus* para a abertura dos dados governamentais: é neste ambiente que os dados estão disponíveis para que cidadãos, organizações e quaisquer partes interessadas possam fazer uso deles.

No âmbito dos governos, um marco internacional para a maior conscientização e o incentivo ao desenvolvimento de práticas de fomento à transparência foi a criação da *Open Government Partnership* (OGP), que surgiu em 2011 e hoje reúne mais de 75 países¹, tendo criado posteriormente um grupo de trabalho voltado especificamente à transparência legislativa. Àquela ocasião, o desafio da transparência pública versava, portanto, sobre a abertura de dados públicos e a promulgação de legislações que tornassem a transparência a regra, e o sigilo a exceção. Neste âmbito, diversos países realizaram discussões e aprovaram os FOIAS (*Freedom of Information Acts*), ou LAIs (Leis de Acesso à Informação), documentos exigidos para participação de uma nação como signatária da OGP.

O Brasil, por sua vez, conta com previsões legais e marcos normativos que asseguram o provimento de informações claras e completas so-

1 Open Government Partnership. Disponível em: <https://www.opengovpartnership.org/about/> Acesso em 02 jun. 2025.

bre o funcionamento do governo, seus órgãos e agências, bem como as atividades e a tomada de decisão dos parlamentares. Trata-se da Lei de Responsabilidade Fiscal (Lei complementar 101, conhecida como LRF, de 2000)², da Lei da Transparência (Lei Complementar 131, de 2009)³, e, por fim, da Lei de Acesso à Informação (Lei Complementar 12.527, conhecida como LAI, de 2011)⁴, que destaca o uso da internet como forma de dar ampla divulgação a informações⁵. Esse conjunto de dispositivos legais tem exercido impacto significativo sobre a promoção de ferramentas de transparência pública no país, tendo em vista que os órgãos públicos devem oferecer informações com qualidade assegurada por meio da transparência ativa, garantida desde a Constituição Federal de 1988, em seus canais analógicos e digitais, e passiva, modalidade que aparece na LAI, por meio de pedidos de acesso à informação não disponível. Embora tenha sido aprovada mais de 20 anos após a garantia de transparência pública estabelecida pela Constituição, a LAI traz novas diretrizes importantes para o melhor cumprimento desse valor (Carvalho *et al.*, 2022), sobretudo no contexto do mundo digital, que não existia ou era bastante diferente quando da promulgação dos demais dispositivos legais nacionais neste âmbito.

Após essa primeira fase de reflexões e ações em torno da transparência pública no século XXI, aproximando-se de 2020, os desafios postos à administração pública e à transparência pública passaram a ser outros, o que tem suscitado debates acerca da atualização das LAIs. A lida

2 BRASIL. **Lei Complementar nº 101, de 4 de maio de 2000**. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/lcp/lcp101.htm. Acesso em: 02 jun. 2025.

3 BRASIL. **Lei Complementar nº 131, de 27 de maio de 2009**. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/lcp/lcp131.htm. Acesso em: 02 jun. 2025.

4 BRASIL. **Lei Complementar n. 12.527, de 18 de novembro de 2011**. 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm Acesso em: 02 jun. 2025.

5 A transparência passiva, que acontece via pedidos de acesso às informações, está prevista na LAI, enquanto a ativa se dá por meio da divulgação de dados em portais governamentais em tempo real, sem necessitar da demanda dos cidadãos, prevista na LRF e na Lei da Transparência.

com um volume cada vez maior de dados, os chamados *Big Data*, tem sido um deles, assim como a discussão acerca de sua abertura, sobretudo porque a informação é um recurso incontestável de poder (RUIJER *et al.*, 2020, p.261), especialmente quando é abundante. Esses dados podem ser conceituados a partir de diferentes perspectivas e características, de modo que Mergel e colegas os definem da seguinte maneira:

[...] grande volume de dados que frequentemente são a combinação de dados administrativos altamente estruturados coletados ativamente por organizações do setor público, com dados estruturados e não estruturados coletados de forma contínua e automática em tempo real, que muitas vezes são criados passivamente por entidades públicas e privadas, a partir de suas interações online (MERGEL *et al.*, 2016, p.932).

Desse modo, pode-se compreender que os dados públicos atualmente podem ser construídos a partir de diversas fontes, combinando tipos distintos de informações disponíveis, sobretudo, no ambiente online.

1. Armazenamento e processamento de grandes dados: segurança e governança

O crescimento exponencial de dados digitais – inclusive pessoais – disponíveis online, somado à expansão do mercado de análise de dados por governos e empresas, trouxe desafios significativos para a proteção de direitos fundamentais. Identificadores como e-mail, CPF (no caso brasileiro), biometria e localização são cruciais no cruzamento de bases em larga escala (*Big Data*), com potencial para alimentar campanhas de desinformação, *microtargeting* eleitoral e perseguição política, como nos casos da Cambridge Analytica e dos disparos segmentados no WhatsApp.

No setor público, o acesso a serviços – principalmente digitais – depende da entrega de dados pessoais pelos cidadãos. Por isso, é essencial

garantir a estes a segurança (acesso) e a proteção (uso) destes dados, sobretudo em contextos em que a transparência pública pode gerar exposição indevida, como na divulgação de beneficiários de programas sociais, por exemplo. É preciso assegurar que dados pessoais não serão explorados para outras finalidades. A coleta de dados por meio de aplicativos e portais públicos deve exigir regras claras de consentimento, uso e controle por parte dos cidadãos.

Por outro lado, tem-se o fato de que governos de todos os níveis estão hoje entre os maiores produtores de dados e muitos desses dados são abertos e estão disponíveis a quaisquer interessados, o que é, inclusive, resguardado pelo princípio da transparência pública (se o dado em questão, vale ressaltar, dizer respeito àquilo que é de interesse público). Muito do que se usa para a construção de bancos de dados governamentais passa pela recolha de bases informacionais produzidas por outros setores, ou mesmo pelos próprios cidadãos. Nesse âmbito, a prática mais corrente de governos tem sido a coleta de dados para fins técnicos, sobretudo de segurança.

Sabemos que um movimento decorrente da recolha de dados é a sua abertura. Como já argumentamos aqui, os dados abertos governamentais são ferramentas estratégicas importantes para governos, uma vez que possibilitam o desenvolvimento de ferramentas melhores para a fiscalização do Estado, o que pode garantir processos mais *accountable*. Mas, além dos benefícios associados à abertura de dados públicos, há os riscos de exposição e compartilhamento de dados pessoais dos cidadãos. No contexto brasileiro, vimos em muitos casos a Lei Geral de Proteção de Dados (LGPD)⁶ ser usada como prerrogativa por órgãos públicos quando estes decidem não divulgar dados e informações que lhes foram solicitados. Por esta razão, pesquisadores têm apontado para a importância de se atualizar e ajustar as legislações que regem a trans-

6 BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/lei/l14020.htm. Acesso em 10 de junho de 2025.

parência, em especial a LAI, de modo a evitar interpretações errôneas ou manobras jurídicas a fim de descumprir diretrizes fundamentais da transparência pública.

Em nível mundial, nos últimos quinze anos, vimos a transição de um foco exclusivo em transparência pública para a incorporação da proteção de dados como valor central. Embora haja tensões entre esses princípios, eles não são excludentes: a transparência sobre o tratamento de dados pessoais é, inclusive, um dos pilares da proteção de dados. Entretanto, a aplicação prática dessa conciliação pode gerar conflitos entre diferentes culturas institucionais – como a de jornalistas investigativos e ativistas da privacidade.

Aqui, entendemos que a privacidade vai além da ausência de vigilância direta. O problema não está em um dado isolado, como uma imagem em praça pública, mas na capacidade de se cruzar informações e compor perfis detalhados, como ocorre com sistemas de reconhecimento facial. Transparência, acesso à informação e proteção de dados são pilares das democracias contemporâneas e devem ser garantidos simultaneamente, permitindo que cidadãos fiscalizem o poder público enquanto preservam suas liberdades individuais frente ao uso indevido de seus dados por governos ou corporações.

Deve-se, portanto, refletir se há e quais são as tensões e complementaridades entre a promoção da transparência pública e a proteção de dados pessoais, assegurada pela LGPD. Nesse sentido, sabe-se que as instituições públicas são constrangidas pelos dispositivos legais que regulam todo o processamento de dados públicos. Todavia, isso não significa que os grandes bancos de dados não possam ser analisados a partir do cruzamento de variáveis que podem mostrar como transformá-los em informações para basear políticas, por exemplo, sem ferir qualquer diretriz estabelecida pela legislação vigente. Do contrário, como afirmam Ruijer e colegas (2022, p. 316), *“Choices about what data to collect on individuals and groups, on how to store, analyze, and use data for specific purposes can lead to inequities in the management of public services”*.

2. O grande volume de dados como um problema de acesso à informação

No mundo dos dados, exige-se cada vez mais transparência, mas nem sempre as consequências de ser mais transparente são positivas. Prova disso é o crescente descrédito nas instituições a partir dos casos de corrupção, que são revelados quando se lança luz sobre dados, informações, atores políticos, tomadas de decisão etc. Se, por um lado, a transparência contribui para o constrangimento dessas práticas, por outro, não é suficiente para cercear esse tipo de comportamento, e aqui chamamos a atenção a um aspecto salutar, qual seja, o excesso de informação produzida no ambiente digital e suas consequências.

O acesso à informação nunca foi tão ressaltado, sobretudo porque esse recurso flui a partir de diversos canais, não somente das instituições do Estado. As redes de comunicação digital, especialmente as redes sociais digitais, têm capacidade de disseminação, velocidade e alcance sem precedentes. Um exemplo das consequências desse potencial é a chamada *infodemia*, termo usado pela Organização Mundial da Saúde durante a pandemia da Covid-19 (OPAS, 2020), típica de momentos emergenciais e tragédias. Ela é caracterizada pelo aumento excessivo no volume e na circulação de informações, corretas ou não, sobre um tema específico e pode: (1) atrapalhar a confiança nas organizações oficiais; (2) gerar dissonância cognitiva, caracterizada pela dificuldade de identificar as informações verdadeiras; (3) atrasar o processo de tomada de decisão dos governos, profissionais e população em momentos de crise.

Desse modo, governos precisam lidar com o fato de que, ao se tornarem mais transparentes, sobretudo a partir da ampla organização e divulgação dos grandes dados públicos, precisam também atuar para garantir a promoção de mecanismos de educação política, a fim de permitir que esses dados sejam lidos corretamente pelos cidadãos, uma vez que dados abertos nem sempre são intuitivos para o público geral, e nem todos os dados e informações disponíveis são relevantes do ponto de vista do interesse público.

3. IA como instrumento de ampliação da transparência

Grandes dados só podem ser processados por máquinas. Desse modo, a inteligência artificial se torna uma consequência do volume e da complexidade dos processos que envolvem as informações armazenadas pelos governos. Essas tecnologias possibilitam o processamento de grandes volumes de dados em tempo real, auxiliando a administração pública a tomar decisões informadas e personalizar a prestação de serviços para diferentes segmentos da população.

A administração pública e os governos de todo o mundo têm avançado na utilização de novos instrumentos e recursos para a promoção da transparência, impulsionados pela mineração de dados (*data mining*) e por aplicativos de monitoramento e raspagem de dados (*data scraping*) (GOMES, AMORIM, ALMADA; 2018). Em países como o Brasil, isso pode significar a democratização do acesso a serviços públicos, reduzindo as barreiras impostas por fatores geográficos e socioeconômicos (MENDONÇA, FILGUEIRAS, ALMEIDA, 2023). Essa evolução busca tornar a gestão pública mais aberta, acessível e fácil de ser fiscalizada pela sociedade.

Especificamente, a mineração de dados (GOLDSCHMIDT, PASSOS, BEZERRA; 2015) consiste basicamente em analisar grandes volumes de informações para descobrir padrões, tendências e relações ocultas e pode ser aplicada por governos para:

- Identificação de Fraudes e Irregularidades: Órgãos de controle e tribunais de Contas podem utilizar a mineração de dados para cruzar informações de diferentes bancos de dados (licitações, pagamentos, contratos, folhas de pessoal) e identificar indícios de desvio de recursos, superfaturamento, nepotismo ou outras irregularidades, possibilitando uma fiscalização mais direcionada e eficiente.
- Análise de Gastos Públicos: A mineração de dados auxilia na análise de grandes conjuntos de dados de despesas públicas, permitindo que a sociedade e os órgãos de controle compreendam melhor onde

e como o dinheiro público está sendo gasto. Isso pode revelar padrões de consumo, ineficiências ou desvios.

- **Monitoramento de Políticas Públicas:** Ao analisar dados sobre a execução de programas sociais, investimentos em infraestrutura ou outras políticas públicas, a mineração de dados pode ajudar a avaliar a efetividade dessas ações e a identificar áreas que precisam de melhoria.

Ao passo que o *web scraping* (ASSIS, GOMIDE; 2021), consiste na técnica de extrair informações de websites de forma automatizada. Ao ser combinada com aplicativos de monitoramento, ela se torna uma ferramenta poderosa para a transparência ao possibilitar:

- **Coleta e Consolidação de Dados de Portais da Transparência:** No caso brasileiro, muitos governos e órgãos públicos no Brasil possuem seus próprios Portais da Transparência. Ferramentas de *data scraping* são utilizadas para coletar e consolidar esses dados, que muitas vezes estão dispersos e em formatos variados. Isso permite a criação de bases de dados mais completas para análise.
- **Criação de Ferramentas de Controle Social:** Diversas iniciativas da sociedade civil e de organizações não governamentais utilizam o *data scraping* para extrair dados públicos e criar plataformas e aplicativos que facilitam o acesso e a compreensão das informações por parte do cidadão. Exemplos: Portal Brasileiro de Dados Abertos, “Operação Serenata de Amor” e Radar da Transparência Pública (ATRICON).
- **Fiscalização de Contratos e Licitações:** Aplicativos podem raspar dados de diários oficiais e portais de licitações para monitorar contratos, identificar aditivos suspeitos, atrasos ou outras inconsistências.
- **Monitoramento em Tempo Real:** A capacidade de automatizar a coleta de dados permite um monitoramento mais frequente e em tempo real, alertando sobre novas informações ou alterações em dados já existentes, colaborando para uma fiscalização eficaz.

Como vimos, são diversas as aplicações de IA, de maneira mais ampla, para a administração pública com o objetivo do aumento da transparência. A Ouvidoria do Estado de Goiás, por exemplo, utiliza de IA para melhorar o serviço de Pedido de Acesso à Informação (PAI), facilitando as demandas dos cidadãos. O serviço, que anteriormente “exigia que o usuário soubesse diferenciar oito tipologias de manifestações de ouvidoria e depois escolher, entre 48 órgãos e entidades, aquele que seria o responsável pela sua demanda e ainda mais de 3 centenas de possíveis assuntos” (CARVALHO *et al.*, 2022, p.231), hoje consegue identificar no conteúdo enviado pelo cidadão, por meio de Processamento de Linguagem Natural, qual é a intenção do usuário e a qual órgão pertence sua demanda⁷. Além de permitir um atendimento mais fácil, ela também realiza um acompanhamento da evolução de cada manifestação, disponibilizado a quem faz a solicitação, mostrando que esse tipo de iniciativa pode contribuir com a facilitação do dispositivo de transparência passiva e a *accountability* acerca de todo o processo.

Além de auxiliar os cidadãos com suas demandas às instituições, uso comum de IA no serviço público como mostra o exemplo acima e a recorrente associação desse uso à adoção de *chatbots*, essas ferramentas poderiam ser usadas, ainda, para organizar as próprias informações dos órgãos públicos e identificar o que lhes falta disponibilizar de modo ativo em seus canais de comunicação à sociedade.

Por outro lado, qualquer ação que leve alguma tomada de decisão a passar por ferramentas de aprendizado de máquinas exige estabelecimento de princípios e monitoramento para esse trabalho, uma vez que, se as instâncias envolvidas no processo não compreendem “o que está acontecendo nos complexos processos a partir dos quais as decisões são tomadas, é impossível haver *accountability*” (MENDONÇA, FILGUEIRAS, ALMEIDA, 2023, p.141).

7 Disponível em: www.expresso.go.gov.br. Acesso em 2 de junho de 2025.

4. Considerações finais

As oportunidades para o uso de *big data* no aprimoramento da transparência pública são imensas, mas, junto com elas, os desafios se mostram igualmente iminentes. Primeiramente, é importante que se compreenda dados abertos governamentais e transparência pública como uma cultura administrativa, ou seja, sociedade e governo devem entender a transparência dos seus dados e informações como regra e sigilo como exceção. Dito isso, entendemos que a continuação do investimento em tecnologias de dados, a colaboração entre governo e sociedade civil, e a promoção de uma cultura de dados abertos podem revolucionar a transparência no Brasil, fortalecendo a democracia e o controle social. Cabe ressaltar, ainda, que a abordagem adequada para a análise dos grandes dados governamentais deve considerar a combinação de diferentes áreas e expertises, tendo em vista o aumento das complexidades envolvendo a governança que o setor público tem vivido.

A maioria da coleta de dados é realizada pelo Estado sem que haja uma preocupação mais que técnica. Todavia, se o poder público se apodera dos dados para gerar informações a partir deles, então falamos de possibilidades analíticas para a construção de políticas. Se as *big techs* usam informações cruzadas para treinar algoritmos e fazê-los prever nossas preferências, o Estado também pode realizar determinados cruzamentos a fim de reconhecer o perfil de seus cidadãos em relação a múltiplas variáveis (como perfil de saúde, acesso a benefícios sociais, empregabilidade etc.) e “prever o presente” (CHOI e VARIAN, 2012), ou seja, fazer diagnósticos precisos do presente e ser capaz de prever o que pode ser melhor no futuro imediato para a população.

Desse modo, pode-se perceber que a transparência, sobretudo a partir do uso mais amplo do ambiente online para sua promoção, abrange novas práticas e, conseqüentemente, desafios que extrapolam sua definição inicial. Isso significa dizer que, em breve, será inadiável repensar a legislação que a regula, que tem se mostrado em alguns casos insuficiente para lidar com questões específicas, como a lógica al-

gorítmica da circulação e de informação, a proteção de dados e o uso de inteligência artificial para análise de dados, a fim de que esse valor democrático siga contribuindo para a fiscalização das práticas e a credibilidade do regime.

Referências

AMORIM, P. K. **Democracia e Internet**: a transparência de gestão nos portais eletrônicos das capitais brasileiras. 2012. Dissertação (Mestrado) – Universidade Federal da Bahia, Salvador, 2012.

ASSIS, W. V.; GOMIDE, J. V. B. Web scraping em dados públicos: método para extração de dados dos gastos públicos dos vereadores da Câmara Municipal de Belo Horizonte. **Informação & Informação**, 26(4), p. 319–341, 2021. <https://doi.org/10.5433/1981-8920.2021v26n4p319>

BOVENS, M. Information rights: citizenship in the information society. **The Journal of Political Philosophy**, v. 10, n. 3, p. 317–341, 2002.

