

internet& sociedade

volume 5 / número 1 / jul 2024

INTERNETLAB



internet & sociedade

A Revista Internet & Sociedade é uma publicação semestral organizada pelo InternetLab, centro independente de pesquisa em direito, políticas públicas e tecnologia localizado em São Paulo (SP). Nosso objetivo é reunir insumos, evidências e argumentos que aprofundem o pensamento crítico em torno de diferentes aspectos sociais, econômicos, políticos e regulatórios envolvendo mídias digitais e tecnologias de comunicação e informação; e, assim, avançar debates acadêmicos e abordar as múltiplas dimensões entre internet e sociedade.



INTERNET &
SOCIEDADE, V.5, N.1
– 2024, 1º SEMESTRE.

Imagem da capa
Joana Resek



Este trabalho está licenciado sob a
Licença Atribuição-Compartilhada 4.0
Internacional Creative Commons.

This work is licensed under a
Creative Commons Attribution-ShareAlike
4.0 International License.

Conselho Editorial

Dalton Lopes Martins
FACULDADE DE CIÊNCIA DA INFORMAÇÃO, UNB

Elias Duarte Jr.
DEPARTAMENTO DE INFORMÁTICA, UFPR

Gisele Craveiro
ESCOLA DE ARTES, CIÊNCIAS E HUMANIDADES, USP

Giselle Beiguelmann
FACULDADE DE ARQUITETURA E URBANISMO, USP

Graciela Nathanson
FACULDADE DE COMUNICAÇÃO, UFBA

José Roberto Xavier
FACULDADE NACIONAL DE DIREITO, UFRJ

Jussara Marques de Almeida
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO, UFMG

Maíra Rocha Machado
ESCOLA DE DIREITO, FGV-SP

Marcelo Thompson
FACULDADE DE DIREITO, UNIVERSIDADE DE HONG KONG

Rogério Christofolletti
DEPARTAMENTO DE JORNALISMO, UFSC

Virgílio Afonso da Silva
FACULDADE DE DIREITO, USP

Editores

Francisco Brito
Cruz e Mariana
Valente

Editor executivo
Rafael C. Sampaio

Comunicação
João Vitor Araújo

Revisão
Vitor Santos
Vilanova

Apoio:
Luisa Brito

Identidade visual,
projeto gráfico
Polar.Itda

Diagramação:
Joana Resek

Site
Atonal Studio

Periodicidade
Semestral

Idiomas em que são
aceitas publicações
Português e Inglês

ISSN
2763-5244

Autor corporativo
InternetLab -
Av. Ipiranga, 344.
São Paulo (SP), Bra-
sil. CEP 01046-010.

revista.internetlab.org.br
contato@revista.internetlab.org.br

**CARTA
DOS
EDITORES**

A inteligência artificial (IA) enquanto campo de estudos e aplicações, possui uma trajetória de várias décadas, mas ganhou notoriedade com o lançamento de ferramentas como ChatGPT e Midjourney no ano de 2022. Esses sistemas de IA possuem a capacidade de gerar textos, imagens, sons e vídeos comparáveis à produção humana. E, de forma inédita, tal capacidade de produção tornou-se pública de forma fácil e acessível para boa parte dos usuários de internet.

Essas aplicações fomentaram discussões sobre os impactos diretos de IAs em diversas profissões e sobre as possíveis transformações estruturais em diversos setores da sociedade. Além disso, reacenderam o debate acerca de sistemas de IAs que já estavam em pleno uso, como aqueles presentes em plataformas de redes sociais digitais, bem como na formulação e implementação de políticas públicas.

No contexto dos impactos da IA, torna-se imprescindível discutir o que caracteriza um sistema de IA, dado que é um conceito em evolução e ainda desprovido de consenso. A Organização para a Cooperação e Desenvolvimento Econômico (OCDE) define sistema de IA como: “[...] um sistema baseado em máquina que, para fins explícitos ou implícitos, deduz, a partir da entrada que recebe, como gerar saídas como previsões, conteúdo, recomendações ou decisões que podem influenciar ambientes físicos ou virtuais” (OCDE, 2024).

Archegas e Maia (2020) sugerem a análise de subconceitos que delimitam e compõem a complexidade do conceito de IA, como estratégia para sua compreensão. Os autores identificam dois subcampos principais. O primeiro é *machine learning*, que se concentra no emprego de dados e algoritmos para simular o processo de aprendizagem humano, permitindo que os sistemas aprimorem, de forma contínua, sua precisão ao responder questões ou solucionar problemas. O segundo subcampo é o *deep learning*, no qual sistemas de IA aprendem a realizar tarefas mais complexas por meio de redes neurais (Archegas y Maia, 2020).

Embora essas abordagens apresentem distinções, é relevante ressaltar que esses sistemas não são autônomos nem dotados de racionalidade para operar sem um treinamento prévio. Conforme argumenta Crawford (2022), os sistemas de IA dependem de contextos políticos e sociais e são desenvolvidos a partir de recursos naturais, logística, infraestrutura e materiais. Assim, estão relacionados a dinâmicas de poder. Dessa forma, discute-se a ideia de que as tecnologias não são neutras, nem podem ser reduzidas a meros instrumentos técnicos. Como defende Hui (2020), “[...] a tecnologia não é antropologicamente universal; seu funcionamento é assegurado e limitado por cosmologias particulares que vão além da mera funcionalidade e da utilidade” (Hui, 2020, p. 15). Assim, não se pode falar de uma tecnologia única, mas de uma pluralidade de *cosmotécnicas*.

**CARTA
DOS
EDITORES**

Por sua vez, na pesquisa, já é possível visualizar as potenciais transformações que acontecerão pela influência da inteligência artificial generativa, como a forma como procuramos material acadêmico e também como lemos, resumimos e lidamos com esses e outros conteúdos pela rede. Igualmente, os chamados grandes modelos de linguagem, como ChatGPT, são treinados em diferentes linguagens de programação, capacitando ainda mais o movimento de ter habilidades em linguagens como R e Python para a análise de dados. Eles ainda são equipados para serem competentes analistas de dados sejam eles quantitativos sejam qualitativos, podendo inclusive responder a comandos diretos sem a necessidade de nenhum conhecimento na área. Consequentemente, também são capazes de gerar inúmeras visualizações de dados, como quadros, tabelas e diversos tipos de gráficos e grafos. E como modelos de linguagem, eles são especialmente aptos para escrever ou revisar os nossos próprios textos. As consequências podem ser diversas, como problemas com plágio, restrições metodológicas, limite à inovação e originalidade, assim como possíveis efeitos negativos na integridade da pesquisa, já que tais modelos são proprietários, fechados. Paradoxalmente, poderemos fazer tarefas mais complexas usando tais modelos ou outras soluções baseadas nos mesmos (SAMPAIO et al, 2024). Algumas dessas questões acabam em maior ou menor medida sendo discutidas pelos artigos presentes neste dossiê.

Pela grande oferta de artigos apresentados, este dossiê será dividido em duas edições. Após a análise editorial e da avaliação dos(as) pareceristas, 10 artigos, incluindo uma produção artística, são publicados neste dossiê, o que confere a excelente qualidade das submissões realizadas. São artigos que propiciam a reflexão sobre sistemas de IA e regulação, governança, riscos e desafios nas mais diversas áreas, como o mercado do trabalho e as campanhas eleitorais. Bem como convidam à imperiosa discussão sobre a necessidade de uma perspectiva interdisciplinar para compreender os usos e impactos de sistemas de IA na sociedade.