BRASIL. **Lei Complementar nº 101, de 4 de maio de 2000**. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/lcp/lcp101.htm. Acesso em: 02 jun. 2025.

BRASIL. **Lei Complementar nº 131, de 27 de maio de 2009**. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/lcp/lcp131.htm. Acesso em: 02 jun. 2025.

BRASIL. **Lei Complementar n. 12.527, de 18 de novembro de 2011**. 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm Acesso em: 02 jun. 2025.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD)**. Brasília, DF: Presidência da República. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/lei/l14020.htm. Acesso em 10 de junho de 2025.

BRIGHT, J.; MARGETTS, H.; WANG, N.; HALE, S. **Explaining usage patterns in Open Government Data**: the case of Data.gov.uk. [S.l.]: [s.n.], 2015.

CARVALHO, Danilo B.G.; ROLIM, B.; BUENO, F.; SILVA Jr., A.J. Inteligência Artificial como ferramenta de acessibilidade à informação: o caso do estado de Goiás. In: GERALDES, Ellen *et al.* (Orgs.). **Dez anos da Lei de Acesso à Informação**: Limites, perspectivas e desafios. São Paulo: Intercom, 2022. pp. 230-240.

CHOI, Hyunyoung; VARIAN, Hal. Predicting the Present with Google Trends. Supplement 1. **Economic Record**, n.88, p. 2–9, 2012.

COSTA PINHO, M. D., & ALMADA, M. P. Governo aberto e dados abertos governamentais: um mapeamento e sistematização da produção acadêmica. **Comunicação & Inovação**, 20 (43). <https://doi.org/10.13037/ci.vol20n43.5711>, 2019.

FENSTER, M. Transparency as a theory of communication. In: **3rd Global Conference on Transparency Research**, 2013. Anais [...]. [S.l.]: [s.n.], 2013. p. 1–24.

GOLDSCHMIDT, R., PASSOS, E., BEZERRA, E. **Data mining**: conceitos, técnicas, algoritmos, orientações e aplicações. 2. ed. Rio de Janeiro: LTC, 2015. E-book. ISBN 9788535278231.

GOMES, W.; AMORIM, P. K.; ALMADA, M. P. Novos desafios para a ideia de transparência pública. **E-Compós: Revista Da Associação Nacional Dos Programas de Pós-Graduação**, 21(2), 21. <https://doi.org/10.1016/j.cyto.2015.08.171>. 2018.

GRIMMELIKHUIJSEN, S.; MEIJER, A. Effects of transparency on the perceived trustworthiness of a government organization: Evidence from an online experiment. **Journal of Public Administration Research and Theory**, 24, p. 137–157, 2012.

HANSEN, H. K.; CHRISTENSEN, L. T.; FLYVERBOM, M. Introduction: logics of transparency in late modernity: paradoxes, mediation and governance. **European Journal of Social Theory**, v. 18, n. 2, p. 117–131, 2015. Disponível em: <https://doi.org/10.1177/1368431014555254>. Acesso em: 9 jun. 2025.

HEALD, D. Varieties of transparency. In: HOOD, C.; HEALD, D. (org.). **Transparency: the key to better governance**. Oxford: Oxford University Press, 2006. p. 25–43.

MENDONÇA, R. F.; FILGUEIRAS, F.; ALMEIDA, V. **Algorithmic Institutionalism**: the changing rules of social and political life. 1.ed. Oxford: Oxford University Press, 2023. v. 1. 191p.

MERGEL, Ines; RETHEMEYER, R.K.; ISETT, Kimberley. Big Data in Public Affairs. **Public Administration Review**, vol.76, n.6, p. 928–37, 2016.

OMS, 2020.

NAURIN, D. Transparency, publicity, accountability – the missing links. **Swiss Political Science Review**, mar. 2007. p. 1–12. Disponível em: http://www.mzes.uni-mannheim.de/projekte/typo3/site/fileadmin/research_groups/2/Uppsala2007/Paper_Naurin_II.pdf. Acesso em: 9 jun. 2025.

OPAS – Organização Pan-Americana da Saúde. Fatos sobre a Infodemia. Brasília: OPAS, 2020. Disponível em: https://iris.paho.org/bitstream/handle/10665.2/52054/Factsheet-Infodemic_por.pdf. Acesso em: 9 jun. 2025.

RUIJER, Erna *et al.* The Politics of Open Government Data: Understanding Organizational Responses to Pressure for More Transparency. **American Review of Public Administration**, Vol. 50 (3), p. 260–274, 2020.

Ruijter, Erna; Porumbescu, Gregory; Porter, Rebecca; Piotrowski, Suzanne. Social equity in the data era: A systematic literature review of data-driven public service research. **Public Administration Review**, n.83, p. 316–332, 2023.

SILVA, S. P. da. Estado, democracia e internet: requisitos democráticos e dimensões analíticas para a interface digital do Estado. 2009. **Dissertação (Mestrado) – Universidade Federal da Bahia**, Salvador, 2009.

Maria Paula Almada · Co-diretora Executiva do Aláfia Lab. Doutora em Comunicação e Cultura Contemporâneas pelo Póscom-UFBA.

Isabele B. Mitozo · Professora do Departamento e do Programa de Pós-Graduação em Ciência Política da Universidade Federal de Minas Gerais. Doutora em Ciência Política pela UFPR. Coordenadora do Grupo de Pesquisa em Democracia e Comunicação Digital (DECODE).

***(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

A pesar 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.

Segurança pública na era do Big Data: policiamento algorítmico, suas limitações operacionais e seus dilemas éticos

Daniel Edler

Resumo

Novas tecnologias de segurança têm transformado a rotina de forças policiais, trazendo mais eficácia para o serviço de investigação e permitindo ao policiamento ostensivo aumentar a eficiência no controle preventivo do crime. Avanços recentes na área de ciência de dados, além da disponibilidade de sistemas de processamento computacional mais baratos, permitem que departamentos de polícia aprimorem o serviço prestado à população, auxiliando as forças de segurança a “fazerem mais com menos”. Contudo, essas inovações também criam desafios na prática policial. Além de apresentar aplicações recentes de *big data* no campo do policiamento e suas limitações operacionais, esse artigo debate os riscos da vigilância em massa e do aprofundamento de padrões discriminatórios nas estratégias de controle do crime.

Abstract

New security technologies have transformed the routine of police forces, making investigative work more effective and allowing visible policing to increase efficiency in preventive crime control. Recent advances in data science, in addition to the availability of cheaper computer pro-

cessing systems, allow police departments to improve the service provided to the population, helping security forces to “do more with less.” However, these innovations also create challenges in police practice. In addition to presenting recent applications of big data in the field of policing and their operational limitations, this article discusses the risks of mass surveillance and the deepening of discriminatory patterns in crime control strategies.

1. Introdução

As vantagens de inovações tecnológicas no campo da segurança pública são colhidas todos os dias nas grandes cidades brasileiras. Por exemplo, até poucos anos atrás, sistemas de videomonitoramento tinham utilidade bastante restrita, oferecendo imagens de baixa resolução, o que limitava a polícia a análises superficiais sobre a dinâmica criminal. A dificuldade de armazenamento de dados também representava um desafio, já que novos registros eram muitas vezes gravados sobre fitas VHS utilizadas nos dias anteriores. Se a polícia não agisse rápido, perdia as evidências. Atualmente, não apenas as câmeras espalhadas pela paisagem urbana produzem imagens em alta resolução, como sistemas de processamento de vídeo em tempo real são capazes de gerar alertas para situações suspeitas e identificar indivíduos envolvidos em ações delituosas. Ou seja, as polícias militares podem responder mais rapidamente aos eventos criminais e as polícias civis são capazes arrolar evidências robustas em seus inquéritos enquanto economizam horas de serviço. Esses avanços se devem, em grande medida, à disponibilidade massiva de dados e à possibilidade de processamento algorítmico para a identificação de padrões criminais e eventos de interesse ao trabalho policial. Em outras palavras, as inovações descritas acima se devem ao advento do *big data*.

Big Data pode ser definido como a capacidade de coletar, armazenar, processar e analisar grandes volumes de dados, frequentemente em tempo real, a partir de fontes diversas e heterogêneas (MAYER-SCHÖNBERGER

& CUKIER, 2017). Sua principal característica está na combinação dos chamados cinco “Vs”:

1. Volume: a escala dos dados coletados e disponibilizados para análise e processamento é cada vez maior, sendo medida não mais em *kilobytes* ou *megabytes*, mas, geralmente, em *petabytes* (10^{15} bytes) e *exabytes* (10^{18} bytes).
2. Velocidade: o processamento de dados ocorre em tempo real, o que permite a atualização de análises em poucos segundos, aumentando, por exemplo, a consciência situacional da polícia e sua capacidade de intervenção.
3. Variedade: os dados coletados podem ser estruturados (e.g., uma base de dados como nome e idade de indivíduos condenados pela justiça) e não-estruturados (e.g., vídeos produzidos por sistemas de monitoramento).
4. Veracidade: a maior disponibilidade de dados não vem necessariamente com maior certeza sobre sua qualidade e acurácia, sendo necessário avaliar a adequação dos dados à análise pretendida. Porém, o cruzamento de diferentes bases pode proporcionar ganhos de confiança sobre a informação analisada.
5. Valor: as vastas bases de dados tanto têm valor em si, formando um enorme mercado secundário de *data brokers*, como agregam enorme valor a processos de decisão nos setores público e privado.

Frente a esse cenário, entusiastas apontam que o *big data* deve revolucionar a maneira como trabalhamos, nos relacionamos e pensamos (DOMINGOS, 2015; FRY, 2019). Como descrevem Mayer-Schönberger e Cukier (2013, p. 132):

Antigas certezas estão sendo questionadas. O *big data* requer nova discussão quanto ao caráter das tomadas de decisões... Uma visão de mundo que pensávamos estar relacionada com causas é desafiada pela preponderância das correlações. A posse do conhecimento, que já signi-

ficou o entendimento do passado, se transforma na capacidade de prever o futuro.

A definição de *big data* pode ser relativamente simples, mas entender seu impacto na sociedade é bem mais controverso. Dois pontos levantados na descrição de Mayer-Schönberger e Cukier (2013) são centrais para a discussão: a produção de conhecimento a partir de correlações e a capacidade de previsão de fenômenos sociais.

No primeiro caso, alguns autores apontam que o *big data* induz novas abordagens epistemológicas sobre formas de conhecer o mundo (KITCHIN, 2014), o que, quando levado ao extremo, gera argumentos sobre o “fim da teoria”, já que o “dilúvio de dados torna[ria] o método científico obsoleto” (ANDERSON, 2008). Nessa perspectiva, a profusão de dados nos permitiria abandonar a obsessão por compreender mecanismos causais e abraçar um conhecimento pautado exclusivamente na identificação de padrões (na natureza e na sociedade) e em suas mais variadas correlações (ZAVRŠNIK, 2018). O cruzamento de *exabytes* de dados sobre consumo e meteorologia, por exemplo, nos permitiria descobrir que indivíduos comprem mais doces em dias de chuva. Essa correlação pode indicar que as condições climáticas causam variações na dieta da população, mas isso não é necessariamente verdade e nem relevante. O que importa para um gerente de mercado, digamos, é que ele deve ter um bom estoque de doces nos dias em que a previsão for de chuva.

No campo da segurança pública, o foco em correlações tem transformado a compreensão sobre o tipo de informação considerada relevante para o trabalho policial, o que leva à pressão pela expansão dos dados disponibilizados para as forças de segurança. Dito de outra forma, se antes serviços de inteligência e análise criminal focavam em diminuir o tamanho do palheiro para tornar mais eficiente a busca pela agulha, atualmente não se dispensa nada. Como podemos encontrar correlações surpreendentes entre variáveis que julgávamos distantes, qualquer dado é a priori importante para o combate ao crime. A busca

por suspeitos de lavagem de dinheiro pode obter avanços, por exemplo, pelo cruzamento de dados da Receita Federal, de empresas de cartões de crédito, instituições bancárias, redes sociais, buscadores na internet (e.g., Google), e serviços de compras online (e.g., Amazon). De modo semelhante, a coleta massiva de dados meteorológicos, de transporte e de interações nas redes sociais, somado ao uso de bases cartográficas com geolocalização de escolas, parques e hospitais, pode levar à descoberta de padrões criminais antes ignorados. Nessa nova forma de “conhecer” o crime, importam menos as causas do comportamento delituoso e mais as correlações que permitem a identificação de perfis ou locais de risco (WEISBURD, 2015).

E é justamente a lógica do risco que nos leva para o segundo ponto da definição de Mayer-Schönberger e Cukier (2013). Embora a experiência policial e os dados criminais já permitissem identificar determinados padrões, como concentração de roubos em áreas de grande circulação e em dias de pagamento, o cruzamento de novas bases de dados promete ir além na granularidade das informações fornecidas, permitindo a polícia a passar de estratégias preventivas e investigativas, para ações preditivas (EDLER & LOBATO, 2021). Por exemplo, ao cruzar dados de desempenho escolar, de serviços de assistência social e de saúde, algumas empresas prometem construir avaliações de risco sobre a chance de determinado jovem cometer crimes no futuro próximo (STATEWATCH, 2025). Assim sendo, o poder público pode acionar um conjunto de medidas que intervenham na situação desse jovem antes que o crime seja, de fato, cometido. Valendo-se de dados semelhantes, forças policiais têm buscado construir sistemas que avaliam também a chance de indivíduos integrarem redes criminais (DODD, 2018). Policiais, investigadores, promotores e juízes têm usado essa informação para auxiliar em processos de tomada de decisão, seja concentrando esforços na busca por mais evidências, seja alertando as patrulhas locais (RYBERG & ROBERTS, 2022).

Observando esse contexto, profissionais do sistema de justiça criminal e desenvolvedores de aplicações de *big data* para a segu-

rança pública argumentam que as inovações recentes representam a substituição do “faro policial” e da experiência construída na rotina de trabalho pelo conhecimento objetivo sobre a ação criminal que brotaria a partir dos dados, de modo que o policiamento das grandes cidades se tornaria mais preciso, eficaz e justo (BECK & MCCUE, 2009).

No entanto, essa “epistemologia empirista” e o entusiasmo em torno da possibilidade de prever e se antecipar ao crime se baseiam em uma série de premissas problemáticas (KITCHIN, 2014), entre elas: (1) a ideia de que as bases de dados usadas para alimentar os algoritmos são capazes de capturar e descrever todas as dimensões do fenômeno que buscam analisar; (2) o argumento de que o conhecimento surge puramente a partir dos dados, dispensando qualquer tipo de teoria social sobre o crime, de modo que os dados “falariam por si sós”; e (3) uma percepção de que os métodos de análise transcenderiam o conhecimento especializado, o que permitiria a qualquer um tirar conclusões acuradas a partir de *dashboards* de visualizações de dados. Essas premissas emprestam uma imagem de objetividade às soluções de *big data* empregadas por forças policiais, negligenciando o fato de que os dados são eles próprios produzidos por ferramentas políticas (i.e., decorrem do que queremos e conseguimos compilar) (BOWKER & STAR, 2000). Como resume Završnik (2018, p. 5), “estatísticas são produzidas por humanos e para humanos”.

Para discutir sobre o impacto de novas aplicações de *big data* no campo da segurança pública, esse artigo se divide em duas partes, além dessa introdução e de uma breve conclusão. Na próxima seção, apresento aplicações de *big data* que permitiram o desenvolvimento de duas tecnologias que têm sido alvo de largos investimentos pelas forças policiais, os sistemas de alerta para dinâmicas criminais e as ferramentas de policiamento preditivo. Em seguida, o artigo levanta alguns dos dilemas éticos que surgem ou se aprofundam com o uso de *big data* em ações de controle do crime, entre eles a automação de práticas discriminatórias e os riscos da vigilância intrusiva para a democracia.

2. Novas tecnologias e novas formas de policiamento

Nos últimos anos, diversas metrópoles brasileiras investiram na construção de centros de comando e controle, onde as forças policiais operam uma panóplia de sistemas de pronta-resposta e vigilância. Entre os dispositivos à disposição dos agentes estão câmeras que realizam cercamento eletrônico – capazes de identificar em tempo real placas de veículos e o rosto de indivíduos suspeitos – e softwares de mapeamento de ocorrências e análise criminal. Em conjunto, essas “soluções inteligentes” têm afetado estratégias de patrulhamento e ampliado os horizontes de políticas punitivas (PERON & ALVAREZ, 2021). Através de exemplos de casos concretos, esta seção aborda brevemente como o *big data* e a automação algorítmica têm prestado auxílio à atividade policial ao identificar áreas e alvos prioritários para o policiamento.¹

Em 2014, o estado de São Paulo anunciou a implementação do *Detecta*, uma ferramenta de monitoramento que combinava circuitos de câmeras de vigilância, a busca em múltiplas bases de dados (incluindo registros de ocorrência, identificação civil, etc.) e a integração com os demais sistemas empregados pelas polícias do estado.² Inspirado no *Domain Awareness System* (DAS), plataforma de integração de dados, mapeamento criminal e videomonitoramento desenvolvido no departamento de polícia de Nova York, o *Detecta* tinha como objetivo principal disparar alertas automatizados para a polícia paulista ao identificar dinâmicas de interesse (EDLER et al., 2023).

1 Automação algorítmica se refere a sistemas capazes de realizar uma leitura da imagem em tempo real, identificando-a com padrões previamente classificados de interesse. A partir desse “reconhecimento” automatizado, um alerta é criado para anunciar a ocorrência de determinado evento para o usuário, como uma conduta suspeita praticada em um determinado perímetro de abrangência da câmera, a invasão desse perímetro, ou a identificação de objetos de interesse, como facas e armas de fogo.

2 Entre os múltiplos sistemas integrados ao *Detecta* estavam: Infocrim (1999), FotoCrim (2002), o COPOM Online (2002) e o Sistema Omega (2003).

Com a promessa de ser o “sistema nervoso” da polícia, o *Detecta* teve um custo aproximado de 28 milhões de reais.³ O alto valor se justificava pela perspectiva de ganhos para a segurança pública, já que além de ser capaz de identificar veículos roubados e pessoas com pendências na justiça, o sistema permitia ainda o reconhecimento de “comportamentos suspeitos” que poderiam requerer atenção especial (e.g., indivíduos caminhando de capacete na calçada, motos com passageiros emparelhadas com carros, etc.). A expectativa era que, na medida em que mais crimes eram registrados pelas câmeras de segurança, o sistema se tornaria mais acurado para identificar as dinâmicas delituosas no estado, podendo reconhecer, por exemplo, abordagens suspeitas e a presença de armas de fogo. Além disso, a disseminação das câmeras no espaço urbano permitiria seguir o indivíduo suspeito durante sua movimentação, fornecendo informações preciosas para que a polícia realizasse a prisão. Como descreveu o então governador: “Antes, [o centro de operações] era um arquipélago isolado e agora integramos todos os bancos de dados das polícias civil, científica e militar. O sistema [*Detecta*] é um ‘*Big Data*’ da polícia, extremamente eficiente” (SÃO PAULO, 2016).