Neste contexto, o que estamos assistindo é um momento histórico no qual organizações e corporações estão acumulando e estruturando dados de todos os tipos, sobre todas as coisas e fenômenos sociais, naturais ou sintéticos. Esta coleta silenciosa e onipresente culminará, nos próximos anos, em modelos automatizados cada vez mais robustos e mais conectados ao comportamento humano. Ao mesmo tempo, o avanço na computação quântica, ainda em fase experimental, representará uma guinada na capacidade desses sistemas, estabelecendo um novo patamar. Tudo isso está intimamente conectado a novas estruturas de poder tendo em vista a tendência de enraizamento dessas ferramentas na cultura, política, economia e nos diversos outros campos de ação humana. Diante disso, o Estado tem o papel fundamental de estabelecer normatizações e políticas públicas capazes de dar conta de dois movimentos concomitantes: por um lado, fomentar infraestrutura e know how para

**CARTA
DOS
EDITORES**

viabilizar esses sistemas do ponto de vista operacional com foco em seu caráter inclusivo. Países que não investirem em infraestrutura de transporte de dados (backbones, backhauls, 5G e 6G), armazenamento e processamento (especialmente datacenters) enfrentarão gargalos técnicos e sofrerão com forte dependência geopolítica, acirrando disparidades. Por outro lado, garantir a proteção de direitos, combater violações e canalizar esses avanços na forma de benefícios coletivos é uma outra obrigação do Estado. Neste sentido, a regulação de IA se torna uma premissa obrigatória, não apenas uma opção facultativa. Requer leis e normatizações com sólidos princípios baseados no interesse público, com claras responsabilidades atribuídas e limites delimitados para grandes players, incluindo o próprio Estado. A busca deve ser por um ecossistema para IA que esteja atento aos riscos e aberto à inovação. Regulações genéricas e mal desenhadas podem ser extremamente danosas pois tendem a aumentar a concentração de poder, ampliar desigualdades, potencializando violações sob o respaldo da lei.

Rafael Sampaio

Maria Alejandra Nicolás

Sivaldo Pereira Silva

Os editores

Referências bibliográficas

- Arhegas, J. V., & Maia, G. (2022). O que é inteligência artificial (IA)? Análise em três atos de um conceito em desenvolvimento. *Cadernos Adenauer*, 23(2), 14-17.
- Crawford, K. (2022). *Atlas de inteligencia artificial: poder, política y costos planetarios*. Buenos Aires: Fondo de Cultura Económica (FCE).
- Hui, Y. (2020). *Tecnodiversidade*. São Paulo: Ubu Editora.
- OCDE. (2024). Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/0449. Disponible en: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
- Sampaio, R. C., Nicolás, M. A., Junquillo, T. A., Silva, L. R. L., Freitas, C. S. D., Telles, M., TEIXEIRA, J.; ESCÓSSIA, F.; & Santos, L. C. D. (2024). ChatGPT e outras IAs transformarão a pesquisa científica: reflexões sobre seus usos. *Revista de Sociologia e Política*, 32, e008.

SUMÁRIO

- ↘ p. 8
Afiml, o que é
privacidade?
Um panorama
histórico do direito
à privacidade
no ordenamento
constitucional
brasileiro
Isadora Valadares
Assunção e
Silvio Fernando
Janson
- ↘ p. 33
Análise de podcast
pela perspectiva
metodológica da
roleta interseccional
Carolina Salles
Carvalho e João
Alexandre Peschanski
- ↘ p. 56
Queering ai:
uma epistemologia
queer para a ia com
uma perspectiva
de maioria global
Umut Pajaro
Velasquez
- ↘ p. 71
A Plataformização das
Escolas Brasileiras:
impactos à educação e
proteção de dados de
crianças e adolescentes
Anderson Batista,
Eduardo Ariento e
Natália Ribeiro
- ↘ p. 94
Erisictão Já Invadiu
o Bosque de
Deméter: análise
dos riscos sociais do
colapso de IAs por
retroalimentação
Pedro Khauaja
- ↘ p. 108
Gênero e
Discriminação
Algorítmica: o papel da
governança corporativa
para mitigar vieses
comportamentais em
inteligência artificial
Débora Schwartz
- ↘ p. 126
Opacidade algorítmica,
segregação epistêmica
e alterações no
conceito de trabalho
Catarina Spagnol
- ↘ p. 148
Inteligência artificial e
desigualdade social:
o impacto do
colonialismo digital
nas políticas públicas
Letícia Lé Oliveira
- ↘ p. 168
Participação Cidadã
e Regulação da
Inteligência Artificial:
Análise da Composição
Setorial das Audiências
Públicas na CJSUBIA
Juan Manuel García
e Jamila Venturini
- ↘ p. 194
A imagem na
inteligência artificial
Joana Resek

ARTIGO

Afinal, o que é privacidade? Um panorama histórico do direito à privacidade no ordenamento constitucional brasileiro

Isadora Valadares Assunção

isadoravaladares@usp.br

Graduanda em Direito pela
Universidade de São Paulo.

Silvio Fernando Janson

sfjanson@usp.br

Bacharel e Mestre em Economia
pela FEA/USP. Licenciado em Música
e Doutor em Artes pelo Instituto
de Artes da UNESP. Bacharelado
em Direito pela FD/USP.

Palavras-chave

Privacidade

Constituição

Evolução histórica

Proteção de dados

Resumo

A privacidade é um conceito jurídico cujas discussões remontam a invenções tecnológicas como a fotografia e as escutas ambientais. No Brasil, suas origens históricas recebem pouca atenção acadêmica, o que leva a imprecisões teóricas na interpretação atual dos dispositivos constitucionais e legislativos concernentes a ela. Nesse sentido, propõe-se uma revisão narrativa de literatura, junto à análise documental dos anais da Assembleia Constituinte, para evidenciar a evolução temática do direito à privacidade no ordenamento brasileiro. Para tal, foi analisado o contexto internacional a partir do qual se inseriu o tema na Assembleia Nacional Constituinte e as discussões sobre a privacidade durante esta. Por fim, examinou-se o panorama atual da privacidade a nível nacional a partir do texto promulgado da Constituição Federal de 1988.

Ultimately, what is privacy? A historical panorama of the right to privacy in the Brazilian constitutional order

Keywords

Privacy

Constitution

Historical evolution

Data protection

Abstract

Privacy is a legal concept with discussions that can be traced back to technological inventions such as photography and electronic eavesdropping. In Brazil, privacy's historic roots receive little attention on the academia, which leads to theoretical imprecisions in the current interpretation of the constitutional and legislative provisions concerning it. In that sense, a narrative literature review and the documental analysis of the documents of the National Constituent Assembly is proposed, aiming to demonstrate the thematic evolution of the right to privacy in the Brazilian legal order. In order to do so, it was analyzed the international context in which the theme was inserted on the debates of the Constituent National Assembly and also the discussions about privacy during it. Finally, it was examined the current privacy panorama at the national level based on the enacted text of the 1988 Federal Constitution.

1. Introdução

A hermenêutica de conceitos jurídicos inseridos em contextos de alto dinamismo revela elevada complexidade. Nesse sentido, a metáfora de um romance em cadeia ilustra que o jurista interpreta o direito a partir de um panorama - o conjunto de decisões, as elaborações legislativas, entre outros - que restringe seu espaço de criatividade, mas não o elimina. Pelo contrário, tal criatividade deve ser exercida em direção à melhor construção do romance, tida em seu caráter contínuo e evolutivo¹. Um destes conceitos cuja é a privacidade, cuja evolução interpretativa se destaca em suas nuances, avanços e retrocessos.

Nesse sentido, observa-se que as novas tecnologias aparentemente surgem para trazer mais satisfação, eficiência e tranquilidade para a humanidade. Porém, com o advento da revolução na informática, talvez o direito mais importante possuído pelo ser humano esteja sob ataque. Câmeras de vigilância, o rastreamento da navegação do usuário para criar perfis comportamentais e direcionar anúncios personalizados para ele², sistemas computadorizados que armazenam dados pessoais eliminando o direito ao esquecimento³, tudo isso já são fatos.