Contudo, as promessas em torno dos ganhos operacionais proporcionados pelo *Detecta* não se confirmaram. Em poucos meses, uma auditoria do Tribunal de Contas do Estado (TCE) revelou que o sistema era pouco usado pela polícia, já que enfrentava dificuldades técnicas na integração com os demais sistemas da instituição (TCE, 2017). Além disso, os ganhos com alertas automatizados para atividades suspeitas foram em muito suplantados pelo problema dos “falsos positivos”. O sistema gerava uma quantidade enorme de alertas que se mostravam infundados, o que sobrecarregava os operadores e, em muitos casos, atrapalhava a rotina de monitoramento urbano. Esse problema veio à público durante uma demonstração das funcionalidades do *Detecta* para jornalistas realizada pela secretaria de segurança. Na ocasião, o sistema gerou

3 Valor referente a contratos da PRODESP com a Microsoft e contratos da SSP para compra de equipamentos e prestação de serviços (TCE, 2017).

alerta para “veículo em situação suspeita” no caso de um carro supostamente estacionado em via expressa. Entretanto, após análise do operador, verificou-se que se tratava apenas da sombra de um poste na via (PAGNAN & BARBON, 2017). Segundo o TCE, disputas políticas pelo controle de dados entre as diferentes agências de segurança pública também impediram o fluxo de informações, o que limitou a capacidade de análise do sistema e gerou lentidão. Nesse cenário, o *Detecta* foi aos poucos caindo em desuso, sendo substituído, em 2024, pelo programa Muralha Paulista (PAGNAN, 2024).

As expectativas sobre os ganhos da automação algorítmica em dispositivos de vigilância são, em grande medida, reproduzidas nas análises acerca de ferramentas de policiamento preditivo. Segundo, Chris Sims, ex-comissário de polícia de West Midlands, no Reino Unido, a substituição do tirocínio policial por modelos matemáticos capazes de estimar a incidência futura de crimes e riscos de vitimização é a “a transformação mais radical que já aconteceu na polícia” (citado em DODD, 2014).

Até algumas décadas atrás, dificuldades técnicas se mostravam obstáculos formidáveis ao uso de mapas e abordagens estatísticas na análise criminal e na escolha de alocação de patrulhas (HAGGERTY, 2001). Em muitos países, dados criminais não eram coletados de forma sistemática e padronizada, o que dificultava análise com séries temporais mais amplas e com atenção às dinâmicas de microrregiões. Policiais trabalhavam, então, com bases incompletas e não dispunham de muitas formas para compensar o problema de subnotificação de crimes. Além disso, a produção de mapas requeria investimento de tempo, mão de obra e espaço (geralmente, paredes inteiras) para a visualização adequada da incidência criminal. No Brasil, até poucos anos, mapas de crimes eram representações em papel de áreas de atuação dos batalhões de polícia militar ou delegacias de polícia civil, onde crimes eram geolocalizados com alfinetes coloridos. Na medida em que mais crimes eram cometidos, a visualização de sua distribuição geográfica tornava-se confusa, dificultando a distinção de padrões e séries de eventos relacionados. A

análise de tendências também não era tarefa simples, já que mapas precisavam ser atualizados de tempos em tempos e informações sobre diferentes períodos só estavam disponíveis em fotos de arquivo (DE LIMA, 2005).

De fato, os sistemas modernos de análise criminal alteram radicalmente o cenário descrito acima. Atualmente, não apenas os registros de ocorrência são, em sua maioria, eletrônicos, o que alimenta automaticamente as bases de dados, mas as ferramentas de visualização são também mais adaptáveis às demandas da polícia, permitindo a seleção e análise de determinados territórios ou dinâmicas criminais. Indo além, os sistemas disponíveis no mercado são capazes de cruzar enormes volumes de dados, incluindo variáveis como: densidade populacional, dados censitários, localização de bares, escolas, parques, terrenos baldios, igrejas, áreas comerciais, pontos de ônibus, estações de metrô, tabelas de campeonatos esportivos, agenda de eventos empresariais e culturais, transações bancárias, uso de cartão de crédito, relatórios hospitalares, reservas de hotel, voos, pesquisas de internet, e-mails, redes sociais, ligações telefônicas e até fases da lua (ANDREJEVIC, 2017).

Se antes o policial identificava áreas de concentração de crimes e alocava as patrulhas de acordo, atualmente, a disponibilidade de dados mais granulares e avanços nas ferramentas de modelagem algorítmica permitem a identificação de novos padrões criminais, a ponto de muitos especialistas em segurança pública garantirem que suas dinâmicas futuras também podem ser decifradas. A promessa é que o policiamento ostensivo pode se antecipar a crimes específicos e em áreas específicas (e.g., roubos de celulares em determinada esquina e em determinada hora do dia), de modo que a simples presença policial serviria para dissuadir a ação delituosa. Assim, a polícia atuaria antes do crime, não com o objetivo de reprimir o criminoso, mas tornando o próprio cometimento do crime cada vez mais difícil.

À primeira vista, a implementação desses sistemas na rotina policial parece proporcionar avanços essenciais para a segurança pública. Contudo, o que tem sido verificado é um pouco diferente. Em primei-

ro lugar, não há consenso sobre os impactos positivos de sistemas de policiamento preditivo. Enquanto alguns desenvolvedores defendem que seus produtos são responsáveis pela queda de índices criminais (MOHLER, 2015), grande parte dos analistas aponta que, para instituições policiais que já fazem mapeamento de crimes, o policiamento preditivo não é “uma revolução que vai mudar tudo... mas um ganho incremental” (HOLLYWOOD, 2012). Além disso, pesquisas sobre a implementação de novos sistemas de análise criminal junto às forças de segurança indicam que, entre a prancheta do desenvolvedor e o uso da tecnologia na rotina operacional, há uma série de processos que podem impedir que o objetivo final seja cumprido. Em geral, policiais contestam, resistem, desvirtuam e adaptam os múltiplos sistemas que buscam automatizar seu trabalho e designar as áreas prioritárias de patrulhamento (MANNING, 2011). Por fim, críticos apontam que esses sistemas sofrem com vieses nas bases de dados usadas para seu treinamento. Quando os crimes registrados no passado são a chave para entender sua incidência futura, o que os sistemas fazem é reproduzir o padrão anterior de policiamento. Desse modo, se a polícia costuma reprimir com mais ímpeto crimes em determinada área ou costuma realizar abordagens e prisões de determinado grupo populacional, o que o sistema de análise vai aprender é que é preciso aumentar ainda mais o controle sobre esses mesmos territórios e indivíduos.

Defensores de aplicações de *big data* no campo da segurança pública reconhecem essas limitações, mas apontam que, mesmo como as distorções nas bases de dados, a automatização algorítmica seria ainda melhor do que as alternativas existentes. Como apontam Mayer-Schönberger & Cukier (2013, p. 112):

com o uso do *big data*, esperamos identificar pessoas específicas em vez de grupos, o que nos liberta do problema da “culpa por associação”, no caso do “perfilamento”. Num mundo de *big data*, alguém com um nome árabe, que paga em dinheiro por uma passagem só de ida de primeira classe, talvez não esteja mais sujeito a uma minuciosa investigação no ae-

roporto se outros dados específicos determinarem improvável se tratar de um terrorista. Com o *big data*, podemos escapar da camisa de força das identidades grupais e substituí-las por previsões mais granuladas para cada pessoa. A promessa do *big data* é continuar com a mesma prática – “perfilando” –, mas aperfeiçoada, de forma menos discriminatória e mais individualizada.

As vantagens operacionais advindas de novas aplicações de *big data* não podem ser ignoradas, mas o que se viu com os exemplos acima é que precisamos entender os problemas que elas podem acarretar, especialmente no que tange à supervisão da ação policial e ao seu impacto em populações vulnerabilizadas. Na medida em que aumentam as capacidades das forças de segurança de vigiar e se antecipar a ações criminosas, aumentam também os riscos de usos indevidos das novas tecnologias. E é sobre esses desafios que a próxima seção vai se debruçar.

3. Dilemas éticos e desafios regulatórios no uso de *big data* no campo da segurança pública

Existe alguma forma justa de se atuar sobre um crime que não aconteceu? Quais são os dilemas éticos de realizar perfilamento de risco de reincidência criminal para manter determinada pessoa presa em regime fechado? Os ganhos na capacidade de vigilância e controle a partir de sistemas de reconhecimento facial e identificação de padrões de comportamentos suspeitos são proporcionais ao desafio da violência urbana? Como podemos garantir que essas ferramentas são usadas dentro de limites aceitáveis, evitando, por exemplo, o monitoramento de manifestações pacíficas, a vigilância de adversários políticos ou mesmo o uso dos sistemas de coleta de dados para fins privados (e.g., policiais que usam as câmeras para flagrar traições matrimoniais)? Quando pensamos nos ganhos operacionais das aplicações de *big data*, não podemos negligenciar que as novas ferramentas disponibilizadas para as forças de segurança expandem enormemente sua capacidade de controle, poden-

do interferir nas regras do jogo democrático e nas nossas percepções de liberdade e privacidade.

Não por acaso, muitas das aplicações descritas acima têm sido acompanhadas de críticas relevantes da parte de pesquisadores e membros da sociedade civil. Por um lado, alguns especialistas apontam que muitas das promessas que cercam as tecnologias não se cumprem no mundo real. A adoção de novos dispositivos tecnológicos é sempre mediada pelas diferentes percepções sobre seus efeitos, o que leva policiais a resistir ou se adaptar às novas ferramentas. Mesmo inovações aparentemente simples, como a troca de plataformas de despacho de viaturas, se provam, com frequência, problemáticas, já que policiais tendem a rejeitar mudanças que aumentam a supervisão sobre seu trabalho ou demandem o desenvolvimento de novas habilidades profissionais (MANNING, 2011). Por outro lado, organizações da sociedade civil têm apontado que o processo de automação algorítmica, se não for implementado a partir de mecanismos participativos, transparentes e auditáveis, tende a se tornar simplesmente uma “atualização *high-tech* para o velho e conhecido racismo que está na base do sistema de justiça criminal” (NUNES, 2019). Diversas pesquisas corroboram essa crítica, apontando que sistemas de monitoramento biométrico e de análise da distribuição espacial do crime, por exemplo, carregam vieses que causam erros de identificação de elementos suspeitos e distorções no mapeamento de áreas de risco (FERGUSON, 2017).

Mais especificamente, críticos têm levantado graves alertas para os riscos de aprofundamento dos padrões discriminatórios da ação policial. Sistemas de policiamento preditivo, por exemplo, automatizam formas de classificar, mensurar e visualizar o fenômeno criminal. Estes, portanto, codificam e reproduzem escolhas de patrulhamento, como áreas de baseamento, perfis de suspeitos a serem abordados e crimes prioritários que devem ser registrados e reprimidos. Como lembra Jefferson (2018, p. 2), “estatísticas criminais moldam os mapas de crimes futuros e..., de forma recíproca, os mapas legitimam as estatísticas”, já que os policiais tendem a realizar as prisões em flagrante nas

áreas para as quais foram designados pelos sistemas preditivos. Ou seja, o uso de tecnologias preditivas empresta uma “autoridade algorítmica” à ação policial (GILLESPIE, 2014), negligenciando que as representações do crime carregam vieses de origem, como repressão desproporcional de populações marginalizadas, e que táticas de patrulhamento proativo guardam estreita relação com formas punitivas de governo (ZEDNER, 2007). Contudo, os problemas inerentes às análises criminais preditivas não precisam implicar necessariamente no seu abandono. Se elas trazem conhecimento sobre riscos de criminalização e vitimização, elas podem informar políticas públicas preventivas que não passem pela repressão policial antecipada. Como apontam Edler e Lobato (2021, p. 91):

softwares preditivos poderiam apoiar a retomada de um debate mais amplo sobre as causas sociais do crime, ajudando a guiar investimentos em equipamentos públicos, assistência social, melhorias em serviços básicos e inserção de jovens no mercado de trabalho. Ao invés de contribuir para a criminalização de populações já marginalizadas, [a previsão de crimes] poderia então servir como um diagnóstico de onde o Estado tem sido incapaz de fomentar o desenvolvimento. Nesse sentido, áreas onde há maior risco de crimes não seriam os alvos de mais patrulhas, mas sim de mais atenção aos fatores subjacentes do crime, suas causas menos imediatas e, por isso, mais difíceis de solucionar.

No caso dos sistemas de monitoramento biométrico e alertas automatizados de comportamentos suspeitos, o desafio de reposicionar as soluções tecnológicas para fins não punitivos é mais difícil de contornar. Os investimentos recentes em videomonitoramento nas grandes cidades têm criado redes de vigilância que capturam informações privadas de todos os cidadãos. Embora, as forças de segurança aleguem que monitoram apenas indivíduos com pendências na justiça e suspeitos de envolvimento com crimes, nossa capacidade de supervisionar a atividade policial é extremamente limitada e as instituições atuam, muitas vezes, à revelia de decisões do Ministério Público e de determinações

legais (MENDONÇA, 2025). Além disso, pesquisas confirmam que o risco de vigilantismo não deve ser minimizado, tendo documentado casos de policiais que usam esses sistemas para fins voyeurísticos, para perseguir adversários políticos e, inclusive, para vender informações ao crime organizado (CARDOSO, 2015; EDLER & CEIA, 2023).

Enquanto os usos de ferramentas de monitoramento eram voltados ao serviço investigativo e ocorriam após autorizações judiciais, era possível limitá-los à obtenção de evidências e pistas que pudessem ajudar na resolução de crimes. Já o emprego da vigilância algorítmica no policiamento ostensivo e preventivo pode torná-la uma “ferramenta irresistível de opressão e uma máquina de destruição total de privacidade perfeitamente adequada para governos que exercem controle autoritário sem precedentes” (HARTZOG & SELINGER, 2018). De fato, a disseminação de câmeras de vigilância biométrica nas grandes cidades faz com que nossos hábitos, movimentações e interações possam ser documentados e catalogados sem qualquer esforço operacional. Basta fazer uma pesquisa em um banco de dados para ter acesso a todas as imagens em que determinado indivíduo aparece. Em questão de segundos, a polícia pode descobrir, por exemplo, todas as vezes em que uma pessoa foi a um bar, visitou amigos, chegou atrasada no trabalho, frequentou uma casa de prostituição, fumou na calçada, participou de reuniões dos alcóolicos anônimos ou se fantasiou de *Pikachu* no carnaval. Ou seja, a vigilância algorítmica carrega um potencial de controle ubíquo que pode transformar a natureza das interações sociais e constranger comportamentos individuais, sejam esses ilegais ou não.

4. Conclusão

“Raw data is an oxymoron” (GITELMAN, 2013)

Analisar os impactos políticos e sociais das múltiplas aplicações de *big data* para o campo de controle do crime é um importante antídoto ao otimismo exagerado que costuma cercar as inovações tecno-

lógicas na segurança pública. Assimilar de forma acrítica o discurso “tecno-solucionista” dos desenvolvedores de sistemas de vigilância biométrica ou policiamento preditivo (MOROZOV, 2013), por exemplo, pode levar não apenas a frustrações com seu real impacto na rotina policial, mas também pode esconder muitos de seus efeitos deletérios, como a adoção de padrões desiguais de policiamento em bairros pobres e um monitoramento mais intrusivo da população negra.

Lisa Gitelman (2013) nos lembra que não existem dados puros que refletem a realidade de determinado fenômeno social. Os dados não são recursos naturais, mas produções sociais, culturais e políticas que se moldam a partir de nossos valores e interesses. No campo da segurança pública e da justiça criminal, isso implica em uma série de desafios, como, por exemplo, como lidar com as cifras ocultas de crimes ou com a sobrerrepresentação de determinadas populações nas bases de suspeitos das polícias. Como vimos, o *big data* pode aumentar nossa compreensão sobre as dinâmicas criminais e melhorar a eficácia das forças de segurança, reduzindo a violência urbana. No entanto, o desenvolvimento de novas ferramentas deve vir acompanhado de regulação adequada, com mecanismos efetivos de supervisão e transparência.

Referências

ANDERSON, Chris. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. **Wired**, 23 de junho de 2008. Disponível em: <https://www.wired.com/2008/06/pb-theory>

ANDREJEVIC, Mark. To Pre-Empt A Thief. **International Journal of Communication**, v. 11, p. 879-896, 2017.

BECK, C., McCUE, C. Predictive Policing: What Can We Learn from Wal-Mart and Amazon about Fighting Crime in a Recession? **The Police Chief**, 26(11), p. 18-25, 2009.

BOWKER, G., STAR, S. **Sorting Things Out: Classification and Its Consequences**. Cambridge, MA: The MIT Press, 2000.

CARDOSO, B. **Todos os Olhos: Videovigilâncias, Voyeurismo e (re)produção Imagética**. Rio de Janeiro: Editora UFRJ, 2015.

DE LIMA, R. S. (2005) **Contando crimes e criminosos em São Paulo: Uma sociologia das estatísticas produzidas e utilizadas entre 1871 e 2000**. Tese de Doutorado, FFLCH/USP, 2005.

DODD, V. Police force spends £25m on switch to technology-led crime-fighting. **The Guardian**, 21 de julho de 2014. Disponível em: <https://www.theguardian.com/uk-news/2014/jul/21/west-midlands-police-technology-led-crime-fighting>

DODD, V. Met gangs matrix may be discriminatory, review finds. **The Guardian**, 21 de dezembro de 2018. Disponível em: <https://www.theguardian.com/uk-news/2018/dec/21/metropolitan-police-gangs-matrix-review-london-mayor-discriminatory>

DOMINGOS, P. **The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World**. New York, NY: Basic Books, 2015.

EDLER, D., Lobato, L. A política do policiamento preditivo: pressupostos criminológicos, técnicas algorítmicas e estratégias punitivas. **Revista Brasileira de Ciências Criminais**, 29(183), p. 57-98, 2021.

EDLER, D., CEIA, E. (eds.). **Tecnologia, Segurança e Direitos: Os usos e riscos de sistemas de reconhecimento facial no Brasil**. Rio de Janeiro: Fundação Konrad Adenauer, 2023.

EDLER, D., SIMÕES-GOMES, L., PERON, A. **Nem uma “revolução digital”, nem a distopia do controle total: Os efeitos das inovações tecnológicas na polícia de São Paulo**. Relatório de pesquisa, Instituto Igarapé e Universidade de Essex, 2023.

FERGUSON, A. **The Rise of Big Data Policing: Surveillance, Race and the Future of Law Enforcement**. New York, NY: NYU Press, 2017.

FRY, H. **Hello World: Being Human in the Age of Algorithms**. New York, NY: W. W. Norton & Company, 2019.

GILLESPIE, T. The Relevance of Algorithms. In: GILLESPIE, T., BOCZKOWSKI, P., FOOT, K. (eds.). **Media Technologies: Essays on Communication, Materiality, and Society**. Cambridge, MA: MIT Press, p. 167-194, 2014.

GITELMAN, L. (ed.). **“Raw data” is an oxymoron**. Cambridge, MA.: MIT Press, 2013.

HAGGERTY, K. **Making Crime Count**. Toronto: University of Toronto Press, 2001.

HARTZOG, W., SELINGER, E. Facial Recognition Is the Perfect Tool for Oppression. **Medium**, 02 de Agosto de 2018. Disponível em: <https://medium.com/s/story/facial-recognition-is-the-perfect-tool-for-oppression-bc2ao8fofe66>

HOLLYWOOD, J. **Predictive Policing: What It Is, What It Isn't, and Where It Can Be Useful**. Arlington, Va: RAND Corporation and National Institute of Justice, 2012.

JEFFERSON, B. **Predict and Surveil: Data, Discretion, and the Future of Policing**. New York, NY: Oxford University Press, 2020.

KITCHIN, R. Big Data, new epistemologies and paradigm shifts. **Big Data & Society**, 1(1), p. 1-12, 2014.

MANNING, P. **The Technology of Policing: Crime Mapping, Information Technology, and the Rationality of Crime Control**. Nova York, NYU Press, 2011.

MENDONÇA, J. Minoria dos promotores acha que fiscalizar a polícia é prioridade do Ministério Público. Agência Pública, 29 de abril de 2025. Disponível em: <https://apublica.org/2025/04/minoria-dos-promotores-acha-que-fiscalizar-a-policia-e-prioridade-do-ministerio-publico/>

MOHLER, G., Short, M., MALINOWSKI, S., JOHNSON, M., TITA, G., BERTOZZI, A., BRANTINGHAM, J. (2015) Randomized Controlled Field Trials of Predictive Policing. **Journal of American Statistical Association**, 110(512), p. 1399-1411, 2015.

MOROZOV, E. **To Save Everything, Click Here: Technology, Solutionism, and the Urge to Fix Problems that Don't Exist**. London: Allen Lane, 2013.

NUNES, P. Exclusivo: levantamento revela que 90,5% dos presos por monitoramento facial no Brasil são negros. **The Intercept**, 21 de novembro de 2019. Disponível em: <https://theintercept.com/2019/11/21/presos-monitoramento-facial-brasil-negros/>

PAGNAN, R. Tarcísio apaga marca tucana Detecta do sistema de busca das polícias de SP. *Folha de São Paulo*, 30 de julho de 2024. Disponível em: <https://www1.folha.uol.com.br/cotidiano/2024/07/tarcisio-apaga-marca-tucana-detecta-do-sistema-de-busca-das-policias-de-sp.shtml>

PAGNAN, R., BARBON, J. Alckmin vai relançar sistema que já custou R\$ 30 milhões e não funciona. *Folha de São Paulo*, 30 de junho de 2017. Disponível em: <https://www1.folha.uol.com.br/cotidiano/2017/06/1897306-alckmin-vai-relancar-sistema-que-ja-custou-r-30-milhoes-e-nao-funciona.shtml>

PERON, A., ALVAREZ, M. O Governo da Segurança: Modelos Securitários Transnacionais e Tecnologias de Vigilância na cidade de São Paulo. **Lua Nova**, no. 114, p. 175-212, 2021.

RYBERG, J., ROBERTS, J.V. (eds.) **Sentencing and Artificial Intelligence**. New York, NY: Oxford University Press, 2022.

SÃO PAULO. Sistema Detecta ganha 97 novas câmeras de monitoramento. **Governo do Estado de São Paulo**, 19 de outubro de 2016. Disponível em: https://www.youtube.com/watch?v=EDyAm64qDgc&ab_channel=GovernodoEstadoDeS%C3%A3oPaulo

STATEWATCH. UK: Ministry of Justice secretly developing 'murder prediction' system. **Statewatch**, 08 de abril de 2025. Disponível em: <https://www.statewatch.org/>

[news/2025/april/uk-ministry-of-justice-secretly-developing-murder-prediction-system/](https://www.bbc.com/news/2025/april/uk-ministry-of-justice-secretly-developing-murder-prediction-system/)

TCE. **Relatório de Fiscalização de Natureza Operacional Solução de Consciência Situacional – DAS Detecta**. Tribunal de Contas do Estado de São Paulo, 2017. Disponível em: <https://www.tce.sp.gov.br/sites/default/files/portal/detecta.pdf>

WEISBURD, D. The law of crime concentration and the criminology of place. **Criminology**, 53(2), p. 133-157, 2015.

ZAVRŠNIK, A. (ed.). **Big Data, Crime and Social Control**. Abingdon and New York: Routledge, 2018.

ZEDNER, L. Pre-crime and post-criminology? **Theoretical Criminology**, 11(2), p. 261-281, 2007.

Daniel Edler Duarte é pesquisador associado do Instituto de Ciências Sociais da Universidade do Estado do Rio de Janeiro (ICS-UERJ) e do Núcleo de Estudos da Violência da Universidade de São Paulo (NEV/USP), além de membro da Comissão de Segurança Pública da OAB/SP. Daniel já trabalhou em diversas instituições, incluindo: Universidade de Glasgow, Universidade de Southampton, PUC-Rio e na Fundação Getulio Vargas. Sua pesquisa atual se desdobra em três eixos principais: (1) novas tecnologias e práticas de vigilância; (2) policiamento de protestos; e (3) controvérsias públicas no campo da ciência e tecnologia. E-mail: danieledler@usp.br

Big Data, democracia participativa e o papel da Inteligência Artificial

Filipe Medon

Resumo

O artigo examina como *Big Data* e Inteligência Artificial reconfiguram a deliberação e a participação democráticas no ambiente digital na atualidade. Para tanto, parte de um mapeamento de riscos múltiplos, como a desinformação e a filtragem algorítmica, a qual intensifica bolhas e assimetrias informacionais, afetando a igualdade deliberativa. Na sequência, contextualiza o debate no marco regulatório brasileiro, à luz da responsabilização de plataformas após a decisão do STF sobre a constitucionalidade do artigo 19 do Marco Civil da Internet, e discute inovações democráticas, além de avaliar o potencial positivo do emprego de IA em tarefas como o agrupamento de contribuições e mensuração da qualidade do debate, reconhecendo limites de viés, custo e evidência empírica. Propõe, por fim, uma calibragem entre tecnologia, educação midiática e desenho institucional, com deveres proporcionais às plataformas, a fim de despoluir o espaço informacional e viabilizar uma participação plural e transparente.

Abstract

The article examines how Big Data and Artificial Intelligence are reshaping democratic deliberation and participation in today's digital environment. To this end, it starts from a mapping of multiple risks, such

as misinformation and algorithmic filtering, which intensifies information bubbles and asymmetries, affecting deliberative equality. It then contextualizes the debate within the Brazilian regulatory framework, in light of platform accountability following the Supreme Court's decision on the constitutionality of Article 19 of the Brazilian Civil Rights Framework for the Internet, and discusses democratic innovations, in addition to assessing the positive potential of using AI in tasks such as grouping contributions and measuring the quality of debate, recognizing limits of bias, cost, and empirical evidence. Finally, it proposes a calibration between technology, media education, and institutional design, with proportional duties for platforms, in order to clean up the information space and enable plural and transparent participation.

1. Notas introdutórias: da “legião de imbecis” de Umberto Eco aos *bots* que poluem o ambiente digital

A Internet deu voz a quem antes era deixado alheio ao debate público. Hoje, com apenas um *smartphone* e uma conta em uma rede social, uma pessoa, de qualquer lugar do planeta, pode se manifestar sobre qualquer tema, ressaltados os casos de países em que a Internet não é um espaço democrático e ainda é marcada por limitações baseadas em censura. Dito diversamente: em contextos plurais e em ambientes democráticos, a rápida e acelerada inclusão digital tem permitido que a cidadania seja, cada vez mais, plenamente exercitada na ágora digital da Internet.

Nada obstante, a história recente tem fartamente demonstrado que quantidade não significa necessariamente qualidade. Afinal, como ressaltado por Umberto Eco em 2015, as redes sociais também teriam dado o direito à palavra a uma “legião de imbecis”, que, segundo o filósofo italiano, antes falavam somente “em um bar e depois de uma taça de vinho, sem prejudicar a coletividade.”¹ Agora, no entanto, os “imbecis” teriam

1 REDES sociais deram voz a legião de imbecis, diz Umberto Eco. **UOL Notícias** (Agência ANSA), 11 jun. 2015. Disponível em: <https://noticias.uol.com.br/ulti->

“o mesmo direito à palavra de um Prêmio Nobel”². E assim concluiu o escritor em sua célebre e popular afirmação: “[o] drama da internet é que ela promoveu o idiota da aldeia a portador da verdade”³.

A ácida conclusão do filósofo italiano serve apenas para ressaltar que, apesar de haver cada vez mais informação na Internet, nem sempre ela será útil ou relevante para a construção de uma deliberação coletiva verdadeiramente democrática e voltada à promoção do bem comum. Isso porque, para além dos comentários de ódio dos “imbecis” mencionados por Eco, a Internet hoje está repleta dos chamados *bots*, isto é, robôs normalmente comandados por Inteligência Artificial, e que se valem das mais diversas estratégias para poluir o ambiente digital ou, pelo menos, criar distorções comunicativas graves.

Exemplo disso está na atuação dos “robôs sociais”, que atuam dentro das chamadas redes automatizadas ou *botnets*, e que poderiam ser descritas como “contas controladas por software que geram artificialmente conteúdo e estabelecem interações com não robôs. Eles buscam imitar o comportamento humano e se passar como tal de maneira a interferir em debates espontâneos e criar discussões forjadas.”⁴

Esses robôs podem, por exemplo, criar cortinas de fumaça, simplesmente massificando “o debate sobre um tema irrelevante, afastando o foco de um assunto delicado ou até mesmo de uma notícia contrária a

mas-noticias/ansa/2015/06/11/redes-sociais-deram-voz-a-legiao-de-imbecis-diz-umberto-eco.htm. Acesso em: 29 jun. 2025.

- 2 REDES sociais deram voz a legião de imbecis, diz Umberto Eco. **UOL Notícias** (Agência ANSA), 11 jun. 2015. Disponível em: <https://noticias.uol.com.br/ultimas-noticias/ansa/2015/06/11/redes-sociais-deram-voz-a-legiao-de-imbecis-diz-umberto-eco.htm>. Acesso em: 29 jun. 2025.
- 3 REDES sociais deram voz a legião de imbecis, diz Umberto Eco. **UOL Notícias** (Agência ANSA), 11 jun. 2015. Disponível em: <https://noticias.uol.com.br/ultimas-noticias/ansa/2015/06/11/redes-sociais-deram-voz-a-legiao-de-imbecis-diz-umberto-eco.htm>. Acesso em: 29 jun. 2025.
- 4 FERRARI, Pollyana; FILHO, Alberto Freitas. O Mundo está ao Contrário e Ninguém Reparou. In: TOURAL, Carlos; CORONEL, Gabriela; FERRARI, Pollyana (orgs.) **Big Data e Fake News na sociedade do (des)conhecimento**, 2. ed. Aveiro: Ria Editorial, 2020, p. 170.

determinado político ou pessoa que estaria, se não houvesse a interferência artificial, na pauta do dia.”⁵⁻⁶ Apenas para se ter uma dimensão desse fenômeno, segundo dados de 2024 da 5th Column AI, após analisar 1.269 milhões de contas na plataforma X, antigo Twitter, estimou-se que aproximadamente 64% dessas contas seria potencialmente de *bots*.⁷ E isso em apenas uma única plataforma. Hoje, nem mesmo as preferências do cenário musical são imunes às interferências artificiais de contas inautênticas, como se observa na plataforma Spotify, que vem alertando para o fenômeno conhecido como “artificial streaming”,⁸ o qual pode tanto promover quanto ocultar artistas, alterando a percepção do público sobre o que, de fato, está fazendo sucesso na indústria da música.

-
- 5 ROBL FILHO, Ilton Norberto; MARRAFON, Marco Aurélio; MEDON, Filipe. O impacto das técnicas de inteligência artificial aplicadas no *deepfake* e nas redes automatizadas abalam o mercado livre de ideias e a democracia constitucional e deliberativa. In: **Economic Analysis of Law Review**, v. 13, n. 3, pp. 32-47, out./dez. 2022. Disponível em: <https://portalrevistas.ucb.br/index.php/EALR/article/view/12527>, p. 36.
 - 6 “Something fundamental was happening to threaten democracy, and our collective eye fell on the novel and rapidly changing – technology. Technological processes beyond the control of any person or country – the convergence of social media, algorithmic news curation, bots, artificial intelligence, and big data analysis— were creating echo chambers that reinforced our biases, were removing indicia of trustworthiness, and were generally overwhelming our capacity to make sense of the world, and with it our capacity to govern ourselves as reasonable democracies” (BENKLER, Yochai; FARIS, Robert; ROBERTS, Hal. **Network propaganda: manipulation, disinformation, and radicalization in American politics**. New York: Oxford University Press, 2018, p. 4).
 - 7 ROBINSON, David; JABRAYILOV, Rafiq; MICHAEL K. Elon was right about bots on X.com. **Internet 2.0 – 5th Column**, 25 jan. 2024. Disponível em: <https://internet2-o.com/bots-on-x-com/>. Acesso em: 29 jun. 2025
 - 8 “An artificial stream is a stream that doesn’t reflect genuine user listening intent, including any instance of attempting to manipulate streaming services like Spotify by using automated processes (like bots or scripts). If left undetected, artificial streams dilute the royalty pool, shifting revenue from legitimate artists to bad actors. They also undermine the fair playing field that streaming represents.” (SPOTIFY FOR ARTISTS. **All you need to know about artificial streaming**, 2024. Disponível em: <https://artists.spotify.com/artificial-streaming>. Acesso em: 29 jun. 2025).

Tamanha é a preocupação com essas contas inautênticas, que o Supremo Tribunal Federal brasileiro, em paradigmática e histórica decisão, ao analisar os Recursos Extraordinários (RE) de números 1037396 (Tema 987) e 1057258 (Tema 533), que versavam sobre a constitucionalidade do artigo 19 do Marco Civil da Internet, acabou recrudescendo o regime jurídico aplicável a esse tipo de expediente artificial. Com efeito, as plataformas que antes só podiam ser responsabilizadas civilmente por ilícitos cometidos por usuários após o descumprimento de ordem judicial específica, agora podem vir a ser responsabilizadas mesmo sem ordem judicial ou notificação privada “quando for detectado o uso de redes artificiais de distribuição ilícitas usando robôs. Nesses casos, há uma presunção de que a plataforma tinha conhecimento da ilicitude e ela somente poderá afastar sua responsabilidade se provar que agiu em tempo razoável e com diligência para remover o conteúdo.”⁹

E o que está por trás disso tudo? A Inteligência Artificial. Entretanto, é preciso indagar: será que a mesma ferramenta tecnológica que tem proporcionado essas inegáveis e indesejáveis distorções para a democracia pode, por outro lado, contribuir para aperfeiçoá-la, especialmente por meio da ampliação dos mecanismos de participação individual e coletiva? Em última análise, como indagou Jamie Susskind, pode a democracia deliberativa “sobreviver em um sistema onde a própria deliberação não é mais um privilégio dos seres humanos?”¹⁰ Ou, conforme

9 SUPREMO TRIBUNAL FEDERAL. **Informação à sociedade:** RE 1.037.396 (Tema 987) e 1.057.258 (Tema 533) – Responsabilidade de plataformas digitais por conteúdo de terceiros. Brasília, DF, 2025. Disponível em: https://www.stf.jus.br/arquivo/cms/noticiaNoticiaStf/anexo/Informac807a7710a768SociedadeArt19M-CI_vRev.pdf. Acesso em: 29 jun. 2025.

10 “Can Deliberative Democracy survive in a system where deliberation itself is no longer the preserve of human beings? It’s possible that human voices could be crowded out of the public sphere altogether by bots that care little for our conversational norms. In the future, (...) they could look and sound like humans, endowed with faces and voices and extraordinary rhetorical gifts. How can we, with our feeble brains and limited knowledge, participate meaningfully in deliberations if our views are instantaneously ripped to shreds by armies of bots armed with a million smart-ass retorts? Advocates of bots might put it differently: why

reflexão de Hélène Landemore, “poderia a Inteligência Artificial ajudar a trazer deliberação de qualidade para as massas?”¹¹

É precisamente disso que o presente estudo pretende se ocupar, traçando, inicialmente, linhas estruturais sobre a relação interdependente e vital entre os conceitos de *Big Data* e *Big Data Analytics*.

2. *Big Data* e *Big Data Analytics*: motor e combustível

Em um mundo cada vez mais conectado, a produção de dados tem se tornado massiva. E é precisamente nesse contexto que se alude à noção de *Big Data*, que pode ser compreendida como “um termo em evolução que descreve qualquer quantidade volumosa de dados estruturados, semiestruturados, ou não estruturados que têm o potencial de ser explorados para obter informações.”¹² Apesar da falta de consenso,¹³

spend time deliberating when increasingly sophisticated bots can debate the issues faster and more effectively on our behalf?” (SUSSKIND, Jamie. **Future Politics**: living together in a world transformed by tech. Oxford: Oxford University Press, 2018, p. 233).

11 LANDEMORE, Hélène. Can AI bring deliberative democracy to the masses? *NYU Law Review*, [S.l.], v. 95, pp. 1–20, 2020. Disponível em: <https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>, p. 01).

12 MAGRANI, Eduardo. **A Internet das Coisas**. Rio de Janeiro: FGV Editora, 2018, pp. 21–22.

13 “(...) o *Big Data* é comumente associado a 3 (três) ‘Vs’: volume, velocidade e variedade. Volume e variedade, porque ele excede a capacidade das tecnologias ‘tradicionais’ de processamento, conseguindo organizar quantidades antes inimagináveis dos *bits* aos *yottabytes* – e em diversos formatos – e.g., textos, fotos etc. – e, tudo, isso, em alta velocidade. Tal evolução poderia ser imputada a uma diferença crucial entre o *Big Data* e as outras metodologias comuns de processamento de dados (...), que é o fato da prescindibilidade de os dados estarem previamente estruturados para o seu tratamento. (...) Por isso, os dados passaram a ser analisados não mais em pequenas quantidades ou por amostras, mas em toda a sua extensão.” (BIONI, Bruno Ricardo. **Proteção de dados pessoais**: a função e os limites do consentimento. Rio de Janeiro: Forense, 2019, p. 39–41).

diz-se que o *Big Data* seria explicado e caracterizado por “6vs”: volume, velocidade, variedade, veracidade, valor e visualização.¹⁴

Some-se, então, esse volume crescente de dados variados e de alto valor, com a alta velocidade de produção, análise e visualização¹⁵, associada a mecanismos de Inteligência Artificial que trabalham no processamento¹⁶, e pode-se chegar à importância do *Big Data Analytics*, o qual pode ser compreendido como “a possibilidade de extrair, a partir dos dados, correlações, padrões e associações que possam ser consideradas informações. Para tal objetivo, é grande a importância dos algoritmos e das máquinas responsáveis por tal processamento.”¹⁷ Daí a afirmação de que se os sistemas de Inteligência Artificial representam o motor dessa transformação social, por certo os dados são o combustível e a matéria-prima.¹⁸

14 DING, Guoru; WU, Qihui; WANG, Jinlong; YAO, Yu-Dong. Big Spectrum Data: The New Resource for Cognitive Wireless Networking. In: **ArXiv, Cornell University**, 23 abr. 2014. Disponível em: <https://arxiv.org/abs/1404.6508>. Acesso em: 29 jun. 2025.