No Brasil, com o argumento de proteger a sociedade (explorando o medo da população, semelhantemente ao que ocorre nos EUA justificando-se a vigilância estatal pelo suposto combate ao terrorismo⁴), proliferaram os mecanismos de vigilância, a exemplo dos radares com câmeras espalhados por todas as principais cidades exigindo redução de velocidade e que, muitas vezes, são integrados com sistemas de IA relacionados ao monitoramento criminal, como o Detecta, em SP⁵. Outro exemplo na direção da ubiquidade da vigilância estatal vem da China, onde aparatos de vigilância, como câmeras de reconhecimento facial,

monitoramento da atividade em redes sociais, etc., geram scores de avaliação social no melhor estilo ficção científica, sendo também usados para perseguições étnicas⁶. Os exemplos são diversos e não param de crescer. O Direito, como sempre, corre atrás da realidade tentando adequar-se.

É nesse contexto que o presente artigo visa, em caráter de análise documental, analisar os debates da Assembleia Nacional Constituinte sobre o direito à privacidade e, a partir de uma análise narrativa, evidenciar a evolução de tal direito no ordenamento brasileiro. Assim, não se busca apenas um exame detido das origens genéticas de tal conceito, mas destacar, de maneira historicamente informada, as finalidades de sua tutela, para que se possa, a partir disso, fazer face aos novos desafios advindos do impacto das novas tecnologias na sociedade.

2. O direito à privacidade no campo constitucional

2.1. O contexto internacional do direito à privacidade e a Constituinte

Ainda no século XIX, a discussão sobre o direito à privacidade surgiu nos Estados Unidos frente ao avanço dos meios de comunicação. Nesse contexto, Warren e Brandeis publicaram, em 1890, o texto “*The Right to Privacy*”, no qual estabeleceram o direito à privacidade como o “direito de ser deixado sozinho” (*the right to be let alone*). Tal direito teria uma dimensão negativa,

já que ofereceria proteção contra intromissões não desejadas sobre pensamentos, sentimentos e emoções de um indivíduo⁷.

Apesar deste ser o marco inicial da discussão sobre o direito à privacidade, persistia, ainda, a necessidade de um maior aprofundamento e delimitação de tal direito, com sua diferenciação de outros, como o direito à liberdade⁸. Assim, durante o século XX, diversos autores empenharam-se no estabelecimento de um direito único, cujo centro gravitava, ora na informação ora no acesso, entre outros, como, respectivamente, em Westin e Gavison.

Em 1967, Alan Westin definiu a privacidade como “a reivindicação de indivíduos, grupos ou instituições para determinar por eles mesmos quando, como e até qual ponto informação sobre eles é comunicada com outros”⁹. Em 1980, Ruth Gavison propõe, por outro lado, que a privacidade é uma limitação do acesso de terceiros a um indivíduo, envolvendo a dimensão do conhecimento sobre alguém e do acesso físico a ele, além da medida em que alguém é objeto da atenção de outrem¹⁰.

Concomitantemente, na Europa, as discussões sobre o direito à privacidade eram permeadas por uma preocupação em evitar abusos estatais advindos da coleta de informações individuais, como os ocorridos na Segunda Guerra Mundial¹¹. Na Holanda, por exemplo, em 1938, um sistema de registro populacional foi estabelecido para simplificar a administração pública, seguindo o cidadão “do berço ao túmulo”¹² (*from cradle to grave*). Entretanto, com a invasão nazista em 1940, foi usado para identificar a população judia e cigana, levando ao extermínio dessas¹³.

Assim, enquanto nos Estados Unidos as discussões sobre a privacidade centravam-se na dimensão individual do direito à privacidade, na Europa o viés político-social estava em destaque, ainda mais com o desenvolvimento da informática possibilitando o processamento de dados em grande escala¹⁴. Nesse sentido, o modelo

européu de consideração da privacidade foi, em grande parte, direcionado por uma decisão do Tribunal Constitucional Alemão em 1983.