15 MAGRANI, Eduardo. **A Internet das Coisas**. Rio de Janeiro: FGV Editora, 2018, p. 22.

16 Yuval Noah Harari identifica quatro métodos básicos de melhora da eficiência dessa análise de dados: (i) o aumento do número de processadores; (ii) o aumento da variedade de processadores; (ii) o aumento do número de conexões entre processadores; e (iv) o aumento da liberdade de movimento ao longo das conexões existentes. (HARARI, Yuval Noah. **Homo Deus: uma breve história do amanhã**. (Trad. Paulo Geiger). São Paulo: Companhia das Letras, 2016, pp. 195-196).

17 FRAZÃO, Ana. Plataformas digitais, big data e riscos para os direitos da personalidade. In: MENEZES, Joyceane Bezerra de; TEPEDINO, Gustavo (Coord.), **Autonomia Privada, liberdade existencial e direitos fundamentais**. Belo Horizonte: Fórum, 2019, p. 337.

18 “There is a saying in Silicon Valley that “Big Data is the new oil.” What do people mean by this? Big Data is crucial to the use and development of algorithms and artificial intelligence (“AI”). Algorithms and AI are the machines; Big Data is the fuel that makes the machines run. Just as oil made machines and factories run in the Industrial Age, Big Data makes the relevant machines run in the Algorithmic Society.” (BALKIN, Jack M. Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation. Yale Law School: Faculty Scholarship Series, n. 5160, 2018, p. 1154. Disponível em: [163](https://digitalcom-</p>
</div>
<div data-bbox=)

Em razão do *Big Data*, passa-se a ter “um salto quanto ao volume de dados processados, tornando-se possível correlacionar uma série de fatos (dados), estabelecendo-se entre eles relações para desvendar *padrões* e, por conseguinte, inferir, inclusive, *probabilidades* de acontecimentos futuros.”¹⁹ E isso pode ser utilizado para as mais diversas finalidades: desde a realização de cálculos atuariais, passando pela disseminação orquestrada de desinformação por razões político-eleitorais²⁰, até, enfim,

mons.law.yale.edu/cgi/viewcontent.cgi?article=6159&context=fss_papers. Acesso em 02 mar. 2019).

- 19 BIONI, Bruno Ricardo. **Proteção de dados pessoais: a função e os limites do consentimento**. Rio de Janeiro: Forense, 2019, p. 41.
- 20 “É possível ter uma abordagem ainda mais tangível dos problemas acima descritos – envolvendo análises de big-data e dados pessoais de usuários – quando os relacionamos com política e eleições populares. As análises de dados já são utilizadas como ferramentas eleitorais há pelo menos seis anos. Bimber (2014) aponta que, já nas eleições presidenciais de 2012, nos Estados Unidos, a campanha eleitoral de Barack Obama – o então presidente norte-americano – “introduziu uma onda de inovação técnica” (BIMBER, 2014, p. 141), adotando, entre outros procedimentos, análises de dados em larga escala. Uma das principais vantagens competitivas que uma organização adquire ao utilizar análises de big-data, seja para fins políticos, seja para fins comerciais, é a possibilidade de personalização e difusão de mensagens em termos de indivíduo ou de grupo específico. A partir do momento em que um usuário tem milhares de dados a seu respeito sendo coletados, organizados, relacionados e analisados continuamente, cria-se, então, a oportunidade de lhe transmitir mensagens que sejam mais precisas (BIMBER, 2014) e que abordem temas os quais ele já tenha predisposições a reagir e a interagir. Dessa maneira, organizações como a da campanha de Barack Obama poderiam direcionar mensagens políticas e ideológicas aos usuários de redes sociais, por exemplo, na tentativa de influenciar mudanças de posição e tomadas de decisão – mensagens estas que são, ao mesmo tempo, precisas e inúmeras. Um exemplo ainda mais recente de como o big-data pode ter influência direta sobre a política e, mais especificamente, sobre pleitos democráticos, é o caso da empresa Cambridge Analytica, que esteve por trás de campanhas como a do Brexit, no Reino Unido, e do atual presidente dos Estados Unidos, Donald Trump. Tratava-se de uma organização privada, de origem inglesa, que esteve sob os holofotes nos últimos meses às custas de um escândalo envolvendo a corporação Facebook e dados pessoais de milhões de usuários. Segundo matéria publicada pelo jornal El País (GUIMÓN, 2018), a Cambridge Analytica teria utilizado irregularmente os dados de cerca de oitenta milhões de usuários da rede social Facebook para fins eleitorais. A empresa realizava análises de dados provenientes de big-data tendo

para a construção de mecanismos colaborativos de participação e deliberação democrática. Com efeito, dessas utilizações podem advir tanto benefícios, quanto problemas.

Especificamente em relação à participação democrática, David M. Douglas e Dave Kinkead observam que o *Big Data Analytics* e a Internet acabaram transformando a própria noção de comunicação política, uma vez que passaram a permitir a disseminação em larga escala de mensagens com elevadíssimo grau de personalização, sem que, necessariamente, estas tenham sido devida e suficientemente expostas ao amplo debate público.²¹

Nas palavras dos autores:

The Internet and big data analytics have changed the nature of political communication by decoupling message reach from message targeting. It allows individually tailored content to be distributed on a global scale without fear that the content of those messages will be overheard by others who may contest or reject it. Politicians are freed from the constraints imposed on what they say by the public presentation and discussion of

por objetivo obter informações proveitosas para a disputa política, e então as vendia assegurando a seus clientes maior possibilidade de vitória por meio de tais análises. Uma vez feita a coleta e a análise, os resultados eram direcionados aos eleitores em formato de ações via redes sociais, visando obter o voto daqueles que se mostravam mais sujeitos à mudança de opinião – chegava-se ao resultado de quais eram os usuários mais suscetíveis a mudanças fazendo o uso das análises de big-data. Devido ao uso não autorizado de dados pessoais dos usuários, tanto a empresa Cambridge Analytica, quanto a corporação Facebook, responderam à justiça norte-americana por possível influência irregular nas eleições presidenciais de 2016 nos Estados Unidos.” (CALDAS, Camilo Onoda Luiz; CALDAS, Pedro Neris Luiz. Estado, democracia e tecnologia: conflitos políticos e vulnerabilidade no contexto do big-data, das fake news e das shitstorms. Perspectivas em Ciência da Informação, Belo Horizonte, v. 24, n. 2, abr./jun. 2019. Disponível em: <http://dx.doi.org/10.1590/1981-5344/3604>, p. 201).

- 21 KINKEAD, Dave; DOUGLAS, David M. The Network and the Demos: Big Data and the Epistemic Justifications of Democracy. In: MACNISH, Kevin; GALLIOTT, Jai (orgs.). **Big Data and Democracy**. Edimburgo: Edinburgh University Press, 202, DOI: 10.3366/edinburgh/9781474463522.003.0009, p. 120.

arguments and claims. It frees political actors to engage in sophistry and demagoguery.²²

Acresça-se a isso, a ideia de que a existência de uma sobrecarga de informações ou *information overload*, causada pelo excesso de dados, pode acabar prejudicando a capacidade de tomar decisões adequadas, o que, em última instância, afeta a deliberação digital em contextos democráticos, já que se limita a transparência e se dificulta a troca significativa de ideias entre cidadãos. Resultaria daí, assim, o comprometimento da qualidade da participação democrática no ambiente virtual.²³ Isso

22 KINKEAD, Dave; DOUGLAS, David M. The Network and the Demos: Big Data and the Epistemic Justifications of Democracy. In: MACNISH, Kevin; GALLIOTT, Jai (orgs.). **Big Data and Democracy**. Edimburgo: Edinburgh University Press, 202, DOI: 10.3366/edinburgh/9781474463522.003.0009, p. 120.

23 “Among these limitations is its reduced capacity to evaluate and process large amounts of information. Information overload, or simply “receiving too much information”, produces cognitive biases as the human brain will routinely resort to heuristics in order to make decisions, which increase the likelihood of systematic errors in judgement [84]. The concept – sometimes discussed as cognitive overload [87], knowledge overload [39], or communication overload [7] – has been applied to a variety of contexts, but predominantly in the field of economics and management [23, 72]. More recently, the concept has been used to describe the challenges of online deliberation in democratic contexts [69] and digital citizen participation in law- and policy making [6, 18, 38, 51]. Information overload is a key challenge for digital mass participation, which jeopardises its intended democratic, epistemic, and economic values. For example, Information overload limits horizontal transparency as participants are not able “see and understand each other’s” [1]. This lowers the capacity for meaningful exchange of ideas and learning and leads to a decline in the overall decline quality of mass participation. Information overload also limits digital mass participation’s vertical transparency. Users of digital democracy platforms have found difficulties with access to and comprehension of online governmental information and in following debates and processes, thereby limiting the exchange of ideas, coordination of political action, and collaboration [24, 74, 82].” (ARANA-CATANIA, Miguel; VAN LIER, Felix-Anselm; PROCTER, Rob; TKACHENKO, Nataliya; HE, Yulan; ZUBIAGA, Arkaitz; LIAKATA, Maria. Citizen Participation and Machine Learning for a Better Democracy. **Digital Government: Research and Practice**, v. 2, n. 3, p. 27:1–27:22, 2021. DOI: 10.1145/3452118. Disponível em: <https://dl.acm.org/doi/10.1145/3452118>).

porque a democracia deliberativa exigiria “não só a igualdade de ‘possibilidades’ para participar da vida pública, mas também igualdade de ‘capacidades’ para fazê-lo efetivamente”.²⁴

E, nesse contexto, “a utilização das novas tecnologias, de redes sociais, dos aplicativos e da internet altera substancialmente a deliberação pública e a pauta republicana nas democracias contemporâneas.”²⁵ Há, por exemplo, “o fenômeno das “bolhas” nas interações sociais pelas redes sociais e atualmente por meio de aplicativos de mensagens, reduzindo substancialmente a pluralidade e a diversidade nas interações comunicativas.”²⁶

Como já se teve a oportunidade de destacar em outra sede:

As novas tecnologias permitem que o indivíduo escolha os fóruns que deseja engajar-se e as pessoas com as quais busca dialogar, fazendo com que os debates e as comunicações se encontrem restritas às áreas de interesse individual. A lógica adotada passa a ser da comunicação-consumo, deixando em segundo plano as preocupações e práticas republicanas. A internet, os aplicativos e as redes sociais potencializaram a radicalização dos discursos e diálogos entre pessoas que pensam iguais, não produzindo um

24 NETO SOUZA, Cláudio Pereira de. **Teoria Constitucional e Democracia Deliberativa**: um Estudo sobre Papel do Direito na Garantia das Condições para a Cooperação na Deliberação Democrática. Rio de Janeiro: Renovar, 2006, pp. 296-297.

25 ROBL FILHO, Ilton Norberto; MARRAFON, Marco Aurélio; MEDON, Filipe. O impacto das técnicas de inteligência artificial aplicadas no *deepfake* e nas redes automatizadas abalam o mercado livre de ideias e a democracia constitucional e deliberativa. In: **Economic Analysis of Law Review**, v. 13, n. 3, p. 32-47, out./dez. 2022. Disponível em: <https://portalrevistas.ucb.br/index.php/EALR/article/view/12527>, p. 40

26 ROBL FILHO, Ilton Norberto; MARRAFON, Marco Aurélio; MEDON, Filipe. O impacto das técnicas de inteligência artificial aplicadas no *deepfake* e nas redes automatizadas abalam o mercado livre de ideias e a democracia constitucional e deliberativa. In: **Economic Analysis of Law Review**, v. 13, n. 3, pp. 32-47, out./dez. 2022. Disponível em: <https://portalrevistas.ucb.br/index.php/EALR/article/view/12527>, p. 40

fórum público republicano na internet ou por meio de aplicativo, e sim uma enorme fragmentação social com grande dificuldade para permitir que diversas visões e tendências participem de debates sociais. Em síntese, ocorre qualitativamente o empobrecimento das reflexões, análises e discussões sociais. Ainda, o modelo de negócio das redes sociais impulsiona a segmentação dos grupos sociais e a degradação do espaço público, (...).²⁷

Como utilizar, então, as ferramentas de Inteligência Artificial para fortalecer a noção de democracia participativa?²⁸⁻²⁹ Como fazer com que a participação política se dê também de forma direta para além da escolha de representantes, com a possibilidade de iguais condições para influenciar decisões legítimas e livres de manipulações?³⁰

-
- 27 ROBL FILHO, Ilton Norberto; MARRAFON, Marco Aurélio; MEDON, Filipe. O impacto das técnicas de inteligência artificial aplicadas no *deepfake* e nas redes automatizadas abalam o mercado livre de ideias e a democracia constitucional e deliberativa. In: **Economic Analysis of Law Review**, v. 13, n. 3, pp. 32-47, out./dez. 2022. Disponível em: <https://portalrevistas.ucb.br/index.php/EALR/article/view/12527>, pp. 40-41. Na mesma direção: CALLEJÓN, Francisco Balaguer. As duas grandes crises do constitucionalismo diante da globalização no século XXI. **Espaço Jurídico Journal of Law**, 19 (3), 684, 2019.
- 28 “At an abstract level and as defined by so-called “classical” deliberative democrats (Mansbridge et al., 2010), democratic deliberation is “the public use of arguments and reasoning among free and equal individuals” (adapted from Cohen, 1989). The “use of arguments and reasoning” can be further specified as an exchange of arguments in which the participants aim to convince their interlocutors of the validity of a claim or, conversely, to refute a given claim.¹ Deliberation in that sense is distinct from bargaining, which consists in appealing to self-interest, or threatening (Elster, 1986). Democratic deliberation is also here meant as an intersubjective exercise among at least two individuals, as opposed to an internal dialogue in the vein of “deliberation within” (Goodin, 2005) or a deliberation occurring among entities larger than individuals, as in system-thinking.” (AITAMURTO, Tanja; LANDEMORE, Hélène. Crowdsourced deliberation: the case of an off-traffic law reform in Finland. **Policy & Internet**, [S.l.], v. 8, n. 2, p. 174-196, maio 2016).
- 29 Por mais sobre o conceito de democracia participativa: COSTA, Pietro. **Poucos, Muitos, Todos: Lições de História da Democracia**. Curitiba: Editora UFPR, 2012, p. 299-307.
- 30 MELLO, Marco Aurélio. A democracia participativa e a inconstitucionalidade do financiamento privado das campanhas eleitorais por pessoas jurídicas. **Revista**

3. A Inteligência Artificial como promotora da democracia participativa em tempos de “infocracia”

Como visto, o ambiente digital tem se caracterizado pelo aumento de informações, o que não necessariamente se traduz em qualidade, assim como por manipulações das mais diversas, sobretudo em razão da utilização maliciosa de ferramentas de Inteligência Artificial, como as já examinadas redes automatizadas e *bots* sociais. No entanto, a mesma ferramenta que pode ser utilizada para desinformar e criar instabilidade na rede, também pode ser uma agente na promoção da democracia participativa.

Em contributo seminal acerca do tema, Hélène Landemore explora como a Inteligência Artificial pode servir como uma possível solução para a tensão havida entre “duas condições aparentemente igualmente importantes da legitimidade democrática: de um lado, a deliberação; de outro, a participação em massa.”³¹ Para tanto, a autora analisa duas vi-

Populus, Salvador, n. 1, pp. 11-25, set. 2015. Na mesma direção, “Nada obstante isso, dentro de uma perspectiva jurídica que não abstrai de questões como legitimidade e justificação das decisões políticas, cabe destacar a contribuição da teoria da democracia deliberativa para a democracia constitucional contemporânea. Assim é que se incorporou ao conhecimento convencional que o processo democrático não se limita às eleições e à atuação das estruturas oficiais, mas deve ser exercido em um ambiente de permanente deliberação pública. Como tal, deve-se entender o debate acessível a todos, dentro de uma esfera pública (não estatal) que permita a participação ampla dos interessados em geral. Tal processo deliberativo tem o condão de pautar, racionalizar e legitimar as decisões tomadas pelas instâncias formais de poder. Instrumentos típicos dessa nova configuração da democracia incluem os meios de comunicação tradicionais, meios de comunicação alternativos (blogs, rádios comunitárias), a rede mundial de computadores e, mais recentemente, as mídias sociais. Mecanismo institucional que desempenha papel relevante nessa abertura dos órgãos de poder para a sociedade são as audiências públicas, presentes nos processos decisórios do Congresso Nacional, das agências reguladoras e até do Supremo Tribunal Federal.” (BARROSO, Luís Roberto. **Curso de Direito Constitucional Contemporâneo**. Rio de Janeiro: Saraiva Jur, 2024, 12. ed., p. 417)

- 31 LANDEMORE, Hélène. Can AI bring deliberative democracy to the masses? **NYU Law Review**, [S.l.], v. 95, pp. 1–20, 2020. Disponível em: <https://www.law>.

sões principais para uma forma de deliberação em massa ampliada por algoritmos, em que a Inteligência Artificial poderia desempenhar papéis cruciais: a chamada Deliberação Online em Massa, ou *Mass Online Deliberation* (MOD), e uma multiplicidade de mini públicos rotativos selecionados aleatoriamente.³²

Em primeiro lugar, Landemore alude ao conceito de *Mass Online Deliberation*, proposto pelo engenheiro russo aposentado Cyril Velikanov e depois teorizado por este em conjunto com Alexander Prossove. A ideia central aqui seria utilizar o *Big Data* e os algoritmos para criar uma plataforma de deliberação democrática em larga escala. Reunindo milhares de participantes em um espaço virtual comum de debates e tomada de decisões apoiadas na tecnologia, os algoritmos poderiam ajudar a organizar e agrupar as contribuições individuais, o que acabaria facilitando a compreensão das proposições de cada um.³³

Como explica a autora:

The concept of “Mass Online Deliberation” has been developed by a retired Russian engineer named Cyril Velikanov. His vision combines human judgment and the capacity of AI for clustering ideas and proposals into one possibly attractive vision of deliberative democracy for the masses. In a paper written with co-author Alexander Prossove, they theorize Mass Online Deliberation (MOD) as a process whereby thousands, perhaps hundreds of thousands of people could be brought into a common virtual space, where they are able to engage in a multi-stage deliberative pro-

nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf, p. 01.

- 32 LANDEMORE, Hélène. Can AI bring deliberative democracy to the masses? **NYU Law Review**, [S.l.], v. 95, pp. 1–20, 2020. Disponível em: <https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>, pp. 01.
- 33 LANDEMORE, Hélène. Can AI bring deliberative democracy to the masses? **NYU Law Review**, [S.l.], v. 95, pp. 1–20, 2020. Disponível em: <https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>, pp. 17-19.

cess of ideation, commenting and exchanging arguments, evaluation, and decision-making with the help of algorithms. They emphasize that for them “mass” means many people deliberating together in one common “room.” This is by contrast with the common method of having several small groups deliberating separately in several “rooms” (2022: 3 and section 6). They distinguish their model from Deliberative Polls (and by implication Citizens’ Assemblies), which typically break up large groups into smaller deliberative units and then develop various strategies to integrate the outputs of these separate units at the collective level afterwards. (...) In this vision of mass online deliberation, algorithms play an important albeit limited role. It sorts out and clusters proposals to offer “a “bird’s eye view onto the whole sea of participants’ contributions” so as to “make it easy for any participant to navigate across it” (Velikanov and Prossove 2022: 32). The exchanges among self-selected participants in the process are supposed to be content-moderated, facilitated, partially structured and organized, and ultimately judged and evaluated by random draws of humans themselves, on the basis of prodding by the algorithm. Velikanov and Prossove also entertain the possibility for AI to play a role as a translator in multilinguistic MODs, but again not so much to provide the translation as to distribute the function of translators to volunteers with the right skills.³⁴

No entanto, o modelo tem limitações claras. Em primeiro lugar, a ausência de testes e evidências empíricas, sobretudo em relação à sua capacidade de acomodar a população inteira de um determinado país. Além disso, há dúvidas razoáveis se o modelo conseguiria viabilizar deliberação de qualidade em vez de simplesmente acomodar engajamentos superficiais. Como assegurar, ainda, que as pessoas conseguiriam in-

34 LANDEMORE, Hélène. Can AI bring deliberative democracy to the masses? **NYU Law Review**, [S.l.], v. 95, pp. 1–20, 2020. Disponível em: <https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>, pp. 17-19.

fluenciar umas às outras no debate público?³⁵ Pode-se alegar, ademais, preocupações legítimas em relação ao enviesamento dos modelos algorítmicos, que eventualmente poderiam perpetuar discriminações caso não fossem adequadamente testados.

Finalmente, pode-se aludir à ideia de “erosão da esfera pública” descrita por Byung-Chul Han no livro “Infocracia: digitalização e a crise da democracia na filosofia”.³⁶ Segundo Han, o excesso de informações, rapidamente disseminadas, proporcionado, sobretudo, pelas redes sociais, teria contribuído para que a opinião pública, em vez de se fortalecer, se tornasse cada vez mais fragmentada e polarizada, em verdadeira massificação do pensamento. É assim que surgem expressões como a *infodemia*, que é a denominação conferida:

ao volume excessivo de informações, muitas delas imprecisas ou falsas (desinformação), sobre determinado assunto (como a pandemia, por exemplo), que se multiplicam e se propagam de forma rápida e incontrolável, o que dificulta o acesso a orientações e fontes confiáveis, causando confusão, desorientação e inúmeros prejuízos à vida das pessoas.³⁷

Como resultado desse processo, em vez de as pessoas desenvolverem seus próprios pensamentos a partir da interação com as demais no ambiente digital, elas acabariam sendo expostas a “filtros bolha” ou “bolhas informacionais”, com cristalização e reforço de suas próprias convicções, o que as levaria, em muitos cenários, a seguir influenciadores digitais em vez de construírem pensamentos pessoais, o que seria

35 LANDEMORE, Hélène. Can AI bring deliberative democracy to the masses? NYU Law Review, [S.l.], v. 95, pp. 1–20, 2020. Disponível em: <https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>, p. 19.

36 HAN, Byun-Chul, *Infocracia: digitalização e a crise da democracia*. Petrópolis: Vozes, 2022.

37 ACADEMIA BRASILEIRA DE LETRAS. **Nova palavra:** Infodemia. Disponível em: <https://www.academia.org.br/nossa-lingua/nova-palavra/infodemia>. Acesso em: 30 jun. 2025.

útil para o debate público. E tudo isso seria alcançado por meio de algoritmos devidamente treinados para priorizar conteúdo que gere engajamento em vez de oportunizar o pluralismo de ideias e visões de mundo.³⁸ Pode-se criticar, assim, a ideia de que tudo deveria ser resolvido por meio da tecnologia, como pontuado por Evgeny Morozov.³⁹

Em segundo lugar, Hélène Landemore analisa aquilo que ela designa como “uma multiplicidade de assembleias rotativas selecionadas aleatoriamente.”⁴⁰ Apresentada como uma alternativa ao modelo das *Mass Online Deliberation*, a autora cita como exemplo os debates ocorridos na França no ano de 2018, naquilo que se convencionou designar de “Grande Debate Nacional”, o qual foi lançado pelo Presidente francês Emmanuel Macron em resposta à chamada crise dos “Coletes Amarelos”, resultante de um aumento no imposto sobre combustíveis. A iniciativa consistiu em “um processo de dois meses que tentou envolver toda a população em uma deliberação em larga escala sobre quatro temas abrangentes, incluindo tributação, serviços e organização do Estado, transição ecológica e democracia e participação.”⁴¹

38 HAN, Byun-Chul, **Infocracia: digitalização e a crise da democracia**, Petrópolis: Vozes, 2022.

39 “Como o Vale do Silício continua a corromper nossa linguagem com sua interminável exaltação da disrupção e da eficiência – conceitos em desacordo com o vocabulário da democracia –, nossa capacidade de questionar o “como” da política fica debilitada. No Vale do Silício, a resposta padronizada para isso é o que chamo de solucionismo: os problemas devem ser resolvidos por meio de aplicativos, sensores e ciclos infinitos de retroalimentação – todos fornecidos por startups.” (MOROZOV, Evgeny, **Big Tech: A ascensão dos dados e a morte da política**, São Paulo: Ubu, 2018, versão digital).

40 LANDEMORE, Hélène. Can AI bring deliberative democracy to the masses? **NYU Law Review**, [S.l.], v. 95, pp. 1–20, 2020. Disponível em: <https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>, p. 22.

41 LANDEMORE, Hélène. Can AI bring deliberative democracy to the masses? **NYU Law Review**, [S.l.], v. 95, pp. 1–20, 2020. Disponível em: <https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>, p. 12.

Os debates foram, então, conduzidos em milhares de reuniões locais envolvendo grupos entre 10 e 300 pessoas, o que foi combinado a outras iniciativas, como plataformas governamentais online para recebimento de contribuições e conferências a nível nacional, com pessoas selecionadas aleatoriamente. Com efeito, cerca de 1,5% da população francesa acabou participando diretamente dos debates. No entanto, Landemore chama a atenção para importante fato: apesar de os debates não terem sido amplamente conduzidos por meio de ferramentas tecnológicas, as tecnologias digitais e o *big data* (embora não a IA), “mostraram-se essenciais e, de fato, absolutamente necessários para analisar e processar a enorme quantidade de dados gerada pelo Grande Debate.”⁴² Todavia, a autora faz a ressalva de que a forma como os dados foram estruturados acabou tornando difícil a extração de conclusões significativas.⁴³

Após analisar os aspectos positivos e as limitações dos dois cenários apresentados, Landemore sistematiza algumas funções principais da Inteligência Artificial que poderiam ajudar a escalar e melhorar a deliberação nos modelos. São elas: (i) a facilitação; (ii) tradução; (iii) verificação de fatos (*fact-checking*); (iv) agrupamento e organização de dados (*data clustering* e *aggregation*); (v) rastrear todas as interações que qualquer pessoa teve com qualquer outra e medir o grau de sobreposição de conteúdo; (vi) medir a qualidade da deliberação; (vii) captar fotografias cognitivas do grupo; (viii) compartilhar consensos entre grupos; (ix) semear assembleias com ideias de alto potencial.⁴⁴

42 LANDEMORE, Hélène. Can AI bring deliberative democracy to the masses? *NYU Law Review*, [S.l.], v. 95, pp. 1–20, 2020. Disponível em: <https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>, p. 15.

43 LANDEMORE, Hélène. Can AI bring deliberative democracy to the masses? *NYU Law Review*, [S.l.], v. 95, pp. 1–20, 2020. Disponível em: <https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>, p. 15.

44 LANDEMORE, Hélène. Can AI bring deliberative democracy to the masses? *NYU Law Review*, [S.l.], v. 95, pp. 1–20, 2020. Disponível em: <https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>, p. 15.

Tais funções possibilitadas pela Inteligência Artificial podem ser especialmente úteis nessa árdua missão de ordenar a grande massa de dados obtidos. Merece destaque, por exemplo, a função de facilitação, segundo a qual a Inteligência Artificial poderia atuar como uma moderadora imparcial – apesar das ressalvas em relação a essa possibilidade do ponto de vista técnico – em discussões em grupo online. Desse modo, auxiliaria a assegurar a igualdade de voz entre as pessoas presentes nas deliberações, o que nem sempre é alcançado no mundo real, já que os facilitadores de carne e osso são frequentemente marcados por vieses e não conseguem se distanciar da indesejada parcialidade.⁴⁵

Além disso, a Inteligência Artificial poderia reduzir barreiras de linguagem, assim como atuar na verificação de fatos, reduzindo os riscos de desinformação que possa contaminar o debate público. No fundo, é a tecnologia auxiliando a combater os malefícios trazidos pela própria tecnologia.

Malgrado tais possibilidades, Landemore conclui no sentido de que, “no fim das contas, precisamos abandonar o ideal original de uma deliberação totalmente inclusiva e simultânea de todos com todos.”⁴⁶

Em suas palavras:

AI can help scale deliberative democracy in multiple capacities. But in the end, we have to let go of the original ideal of full inclusive, simultaneous deliberation of all with all. The best we can probably have is either millions of self-selected people processing arguments and ideas together on an AI-

nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf, pp. 28-33.

45 LANDEMORE, Hélène. Can AI bring deliberative democracy to the masses? **NYU Law Review**, [S.l.], v. 95, pp. 1–20, 2020. Disponível em: <https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>, pp. 28-29.

46 LANDEMORE, Hélène. Can AI bring deliberative democracy to the masses? **NYU Law Review**, [S.l.], v. 95, pp. 1–20, 2020. Disponível em: <https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>, p. 33.

augmented online platform (Mass Online Deliberation) or a substantial percentage of the population engaged in deliberative mini-publics being rotated a number of times (the regional assemblies of the Great National Debate on steroids). Even these approximations of all-inclusive deliberation, however, are incredibly demanding, time-consuming, and presumably costly. But they may be the bridges we need to bridge the ideal of deliberative democracy and that of mass participation.

4. Síntese conclusiva: uma ferramenta imperfeita, que demanda calibragem

Num mundo marcado pela infodemia, bolhas informacionais, influenciadores digitais, “vozes imbecis”, desinformação e redes automatizadas de *bots*, alcançar a plena e livre participação democrática torna-se um colossal desafio. Não obstante, a própria tecnologia pode ser utilizada para combater – ou ao menos atenuar – os males por ela produzidos, despoluindo um ambiente marcado pela disseminação de conteúdo falso e de ódio. Trata-se, assim, de uma questão de calibragem.

Ferramentas de participação e deliberação diretas, se adequadamente executadas, podem ser formas úteis de se alcançar debates de qualidade e que promovam o avanço na sociedade. Para isso, no entanto, é necessário assegurar, tanto quanto possível, que o ambiente digital em que se debate seja verdadeiramente inclusivo. Nessa direção, iniciativas como a de responsabilizar as plataformas digitais pela não eliminação de contas inautênticas e redes automatizadas revela-se como um contributo significativo. Manter a higidez desse espaço deliberativo é missão a ser compartilhada entre Estado, plataformas e sociedade civil.

Em última análise, o ambiente digital acaba sendo um reflexo da sociedade. Contudo, sem a devida atenção para os fenômenos que o conformam e, principalmente, deformam, a deliberação virtual pode acabar se afastando da democracia participativa, justa, livre e solidária, aproximando-se, antes, de uma vocalização de discursos manipulados

ou, como prefere Umberto Eco: imbecis. Por isso, para além de pensar em mecanismos de promoção de democracia participativa aliada à tecnologia, necessário se faz promover também a imprescindível educação digital, que se revela como pilar fundamental para a construção de um ambiente plural.

Filipe Medon • Doutor e Mestre em Direito Civil pela Universidade do Estado do Rio de Janeiro (UERJ). Professor de Direito Civil na FGV Direito Rio, onde também é pesquisador no Centro de Tecnologia e Sociedade (CTS-FGV). Professor convidado de cursos de Pós-Graduação e Extensão de diversas instituições no Brasil e fora. Membro da Taskforce “AI in Education” do Projeto “AI and the Future of Work” da Universidade de Columbia, do Instituto Brasileiro de Estudos de Responsabilidade Civil (IBERC), da govDADOS, da Comissão de Proteção de Dados e Privacidade da OAB/RJ, onde é coordenador de Inteligência Artificial, membro consultor da Comissão de Inteligência Artificial da OAB/RJ, tendo integrado, ainda, a Comissão de Juristas do Senado Federal responsável pela criação do Marco Legal da IA. Seus trabalhos foram citados mais de 300 vezes por inúmeros meios de comunicação internacionais, incluindo a TIME Magazine e The Washington Post. Foi Professor Substituto de Direito Civil na UFRJ, além de palestrante no evento “Artificial Intelligence for Information Accessibility”, realizado pelo Information for All Programme da UNESCO (2021). Membro fundador do Laboratório de Direito e Inteligência Artificial da UERJ (Lab-DIA). Advogado, consultor e parecerista. Autor do livro: “Inteligência Artificial e Responsabilidade Civil: autonomia, riscos e solidariedade” (3ª edição prevista para 2025). Instagram: @filipe.medon.

Democracia movida a dados: valores democráticos e regulação na era da inteligência artificial

João Victor Archegas
Diana de Souza Fernandes

Resumo

O artigo sustenta que a transição da “sociedade em rede” para uma “sociedade algorítmica” colocou os fluxos de dados no centro das dinâmicas econômicas, sociais e políticas – ampliando riscos de concentração informacional. Como contrapeso democrático, defende-se a abertura de dados para fortalecer a accountability e a deliberação pública. Propõe-se, ainda, a tecnodiversidade e a inovação aberta como pilares para redistribuir poder técnico, à luz das assimetrias de infraestrutura computacional e cadeias de semicondutores. No campo regulatório, analisa-se o PL 2.338/2023 (inspirado no AI Act europeu), apontando um trilema entre incentivo à inovação, mitigação de riscos e regras claras, e o risco de sobrecarga regulatória frente às metas do Plano Brasileiro de IA. Como alternativa, destaca-se a “IA aberta” e a experiência de Goiás, que combina proteção de direitos com fomento a capacidades locais. Ao final, o artigo conclui que uma democracia movida a dados exige regulação e infraestrutura abertas, inclusivas e sensíveis a contextos locais.

Abstract

The article argues that the transition from a “network society” to an “algorithmic society” has placed data flows at the center of economic,

social, and political dynamics—increasing the risks of information concentration. As a democratic counterweight, it advocates for open data to strengthen accountability and public deliberation. It also proposes technodiversity and open innovation as pillars for redistributing technical power, in light of asymmetries in computing infrastructure and semiconductor chains. In the regulatory field, Bill 2.338/2023 (inspired by the European AI Act) is analyzed, pointing to a trilemma between encouraging innovation, mitigating risks, and clear rules, and the risk of regulatory overload in view of the goals of the Brazilian AI Plan. As an alternative, it highlights “open AI” and the experience of Goiás, which combines the protection of rights with the promotion of local capabilities. Finally, the article concludes that a data-driven democracy requires open, inclusive, and context-sensitive regulation and infrastructure.

1. Introdução: da sociedade em rede à sociedade algorítmica

A era digital é marcada por tecnologias e técnicas disruptivas que não apenas transformam, mas também estruturam novas relações econômicas, sociais e políticas. Em seu estudo paradigmático sobre o tema, ao analisar a influência da Internet comercial sobre tecido social ainda nos anos 90, Manuel Castells cunhou o termo “sociedade em rede” para se referir aos impactos do ecossistema digital de informações na organização social contemporânea¹. Desde então, entretanto, pelo menos duas mudanças significativas ocorreram: a “plataformização” da Internet, em especial a partir da ascensão de grandes redes sociais como Facebook no início dos anos 2000, e o descobrimento do potencial econômico do “excedente comportamental” que é próprio de grandes plataformas digitais². É dizer, provedores de aplicações passaram a explorar o volume

1 CASTELLS, Manuel. **A Sociedade em Rede**. São Paulo: Paz & Terra, 2013.

2 ZUBOFF, Shoshana. **A Era do Capitalismo de Vigilância**. São Paulo: Intrínseca, 2021.

cada vez maior de dados compartilhados a partir da interação de seus usuários com os serviços digitais disponibilizados, possibilitando a personalização da experiência digital e até mesmo a previsão de padrões de comportamento.

São esses dados, pessoais e comportamentais, que estão por trás de uma verdadeira revolução econômica, social e política. Econômica porque bilhões de dólares são movimentados anualmente no mercado de plataformas digitais a partir do micro-direcionamento de anúncios, produtos e serviços. Social porque relações interpessoais acontecem prioritariamente dentro de serviços digitais que reconfiguram a forma como nos conectamos, comunicamos e consumimos informações. Política porque a Internet assumiu o papel de “praça pública” da era digital e as principais ações políticas, incluindo aquelas que ocorrem em períodos eleitorais, acontecem antes em plataformas digitais do que fora delas. De uma sociedade em rede, passamos a viver em uma sociedade de dados. Alguns autores, considerando o papel central de algoritmos e sistemas de inteligência artificial (IA) treinados a partir dessas informações, se referem a esse novo fenômeno como “sociedade algorítmica”³.

Embora essa questão possa ser analisada a partir de inúmeros ângulos, o presente estudo se debruça sobre o tema da “democracia movida a dados” (*data-driven democracy*) e como valores democráticos podem ser fortalecidos a partir de inovações tecnológicas, em especial quando olhamos para o tema da regulação da IA no Brasil. Vale ressaltar, desde já, que quando falamos em democracia movida a dados não estamos nos referindo apenas ao uso de dados para aprimorar processos democráticos e a construção de decisões públicas, mas também (e principalmente) ao uso de dados para o treinamento de algoritmos e sistemas de IA que impactam diretamente a proteção de valores democráticos como liberdade, pluralismo político, Estado de Direito e respeito aos direitos

3 DE GREGORIO, Giovanni. **Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society**. Cambridge: Cambridge University Press, 2022.

fundamentais. Em outras palavras, não basta olhar para como a tecnologia pode ser implementada pelo setor público para promover valores democráticos, é preciso também compreender que toda e qualquer tecnologia, em especial a IA, pode impactar o futuro da democracia de forma profunda e total.

Para isso, o artigo é dividido em quatro seções. Nesta introdução são apresentados alguns pressupostos da discussão proposta, como a centralidade dos fluxos de dados para a democracia na era digital e a dicotomia entre dados fechados e abertos. Na segunda seção é apresentado o conceito de “tecnodiversidade”, que nos faz refletir sobre as bases técnicas e epistemológicas sobre as quais a IA moderna é construída e a necessidade de apostar em inovação aberta como um vetor de promoção dos valores democráticos diante das transformações digitais. Na terceira seção avaliamos como o debate sobre regulação da IA no Brasil, pelo menos em seu atual estágio, pode levar ao aprofundamento de alguns dos desafios identificados ao longo do artigo e o que podemos fazer para reverter essa situação. Por fim, na quarta seção são apresentadas as considerações finais a partir das conclusões dos tópicos anteriores.