Em 1977, a Alemanha aprovou uma legislação nacional de proteção de dados¹⁵. Com a Lei do Censo Populacional em 1982, previa-se a realização de um censo cujos dados seriam compartilhados com os demais órgãos públicos para fins não estatísticos. Tal lei sofreu intenso escrutínio público, culminando na decisão do Tribunal Constitucional em 1983 que determinou, entre outras disposições, o direito à autodeterminação informativa como desdobramento do direito geral à personalidade. Estabeleceu também o direito de acesso e de retificação dos dados, além do princípio da finalidade. Adicionou-se, então, uma dimensão positiva à privacidade, tradicionalmente entendida como um direito à exclusão de terceiros¹⁶.

Com o reconhecimento da importância do direito à privacidade tanto na dimensão individual quanto coletiva, este passou a ser assegurado também constitucionalmente em países europeus, entre os quais destacam-se Portugal e a Espanha, sendo que aquele influenciou sobremaneira as discussões da Assembleia Nacional Constituinte brasileira sobre o assunto. Na Constituição Portuguesa de 1976, estabelece-se no art. 26 (1) que:

1. A todos são reconhecidos os direitos à identidade pessoal, ao desenvolvimento da personalidade, à capacidade civil, à cidadania, ao bom nome e reputação, à imagem, à palavra, à **reserva da intimidade da vida privada e familiar** e à proteção legal contra quaisquer formas de discriminação. [...] (grifos nossos)

Estabelece também, no art. 35, diretrizes para a utilização da informática, garantindo o direito de acesso a dados informatizados, a proibição de tratamento informatizado de dados

concernentes a opiniões políticas e filosóficas ou à vida privada, entre outros, salvo se para fins estatísticos não individualmente identificáveis. Ainda, proíbe o estabelecimento de um número nacional único aos cidadãos e o acesso aos dados pessoais de terceiros, salvo disposição legal.

Verifica-se, então, uma progressiva ampliação do direito à privacidade, com sua constitucionalização em certos países europeus e a atribuição de uma perspectiva positiva e social ele, ainda que de forma pouco consolidada. Nesse contexto, surge o debate da questão na Assembleia Nacional Constituinte (ANC), contendo disposições sobre quatro eixos centrais da privacidade, como será evidenciado: a intimidade, o sigilo, a dimensão física e a informacional.

2.2. A Comissão Afonso Arinos e o direito à privacidade

O anteprojeto elaborado pela Comissão Afonso Arinos não foi adotado como base para os trabalhos da ANC, mas assumidamente foi utilizado pelos constituintes como inspiração¹⁸.

Apresentava, no seu art. 17, *in verbis*, uma série de determinações sobre a dimensão informacional da privacidade, indo ao encontro das discussões europeias sobre o tema:

Art. 17 – Todos têm direito de acesso às referências e informações a seu respeito, registradas por entidades públicas ou particulares, podendo exigir a retificação de dados, com sua atualização e supressão dos incorretos, mediante procedimento judicial sigiloso.

§ 1º – É vedado o registro informático sobre convicções pessoais, atividades políticas ou

vida privada, ressalvado o processamento de dados não identificados para fins estatísticos.

§ 2º – A lesão decorrente do lançamento ou da utilização de registros falsos gera a responsabilidade civil, penal e administrativa.¹⁹

Englobava, então, tanto o direito de acesso quanto de retificação, assim como na decisão do Tribunal Constitucional Alemão. Refletia também, no § 1º, um receio sobre o processamento de dados pessoais privados armazenados em registros informáticos. Assim como na Constituição de Portugal, apenas se permitia tal registro para fins estatísticos nos quais os dados não fossem individualmente identificáveis. Previa-se, além disso, a responsabilização civil, penal e administrativa sobre registros falsos, além de se assegurar, no art. 48, um instrumento processual para a tutela dos direitos supracitados: o *habeas data*.

Ainda, abordava-se a dimensão física da privacidade, refletida na concepção tradicional norte-americana sobre o tema, com o direito à inviolabilidade da casa e o direito à moradia sendo assegurados, respectivamente, no art. 37 e no art. 368.

Por sua vez, o sigilo da correspondência e das comunicações é garantido pelo art. 38, com exceção para atividades de investigação criminal autorizadas judicialmente.

O texto do anteprojeto Afonso Arinos influenciou sobremaneira as discussões da Subcomissão dos Direitos e Garantias Individuais, tanto é que Cândido Mendes, o relator da comissão que discorreu sobre o tema na Comissão Afonso Arinos, fez extenso discurso na Assembleia Constituinte, como convidado, sobre a questão da privacidade.

2.3. Subcomissão dos Direitos e Garantias Individuais e o direito à privacidade

Na Subcomissão dos Direitos e Garantias Individuais, as pontuações que dizem respeito ao direito à privacidade foram ensejadas, principalmente, pelo discurso de convidados da sociedade civil, em especial: Orlando Coutinho, Cândido Mendes e Arthur Pereira.

Em sua participação na 7ª reunião da subcomissão, Orlando Coutinho, representante da Confederação Geral dos Trabalhadores, enfatizou a preocupação com “o direito à privacidade, à liberdade de opinião e consciência”²⁰ frente a registros informáticos de informações pessoais. Exemplificou que estes poderiam ser utilizados por empresas para criar listas negras de trabalhadores que externassem certas opiniões políticas, impedindo-os de serem contratados. Ademais, afirmou que erros eram comuns em tais registros, levando a implicações práticas como a negação do crédito pessoal pelo Serviço de Proteção ao Crédito²¹.

Assim, apresentou como da mais alta relevância que a responsabilidade civil, penal e administrativa por informações falsas fosse estabelecida, além do direito de acesso e retificação aos bancos de dados. Refletia, então, nesse ponto, posição semelhante àquela presente no anteprojeto da Comissão Afonso Arinos.

Quanto ao sigilo, discordou da Comissão supracitada, defendendo que esse não deveria admitir violações nem mesmo sob ordem judicial, citando caso de escutas telefônicas pelas autoridades policiais de conversas entre advogados paulistas e seus clientes²².

Já na 8ª reunião, Cândido Mendes, representante da Comissão Brasileira de Justiça e Paz, fez

explanação sobre os novos direitos humanos decorrentes do progresso social e tecnológico. Observou, nesse sentido, que: “estamos diante do rapto da alma, os registros eletrônicos e a informática permitem, na instalação desse crime moderno, a ditadura da versão e da ficha sobre a verdade da pessoa”²³. Frente a isso, trouxe, como propostas, de forma alinhada à Constituição Portuguesa e ao anteprojeto Afonso Arinos: (i) o estabelecimento do habeas data, que será posteriormente detalhado; (ii) a inviolabilidade da casa, ameaçada pelas teleobjetivas e por sensores eletrônicos; (iii) a reserva da intimidade²⁴.

Posteriormente, na 10ª reunião, o representante da União Brasileira de Informática Pública, Arthur Pereira Nunes, retomou ao tema das garantias individuais frente à informática. Em relação ao direito à privacidade, defendeu: (i) o direito de acesso e de retificação, (ii) a proibição do tratamento de informações pessoais em registros informatizados, (iii) a responsabilidade dos administradores dos registros, (iv) a garantia do uso da informação apenas para a finalidade para a qual foi coletada e (v) a proibição da atribuição de um número único ao cidadão, nos moldes da Constituição Portuguesa. Justificou a necessidade da positivação desses direitos como advinda do progresso crescente da informática e dos efeitos danosos que este poderia trazer à sociedade, em especial no século XXI²⁵.