1.1 Fluxos de dados e a democracia na era digital

Conforme passamos de uma sociedade em rede para uma sociedade algorítmica, os conjuntos de dados coletados e processados por empresas de tecnologia se tornaram maiores e mais complexos, consumindo proporcionalmente mais poder computacional (ou *compute*) para serem analisados. Surge daí, portanto, a ideia de *big data* para se referir a dados que contam como uma alta variedade, atingindo grandes volumes e com maior velocidade (os chamados “três Vs” do *big data*)⁴.

4 Ver, nesse sentido, a definição de *big data* da empresa de tecnologia e infraestrutura digital Oracle. Disponível em <https://www.oracle.com/br/big-data/what-is-big-data/>.

Se consolida como regra geral que quanto maior o volume de dados e o poder computacional, melhores serão os resultados obtidos a partir da análise e do processamento dessas informações. Essa mesma lógica vale, em certa medida, tanto para plataformas digitais que analisam dados comportamentais de seus usuários para micro-direcionamento de anúncios, quanto para empresas que raspam dados da Internet para treinar modelos fundacionais de IA.

É justamente essa regra geral que está por trás da corrida por dados – e, conseqüentemente, poder – protagonizada nos últimos anos por países como China e Estados Unidos e grandes empresas como Meta, Google, Amazon, OpenAI, Baidu, Alibaba e Huawei. Para além de preocupações legítimas sobre segurança nacional, é justamente isso, por exemplo, que está no pano de fundo das tensões entre China e EUA a respeito da coleta de dados pelo TikTok: “quem tiver mais informações sobre seus cidadãos (e os cidadãos do mundo) terá uma vantagem estratégica na corrida digital. E se, como dizem por aí, *‘data is the new oil’*, então a China é a nova Arábia Saudita. A população chinesa é quatro vezes maior do que a estadunidense, o que multiplica ao menos pela mesma grandeza as possibilidades de coleta de dados sobre o comportamento humano”⁵. Quanto melhores forem seus sistemas algorítmicos e de IA, maior será a capacidade de um país ou de uma empresa de influenciar as relações econômicas e políticas a partir da sua visão de mundo.

Note-se, ainda, o desafio da desinformação em plataformas digitais que vem sendo amplamente discutido nos últimos anos. O combustível que alimenta os sistemas contemporâneos de desinformação não é apenas o conteúdo falso, mas os dados comportamentais que fornecemos constantemente em nossas interações digitais. Como destaca Philip Howard, “políticos costumavam ter dados de pesquisas e

5 ARCHEGAS, João Victor. TikTok e a nova ordem digital. **JOTA**, 31 de julho de 2020. Disponível em <https://www.jota.info/opiniao-e-analise/artigos/tiktok-e-a-nova-ordem-digital-entre-protecao-de-dados-e-concorrenca-desleal>.

enquetes para interpretar o que os eleitores pensavam”; hoje, “temos dados comportamentais sobre o que as pessoas realmente fazem”⁶. A mudança do uso tradicional de pesquisas de opinião e enquetes para a coleta massiva de dados sobre comportamentos reais dos usuários na esfera digital marca uma virada significativa na política moderna. A eficácia das campanhas de desinformação, assim, não reside apenas na sofisticação das mensagens falsas ou distorcidas, mas na capacidade de direcionamento preciso viabilizada pela concentração desses dados em poucas mãos.

Larry Diamond, em sua análise sobre o declínio democrático, alerta que “o maior perigo [para a democracia] é a distorção da verdade em escala industrial, quando governos e grupos políticos lançam operações de informação altamente organizadas”⁷. Nesse contexto, o problema central não é a distorção da verdade em si, mas a forma como os fluxos de dados são explorados para potencializá-la em benefício de poucos. O desafio, portanto, não é apenas combater o conteúdo enganoso que circula nas redes, mas reequilibrar o fluxo de dados dentro das democracias – um ponto enfatizado por Howard ao afirmar que o foco não deve estar no sintoma (as notícias falsas), mas na estrutura que permite sua proliferação⁸. Isso exige pensar a proteção de dados pessoais como estratégia central na contenção das distorções geradas por campanhas coordenadas de manipulação da informação, na preservação da integridade do debate público e na proteção de valores democráticos na era digital de forma geral.

6 HOWARD, Philip N. **Lie Machines**: How to save democracy from troll armies, deceitful robots, junk news operations, and political operatives. New Haven: Yale University Press, 2020, p. 3 (tradução livre).

7 DIAMOND, Larry. **Ill Winds**: Saving democracy from Russian rage, Chinese ambition, and American complacency. Nova Iorque: Penguin Press, 2019, p. 232 (tradução livre).

8 HOWARD, Philip N. **Lie Machines**: How to save democracy from troll armies, deceitful robots, junk news operations, and political operatives. New Haven: Yale University Press, 2020, p. 10.

1.2 Entre dados fechados e abertos

Se a concentração de dados em poucas mãos pode representar um risco para a democracia, sua descentralização ou abertura representa um importante contrapeso em favor dos valores democráticos. Dados abertos tornam visíveis estruturas e padrões antes opacos, permitem o escrutínio público de decisões e possibilitam que cidadãos, organizações e governos aprimorem políticas públicas com base em evidências acessíveis e participação ativa. A abertura não se refere apenas à disponibilização formal de informações, mas à criação de condições técnicas, jurídicas e institucionais para que esses dados sejam compreensíveis, reutilizáveis e efetivamente empregados em benefício coletivo.

No Brasil, os observatórios sociais são um exemplo emblemático de como a abertura de dados orçamentários e de compras públicas pode fortalecer a *accountability* das autoridades locais⁹. Presentes em dezenas de municípios, esses observatórios mobilizam cidadãos voluntários para acompanhar licitações, fiscalizar obras e prevenir desvios com base em dados disponíveis nos portais de transparência e em solicitações feitas a partir da Lei de Acesso à Informação. A partir desse controle social sistemático, milhares de editais foram ajustados e milhões de reais deixaram de ser gastos indevidamente, ao mesmo tempo em que a cultura da participação cívica foi estimulada em comunidades historicamente afastadas dos processos decisórios. O que antes era restrito a órgãos técnicos de controle passou a ser, com os dados abertos, objeto de ação coletiva organizada.

Mas os impactos democráticos dos dados abertos não se limitam à fiscalização de editais e contas públicas. A experiência de Taiwan com a plataforma vTaiwan mostra que a abertura pode ser aplicada diretamente à deliberação política¹⁰. Utilizando ferramentas digitais e dados

9 Para mais informações sobre a metodologia padronizada para a criação e operacionalização de observatórios sociais desenvolvida pelo Observatório Social do Brasil (OSB), ver <https://osbrasil.org.br/o-que-e-o-observatorio-social-do-brasil-osb/>.

10 HORTON, Chris. The simple but ingenious system Taiwan uses to crowdsource its laws. **MIT Technology Review**, 21 de agosto de 2018. Disponível em <https://>

colaborativos, o governo taiwanês conduziu consultas públicas amplas sobre temas polêmicos, como a regulação de plataformas de transporte por aplicativo. Por meio de tecnologias participativas e transparência total nas contribuições, cidadãos influenciaram decisões legislativas concretas. Essa abertura do processo decisório, mais do que uma inovação técnica, representa um novo pacto democrático: o dado não apenas informa, mas estrutura o diálogo entre governo e sociedade, redirecionando os fluxos mencionados acima.

Por fim, iniciativas como o projeto *Follow the Money*¹¹, na Nigéria, e a plataforma *Global Forest Watch*¹² demonstram como dados abertos podem ser essenciais para garantir justiça social e sustentabilidade ambiental. No primeiro caso, dados orçamentários foram usados por comunidades locais para pressionar o governo a cumprir promessas de investimento em saúde pública, revertendo cenários de negligência letal. No segundo, dados de satélite sobre desmatamento permitiram a atuação de jornalistas e ONGs para denunciar crimes ambientais e subsidiar respostas mais rápidas das autoridades. Em ambos os casos, a abertura rompeu o monopólio da informação e permitiu que a sociedade civil atuasse como agente ativo na correção de rumos das políticas públicas.

2. Tecnodiversidade como chave para uma verdadeira democracia digital

Dados abertos fortalecem valores democráticos, como participação, confiança, igualdade e *accountability*, construindo sistemas de go-

www.technologyreview.com/2018/08/21/240284/the-simple-but-ingenuous-system-taiwan-uses-to-crowdsource-its-laws/.

11 MOSCOSO, Sandra et al. Open data on the ground: Nigeria's Follow the Money initiative. **World Bank Blogs**, 14 de agosto de 2014. Disponível em <https://blogs.worldbank.org/en/opendata/open-data-ground-nigeria-s-follow-money-initiative>.

12 Para mais informações sobre a iniciativa *Global Forest Watch* e as bases de dados disponibilizadas de forma aberta, ver <https://data.globalforestwatch.org/>.

vernança mais robustos e resilientes¹³. Nada obstante, a opção por inovação aberta, seja do ponto de vista do setor público ou privado, não está conectada apenas a uma escolha de desenvolvimento tecnológico. Tecnologias de ponta, em especial sistemas de IA que prometem otimizar as iniciativas mencionadas anteriormente e criar inúmeras outras oportunidades para o fortalecimento de valores democráticos, dependem de uma infraestrutura digital adequada para seu treinamento e implementação. Essa infraestrutura, entretanto, está concentrada em alguns poucos países e empresas, o que, conseqüentemente, agrava ainda mais a concentração de dados explorada até aqui.

É nesse contexto que o conceito de *tecnodiversidade*, desenvolvido pelo filósofo Yuk Hui, oferece uma lente crítica para repensarmos a relação entre tecnologia e democracia no século XXI¹⁴. Em oposição à ideia dominante de uma “história universal” da tecnologia – que pressupõe soluções técnicas homogêneas e descontextualizadas – Hui propõe uma inflexão epistemológica que considera os sistemas técnicos como construções historicamente situadas, moldadas por valores culturais, tradições locais e experiências sociais específicas. A tecnologia, assim, não é uma entidade uniforme e universalmente aplicável, mas uma construção multifacetada enraizada nos contextos de onde ela emerge¹⁵.

2.1 Singularidade e democracia como antônimos

A abordagem proposta pela tecnodiversidade, assim, nos afasta das narrativas deterministas e reducionistas que frequentemente dominam o discurso global sobre inovação digital, especialmente no campo

13 VERHULST, Stefaan G. YOUNG, Andrew. Open Data in Developing Economies: Toward building an evidence base of what works and how. **The Governance Lab**, julho de 2017, p. 9-10. Disponível em <https://odimpact.org/files/odimpact-developing-economies.pdf>.

14 HUI, Yuk. **Tecnodiversidade**. São Paulo: Ubu Editora, 2020.

15 Para mais informações sobre o tema, ver o projeto do ITS Rio em <https://www.technodiversity.org/>.

da inteligência artificial. Termos como “singularidade” ou “inteligência artificial geral” (*artificial general intelligence* ou AGI) operam como metanarrativas que naturalizam uma trajetória única de desenvolvimento – geralmente alinhada com os interesses geopolíticos e econômicos do Norte Global – e que invisibilizam outras formas possíveis de imaginar e construir nosso futuro digital. A tecnodiversidade, por sua vez, nos convida a considerar as múltiplas “ecologias tecnológicas” que já existem e que podem florescer quando reconhecidas e nutridas a partir de suas particularidades culturais, geográficas e históricas.

Mais do que uma crítica teórica, a tecnodiversidade oferece uma diretriz normativa e política: a de que a inovação digital não deve estar subordinada apenas a critérios de eficiência ou maximização de valor econômico, mas sim ancorada na valorização de saberes locais, práticas comunitárias e princípios democráticos. Essa postura é especialmente relevante no contexto atual, em que a infraestrutura computacional global é altamente concentrada e os modelos de IA mais avançados refletem dados, línguas, normas e interesses específicos. Ao cultivar uma perspectiva tecnodiversa, somos desafiados a imaginar e construir futuros tecnológicos plurais, em que diferentes visões de mundo possam coexistir – e onde a democracia digital não seja apenas uma promessa, mas uma possibilidade real e distribuída.

2.2 Infraestrutura digital e inovação aberta na era da IA

Essa concentração de infraestrutura digital não é apenas uma metáfora: trata-se de uma realidade empírica e mensurável. Segundo dados de 2021, Estados Unidos e China concentram aproximadamente 50% de todos os data centers de hiperescala do mundo, formando uma espécie de “*Compute North*” que delimita a esfera de poder computacional na era da IA¹⁶. Nenhum dos 100 *clusters* de computação de alto desempenho

16 United Nations. **Digital Economy Report 2021 – Cross-border data flows and development:** For whom the data flows. Disponível em <https://unctad.org/sys->

com capacidade real para treinar grandes modelos de IA está localizado em países em desenvolvimento, e apenas um país africano aparece entre os 300 primeiros¹⁷. Esses dados revelam o abismo tecnológico existente, um abismo que não se limita à infraestrutura, mas impacta diretamente a capacidade de inovação autônoma de países em desenvolvimento.

Além disso, os custos para acesso a essa infraestrutura são proibitivos: estima-se que empresas sediadas no Sul Global precisam investir cerca de 70 milhões de dólares para treinar, por apenas três meses, um grande modelo de linguagem (LLM)¹⁸. Esse cenário é agravado pela hiperconcentração da cadeia de suprimentos de semicondutores, setor crucial para a computação moderna. Mais de 75% da capacidade global de fabricação de chips está localizada na Ásia, que também domina as etapas de montagem e testes com aproximadamente 90% do mercado¹⁹. Empresas chinesas e taiwanesas lideram essas operações, o que torna ainda mais complexa e assimétrica a integração de países fora desse circuito produtivo. Em um contexto em que os avanços em IA dependem de acesso massivo a dados, energia e capacidade computacional, essa desigualdade estrutural impõe barreiras profundas à construção de uma democracia digital atenta aos valores e princípios locais de outros países, tornando urgente o debate sobre acesso equitativo a recursos técnicos e o fortalecimento de capacidades locais de inovação.

Diante desse cenário de assimetrias estruturais, a tecnodiversidade pode servir como fundamento para um novo modelo de inovação que priorize a abertura – de dados, de modelos, de infraestrutura e de governança – como caminho para redistribuir poder e ampliar o

tem/files/official-document/der2021_en.pdf.

17 United Nations. **Governing AI for Humanity**: Final Report 2024. Disponível em https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf, p. 14

18 *Ibidem*, p. 62.

19 Deloitte. 2024 **Semiconductor Industry Outlook**. Disponível em <https://www2.deloitte.com/us/en/pages/technology-media-and-telecommunications/articles/semiconductor-industry-outlook.html>

acesso à IA. Uma abordagem nesses moldes pressupõe que sistemas de IA devem ser treinados não apenas com base em grandes volumes de dados, mas com atenção à pluralidade de contextos, idiomas e valores que compõem as sociedades democráticas. Isso só será possível com políticas públicas que incentivem a criação e compartilhamento de infraestruturas públicas de dados e computação, o financiamento de repositórios abertos e diversos e a articulação de redes internacionais que promovam o desenvolvimento colaborativo de modelos abertos. Ao contrário das iniciativas proprietárias e centralizadas, uma IA aberta – fundamentada em tecnodiversidade – permitiria que diferentes comunidades se apropriem da tecnologia para responder às suas próprias necessidades, fortalecendo os pilares democráticos de participação, inclusão e autonomia.

3. Regulando a IA para fortalecer a democracia

A regulação da IA no Brasil está sendo discutida pelo Congresso Nacional no âmbito do Projeto de Lei 2.338/2023. Após ser aprovado pelo Senado no final de 2024, o texto aguarda votação na Câmara dos Deputados. Inspirado pelo *AI Act* da União Europeia, o PL caminha no sentido de categorizar os usos de IA no Brasil a partir de uma moldura de risco: sistemas considerados de risco excessivo seriam banidos, enquanto sistemas classificados como de alto risco precisariam passar por uma “análise de impacto algorítmico” supervisionada por uma autoridade regulatória ainda a ser definida.

Antes de tudo, entretanto, um primeiro desafio regulatório emerge: qual é a melhor definição de “inteligência artificial”? Essa é uma questão central para o debate aqui proposto. Afinal, como regular sem antes saber exatamente qual é o objeto da regulação? Ainda no século XX, Alan Turing formulou uma das mais influentes perspectivas para essa questão. Para ele, a pergunta correta a se fazer não era se as máquinas poderiam de fato “pensar”, mas sim se elas poderiam vencer um hipotético “jogo da imitação” – ou seja, convencer um interlocutor humano de

que estava falando com outro humano, e não com uma máquina²⁰. Ele sugeriu a possibilidade de construção de máquinas universais, capazes de simular processos cognitivos humanos e aprender com a experiência (algo parecido com o que hoje entendemos por *machine learning*).

John McCarthy, posteriormente, reforçou a definição de IA como “a ciência e a engenharia de criar máquinas inteligentes, especialmente programas de computador inteligentes”²¹, vinculando a inteligência com a habilidade de atingir certos objetivos pré-estabelecidos. Apesar das diversas contribuições do autor, de certa forma apenas alterou-se o foco da dificuldade conceitual: tão difícil quanto definir “inteligência” é definir “objetivo” e “intenção”²². O mesmo pode ser dito em relação ao “agir racionalmente”, termo utilizado por Russel e Norvig²³. Além de apresentarem alto grau de subjetividade, todos os conceitos mencionados até agora dependem da sua inserção em um contexto específico.

A experiência jurídico-legislativa demonstra que a indefinição gera insegurança no campo regulatório. Se o conceito adotado for muito aberto, pode abranger uma variedade enorme de aplicações e usos, os quais podem invadir indevidamente matérias já reguladas de outras formas: significaria um escopo amplo demais para uma regulação eficiente²⁴ e risco de relevante sobreposição de normas. Por outro lado, a adoção de um conceito muito restritivo poderia resultar em insuficiência regulatória, ocasionando graves lacunas jurídicas e rápida desatua-

20 TURING, Alan Mathison. Computing Machinery and Intelligence. **Mind**, v. 59, n. 236, p. 433–460, 1 out. 1950.

21 MCCARTHY, John. **What Is Artificial Intelligence?** Stanford University, 2004. Disponível em <https://www.cs.ryerson.ca/mes/courses/cps721/ln/whatisai.pdf>. Acesso em 10 jul 2025.

22 SCHERER, Matthew U. Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. **Harvard Journal of Law & Technology**, v. 29, n. 2, 2016. P. 361

23 RUSSELL, Stuart; NORVIG, Peter. **Artificial Intelligence: a modern approach**. Pearson, 2020.

24 BUITEN, Miriam C. Towards intelligent regulation of artificial intelligence. **European Journal of Risk Regulation**, v. 10, n. 1, p. 41–59, 2019. P. 45.

lização. Por isso, a tendência é a compreensão deste conceito de forma aberta e dinâmica, mas sujeito a determinadas balizas interpretativas²⁵. Apenas para fins deste artigo, será adotada uma definição propositalmente circular: “inteligência artificial” refere-se, neste trabalho, a máquinas capazes de realizar tarefas que, se realizadas por humanos, demandam inteligência²⁶.

3.1 Dano, responsabilização, inovação e o trilema regulatório

As facilidades e vantagens no uso da inteligência artificial são evidentes. São frequentes notícias de facilitação de diagnósticos médicos, *chatbots* para atendimentos comerciais, automação de atividades rotineiras e repetitivas, análises de padrões para otimização de resultados (como no combate às mudanças climáticas) e uso aprimorado da linguagem, inclusive de programação. Por outro lado, há também informações sobre as falhas e danos, desde acidentes fatais envolvendo veículos autônomos até vieses discriminatórios no uso de sistemas de reconhecimento facial.

Assim, se as vantagens da IA são inúmeras, os riscos também são: por este motivo, há uma crescente pressão por mais transparência e regulação dessas tecnologias. Este, contudo, não é um problema fácil, já que as características dos sistemas de IA trazem dificuldades para regulação prévia e posterior ao dano (*ex ante* e *ex post*). A sua regulação *ex ante* é dificultada porque projetos de IA podem ser: (a) discretos e não necessitarem de grandes estruturas para seu desenvolvimento; (b) difu-

25 ARCHEGAS, João Victor; MAIA, Gabriela. **O que é inteligência artificial? Análise em três atos de um conceito em desenvolvimento**. In: Cadernos Adenauer xxiii (2022), nº2. Inteligência artificial: aplicações e implicações. Rio de Janeiro: Fundação Konrad Adenauer, junho 2022. P. 12.

26 SCHERER, Matthew U. Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. **Harvard Journal of Law & Technology**, v. 29, n. 2, 2016. P. 362.

sos e desenvolvidos em locais e instituições sujeitas a jurisdições diversas; (c) segmentados, valendo-se de diversos componentes que apenas juntos tem determinado resultado, dificultando a localização da origem de falhas; (d) opacos, já que as tecnologias que os fundamentam são obscuras para os reguladores²⁷.

Regular apenas *ex post*, por outro lado, pode ensejar danos de difícil ou até mesmo impossível reparação. Sistemas de IA podem operar de formas imprevisíveis até mesmo para seus desenvolvedores, gerando riscos que não podem ser totalmente antecipados e, portanto, mitigados. Uma IA pode sair do controle não só de seus operadores, legalmente responsáveis, mas também de qualquer outro ser humano, tornando-se imprevisível²⁸. Há, ainda, o problema da imputação de responsabilidade para sistemas híbridos que envolvem humano e máquina. Tais características dificultam sobremaneira a fiscalização preventiva e a responsabilização no mercado, obstaculizando a identificação e localização dos envolvidos, bem como a mensuração e aplicação de sanções. Não é fácil comprovar nexo causal direto ou quantificar o dano, e nem mesmo executar sanções aplicadas quando os envolvidos possivelmente estão em jurisdições diferentes, ou quando a empresa não tem uma estrutura física e centralizada.

Dada a complexidade fática da IA, para sua regulamentação, será preciso enfrentar uma espécie de trilema²⁹. São três os objetivos desejados nesta regulação, mas apenas dois podem ser plenamente alcançados ao mesmo tempo, em detrimento de um terceiro. É preciso trabalhar com uma regulação que busque um equilíbrio dinâmico entre: (1) incentivo à inovação; (2) integridade do mercado e mitigação de

27 SCHERER, Matthew U. Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. **Harvard Journal of Law & Technology**, v. 29, n. 2, 2016. P. 359.

28 Ibidem, p. 359.

29 GASKE, Matthew R. Regulation Priorities for Artificial Intelligence Foundation Models. **Vand. J. Ent. & Tech. L.**, v. 26, p. 1, 2023.

riscos sistêmicos e (3) regras claras³⁰. Segundo Matthew Gaske³¹, o incentivo à inovação (1) não pode ser descartado para a normatização da IA, por motivos geopolíticos e de impactos sociais positivos. Sistemas de IA serão cruciais para estratégia de defesa dos países, e necessariamente esta finalidade tende a ser uma prioridade estatal. Ainda, a IA, enquanto tecnologia de propósito geral, promete revolucionar diversos setores, em especial saúde, educação, indústria e combate às mudanças climáticas.

Ao mesmo tempo, também é essencial a mitigação de riscos (2), em razão do alto potencial danoso da IA, a qual traria riscos sistêmicos em razão de sua aplicação ampla e adaptabilidade. A rapidez operacional e opacidade dos sistemas pode dificultar sua supervisão eficaz; tais qualidades concomitantemente garantem centralidade a esses modelos na sociedade, tornando-os alvos relevantes e vulneráveis para ciberataques. Seu uso generalizado também pode afetar comportamentos humanos de forma inesperada. No âmbito jurídico, já é possível constatar alguns destes efeitos no uso da IA, com novos riscos discriminatórios na tomada de decisões, riscos à confidencialidade de dados, e preocupações sobre algoritmos que funcionam como uma “caixa-preta”, sem que se saiba o caminho tomado para se chegar a determinado resultado³².

Neste cenário em que o incentivo à inovação (1) e a mitigação de riscos (2) não podem ser descartados, resta então a dificuldade de estabelecer regras claras e específicas (3) na regulação da IA. Um “código” – com a pretensão de legislar de forma ampla, abrangente e coesa – pode não ser a forma mais adequada para normatização deste mercado, embora, como se verá abaixo, esse é o caminho pelo qual o Brasil está seguindo.

30 Ibidem.

31 Ibidem.

32 ALQODSI, Enas Mohammed; GURA, Dmitry. High tech and legal challenges: Artificial intelligence-caused damage regulation. **Cogent Social Sciences**, v. 9, n. 2, 24 out. 2023.

3.2 O “modelo de risco” e alternativas regulatórias

Com o aumento do uso da IA tanto pelo setor privado quanto pelo setor público, e em razão dos riscos apresentados como mencionado acima, há certa convergência em torno da adoção de um modelo de regulação baseado no risco, com a utilização de ferramentas importadas de outros setores regulados. A regulação com base no risco possui três categorias de ferramentas: (a) táticas preventivas, incluindo proibições legais, licenciamento e *sandbox* regulatório; (b) análise e mitigação de riscos, envolvendo documentação de processos e relatórios; (c) medidas posteriores à comercialização, como revogação de licenças, registro e monitoramento, e avaliações cíclicas de conformidade e desempenho e modos de segurança³³.

Quando se regula com base no risco, e não no dano, o enfoque é lidar com incertezas futuras, através de prognósticos e prevenção; geralmente, a discussão é mais voltada ao coletivo em detrimento da compensação individualmente considerada. Particularidades são frequentemente desconsideradas e a indenização individual é mais difícil, o que pode levar ao questionamento da legitimidade e eficiência da regulação; há grande regulamentação do processo, e pouca elaboração de critérios materiais de direito³⁴. Apesar da tendência de convergência regulatória com base no risco, há outros modelos que poderiam ser adotados, ou ao menos utilizados de forma híbrida. Nicolas Petit e Jerome de Cooman destacam outros dois modelos pelos quais a IA poderia ser regulada: (a) o modelo “*black letter law*” ou incorporação e (b) o modelo ético ou principiológico³⁵.

33 KAMINSKI, Margot E. Regulating the Risks of AI. **Boston University Law Review**, v. 103, p. 1347, 2023.

34 Ibidem.

35 PETIT, Nicolas; DE COOMAN, Jerome. **Models of Law and Regulation for AI**. Robert Schuman Centre for Advanced Studies Research Paper No. RSCAS, v. 63, 2020.

O modelo de incorporação consiste em verificar como as normas já existentes em determinado ordenamento jurídico podem se aplicar a um sistema de IA, em especial a partir de uma visão setorial. Identificam-se, portanto, quais normas já são aplicáveis aos sistemas de IA, havendo tendência de transversalidade de normas, preponderância da análise teleológica (consideração do objetivo de cada norma) e normatização *ex post*³⁶. Já no modelo ético ou principiológico há preponderância dos princípios sobre as regras, mantendo certa flexibilidade em razão da textura aberta das normas que serão usadas para regular a tecnologia. Diversos países, como o Japão, optaram por esse caminho, apostando na criação de princípios gerais que devem guiar o desenvolvimento e o uso da IA em seus territórios sem, contudo, legislar de forma específica e total pelo menos até que a indústria de IA esteja madura o suficiente, contribuindo com uma melhor compreensão de suas complexidades e especificidades.

3.3 IA aberta como caminho para uma democracia movida a dados

A luz das discussões acima, é preciso refletir sobre o que significa, então, regular a IA a partir de uma perspectiva de fortalecimento da “democracia movida a dados”. Em primeiro lugar, é necessário assumir que sistemas de IA já fazem parte daquilo que podemos chamar de “infraestrutura pública digital” (do inglês *digital public infrastructure* ou DPI) e podem ser explorados por diferentes atores, públicos e privados, para avançar o interesse público (veja-se, por exemplo, o uso de dados abertos para promover o acesso à justiça ou então o combate às mudanças climáticas, em especial a partir do processamento de imagens de satélite de regiões críticas como a Amazônia e o Pantanal). Partindo dessa constatação, resta evidente que a regulação da IA terá um impacto profundo em nossa infraestrutura pública digital e, por isso, deve ser

36 Ibidem.

investigada com cuidado a partir dos pressupostos lançados pelo presente estudo.

A carga regulatória que seria imposta sobre o ecossistema de IA no Brasil caso o PL 2.338/2023 fosse aprovado é parte central desse debate. Segundo levantamento do Instituto de Tecnologia e Sociedade do Rio de Janeiro (ITS Rio), enquanto o *AI Act* da União Europeia prevê 43 obrigações específicas para diferentes atores envolvidos no ciclo de vida da IA (desenvolvedores, distribuidores e implementadores), o PL 2.338, considerando a versão aprovada pelo Senado no final de 2024, prevê 68 obrigações regulatórias, reservando uma parte considerável dessas obrigações para aplicações de alto risco e para aquelas que serão usadas pelo setor público³⁷. Esses dados devem ser lidos ao lado de outra investigação do instituto, que constatou que enquanto na União Europeia se espera que algo em torno de 10 e 18% dos sistemas de IA sejam classificados como de alto risco, no Brasil esse número pode chegar a 35%³⁸. É dizer, uma parcela significativa do ecossistema de IA no país estaria sujeita à mais alta carga regulatória prevista pelo PL.

Mais preocupante ainda é a interface entre o PL 2.338 e o Plano Brasileiro de Inteligência Artificial (PBIA) desenvolvido pelo Ministério da Ciência, Tecnologia e Inovação (MCTI). O PBIA estruturou uma série de ações estratégicas para o desenvolvimento da IA no Brasil, com investimentos bilionários nos próximos anos para sua concretização. Dentre as ações de impacto previstas em seu Anexo 1, por exemplo, estão implementações de IA essenciais para o fortalecimento da infraestrutura pública digital do país e, consequentemente, da democracia brasileira. Veja-se, por exemplo, as propostas de desenvolvimento de sistemas de

37 ITS Rio. **Matriz comparada de obrigações:** PL 2338/23 vs EU AI Act. 23 de junho de 2025. Disponível em https://itsrio.org/wp-content/uploads/2017/01/2025.06.23-Matriz-comparada-de-obrigacoes_-PL-2338_2023-vs.-EU-AI-act_Relatorio_AI-16-jun-1.pdf.

38 ITS Rio. 100IA: PL 2338/23 e a classificação de risco dos usos de IA sob uma perspectiva prática. Outubro de 2024. Disponível em https://itsrio.org/wp-content/uploads/2016/12/Relatorio_100IA.pdf.

IA para auxiliar a fiscalização pela Receita Federal, para a gestão jurídica, para o combate à biopirataria, para otimização de reembolsos na área da saúde, dentre tantas outras. Nada obstante, segundo levantamento do ITS Rio, 54% dessas ações estratégicas serão classificadas como de alto risco pelo PL 2.338, devendo se adequar a algo entre 59 e 68 obrigações regulatórias a depender do uso de IA Generativa para sua implementação³⁹.

O Brasil precisa reconsiderar a rota traçada até então a fim de melhor compatibilizar os objetivos regulatórios perseguidos com a promoção de inovação tecnológica responsável. Parte da solução para o impasse identificado até aqui passa pela inovação aberta – entendida como o processo de desenvolver soluções tecnológicas de forma colaborativa, com base em dados públicos, modelos compartilháveis e infraestrutura acessível –, que deve ser vista como pilar fundamental para a construção de uma democracia movida a dados. A “IA aberta” não é apenas uma alternativa técnica, mas uma estratégia política para democratizar o acesso aos benefícios da IA, permitindo que múltiplos atores – incluindo universidades, cooperativas, governos locais e organizações da sociedade civil – se apropriem da tecnologia de forma criativa, ética e contextualizada. É quando a tecnologia deixa de ser monopólio de grandes plataformas e se torna ferramenta de transformação social que sua potência democrática se revela por completo.

A experiência recente do estado de Goiás, que em 2025 aprovou a primeira lei compreensiva de IA do Brasil, oferece um importante contraponto ao modelo regulatório proposto pelo PL 2.338. Enquanto o projeto federal se concentra majoritariamente na mitigação de riscos e na imposição de obrigações *ex ante*, a lei estadual avança na construção de um ecossistema de inovação voltado ao interesse público, com atenção especial ao fomento à pesquisa aplicada, à educação tecnológica e ao desenvolvimento de soluções locais baseadas em IA. A norma reconhece,

39 ITS Rio. Análise da relação entre o Plano Brasileiro de Inteligência Artificial e o Projeto de Lei para Regular a IA (PL 2338/23): estratégias em conflito. 26 de junho de 2025. Disponível em <https://itsrio.org/wp-content/uploads/2017/01/Relatorio-PBIA-e-regulacao-IA.pdf>.

assim, a importância da tecnodiversidade e da formação de capacidades institucionais distribuídas, abrindo caminho para uma política de IA mais plural, responsiva e orientada ao bem comum.

Essa abordagem revela que é possível (e igualmente desejável) regular a IA com o objetivo de proteger direitos fundamentais sem sufocar a criatividade e o protagonismo local. O desafio do Brasil não é apenas criar salvaguardas contra danos potenciais, mas também garantir que a IA possa ser usada para fortalecer instituições democráticas, ampliar a inclusão produtiva e refletir a diversidade cultural do país. Isso só será possível com uma visão regulatória que valorize a experimentação, apoie a infraestrutura compartilhada e promova ambientes férteis para que uma espécie de “IA achada na rua” – ou seja, apropriada pelas margens – possa florescer como motor de uma democracia orientada por dados e sensível às realidades sociais do século XXI.

4. Conclusão

A análise desenvolvida ao longo deste artigo permitiu compreender como a transição da sociedade em rede para a sociedade algorítmica implicou profundas transformações nos fluxos de dados, nas estruturas de poder e nos valores democráticos. A centralidade dos dados na era digital não apenas redefine a dinâmica econômica e política global, mas também evidencia os riscos de concentração informacional e os desafios da desinformação em escala industrial. Frente a esse cenário, a abertura de dados surge como vetor essencial de fortalecimento democrático, promovendo participação cidadã, accountability e deliberação mais inclusiva, como demonstrado por experiências no Brasil, em Taiwan e outros países do Sul Global.

No entanto, a simples abertura de dados não é suficiente. A construção de uma democracia digital verdadeiramente plural exige a valorização da tecnodiversidade, que pressupõe reconhecer e fomentar infraestruturas técnicas enraizadas em diferentes contextos sociais, culturais e geográficos. O domínio concentrado da infraestrutura com-

putacional por poucos atores do Norte Global revela as desigualdades estruturais que limitam o desenvolvimento autônomo de tecnologias pelo Sul Global. Nesse contexto, a inovação aberta e a IA baseada em princípios de diversidade, acessibilidade e participação comunitária tornam-se estratégias políticas centrais para a redistribuição de poder e o fortalecimento dos pilares democráticos.

Por fim, a regulação da inteligência artificial deve ser encarada não apenas como uma ferramenta de contenção de riscos, mas como instrumento ativo de construção institucional e de fomento à inovação pública. O modelo proposto pelo PL 2.338/2023, ao impor uma elevada carga regulatória sem considerar suficientemente as especificidades do ecossistema brasileiro, pode comprometer os objetivos estratégicos traçados no Plano Brasileiro de IA e, com isso, a própria qualidade da democracia brasileira na era digital. Alternativas como a adotada pelo estado de Goiás apontam para caminhos mais equilibrados, que articulam proteção de direitos e promoção da inovação responsável. A democracia movida a dados, portanto, não é um destino inevitável, mas uma escolha política – e seu êxito dependerá da capacidade de construir um modelo de IA que seja, ao mesmo tempo, aberto, inclusivo e sensível às pluralidades que constituem nossa sociedade.

João Victor Archegas · Professor de Direito na FAE e Coordenador no Instituto de Tecnologia e Sociedade do Rio de Janeiro (ITS Rio). Mestre e Bacharel em Direito pela Universidade Federal do Paraná (UFPR). *Master of Laws* pela Harvard Law School, onde foi Gammon Fellow de excelência acadêmica. E-mail para contato: j.archegas@itsrio.org.

Diana de Souza Fernandes · Advogada. Graduada pela Universidade Federal do Paraná (UFPR). Pós-graduada em Direito Corporativo pela Pontifícia Universidade Católica do Paraná (PUC-PR). Mestrado Profissional em Direito Civil e Processual Civil pela Fundação Getúlio Vargas (FGV). E-mail para contato: dianafernandesadv@gmail.com.

■ A FUNDAÇÃO KONRAD ADENAUER é uma fundação política da República Federal da Alemanha que, naquele país e no plano internacional, vem trabalhando em prol dos direitos humanos, da democracia representativa, do Estado de Direito, da economia social de mercado, da justiça social e do desenvolvimento sustentável.

Os principais campos de atuação da FUNDAÇÃO KONRAD ADENAUER são a formação política, o desenvolvimento de pesquisas aplicadas, o incentivo à participação política e social e a colaboração com as organizações civis e os meios de comunicação.

A FUNDAÇÃO KONRAD ADENAUER está no Brasil desde 1969 e atualmente realiza seu programa de cooperação internacional por meio da Representação no Brasil, no Rio de Janeiro, trabalhando em iniciativas próprias e em cooperação com parceiros locais. Com suas publicações, a FUNDAÇÃO KONRAD ADENAUER pretende contribuir para a ampliação do debate público sobre temas de importância nacional e internacional.

■ Os *Cadernos Adenauer* versam sobre temas de interesse público, relacionados ao desenvolvimento de uma sociedade democrática.

Privilegiam-se artigos que abarcam temas variados nos campos da política, da situação social, da economia, das relações internacionais e do direito.

As opiniões externadas nas contribuições desta série são de exclusiva responsabilidade de seus autores.



adenauer-brasil@kas.de
www.kas.de/brasil