A respeito da última pontuação de tal convidado, o Constituinte Eliel Rodrigues, do PMDB, observou que, apesar de já existirem no Brasil o CPF e o RG, havia uma tendência mundial de centralização das informações a respeito do cidadão, com o processamento por supercomputadores permitindo uma maior ingerência estatal, que implicaria na violação à privacidade do indivíduo²⁶. Em resposta, Arthur Pereira pontuou que de fato o número único apresentava novos riscos frente ao avanço informático, fomentando a perseguição política nos moldes do que ocorreu na ditadura, além de que traria a

possibilidade de erros em escala nos registros públicos centralizados²⁷.

O presidente da Subcomissão, Antônio Mariz, do PMDB, reforçou os riscos políticos advindos do registro de dados individuais sobre os cidadãos, excluindo informações anonimizadas para fins estatísticos. Nesse sentido, afirmou que:

É evidente que esse arquivo de dados pessoais tem um sentido eminentemente político. Esse tipo de registro constitui uma ameaça permanente à sociedade. É claro que nos momentos de recrudescimento - praça aos céus não ocorra - do Estado autoritário esses elementos teriam extrema eficiência na repressão. É claro, se estamos elaborando uma Constituição democrática, não podemos admitir, em hipótese alguma, o registro de dados privados. Que interesse pode haver, do ponto de vista social, nesses dados, a menos que se admita essa discriminação em virtude das particularidades dos indivíduos?²⁸

Percebe-se, então, que a tônica da Subcomissão girou em torno da preocupação com a possibilidade de limitações das garantias individuais em um cenário de desenvolvimento informático. Apesar do reconhecimento da possibilidade de interesses sociais como a realização de pesquisas estatísticas para a formulação de políticas públicas, havia um receio generalizado da perseguição política que poderia se originar do tratamento de dados.

Tais preocupações são expressas no anteprojeto da Subcomissão, inspirado na Constituição Portuguesa e no anteprojeto Afonso Arinos. No inciso VI, previa-se a proteção da dignidade da pessoa humana, de sua honra, privacidade, reputação e imagem pública, além de se assegurar o direito de resposta. Por sua vez, no inciso VIII, *in verbis*, previa-se a proteção da dimensão informacional da privacidade:

VIII - o conhecimento das informações e referências pessoais, e do fim a que elas se destinam, registradas por entidades públicas ou particulares; é vedado o registro de convicções pessoais, atividades político-partidárias, ou acerca da vida privada; é permitido, para fins estatísticos, o registro de dados não identificáveis individualmente; é assegurada a supressão ou retificação de dados incorretos, mediante procedimento administrativo ou judicial, de caráter sigiloso; responde civil, penal e administrativamente todo aquele que determine, realize ou se utilize de registro de dados pessoais incorretos ou falsos; **dar-se-á habeas data ao legítimo interessado, para assegurar-lhe o direito de conhecer as informações e referências pessoais existentes a seu respeito;**²⁹ (grifos nossos)

Percebe-se, portanto, que, nos debates da Subcomissão de Direitos e Garantias Individuais, a preocupação central remetia a que se fosse assegurada uma esfera de liberdade ao indivíduo frente à crescente capacidade estatal e particular de processamento de dados, inclusive por meio da previsão de mecanismos de acesso e retificação dos dados pessoais. Ademais, em algumas situações tidas como especialmente propícias a abusos, como o registro de convicções pessoais ou de atividades político-partidárias, previa-se uma vedação geral, pelo que se percebe um escalonamento das proteções constitucionais à privacidade, inclusive em sua dimensão informacional.

2.4. Comissão da Soberania e dos Direitos e Garantias do Homem e da Mulher e o direito à privacidade

Ao passar para o âmbito da Comissão da Soberania e dos Direitos e Garantias do Homem e da Mulher, as disposições sobre a privacidade assumiram novos contornos. Na nova redação do substitutivo, no inciso VII, intitulado “Da privacidade”, eram garantidas, nas alíneas “a” até “g”, a privacidade da vida particular e familiar, além da casa, tida como inviolável a menos que sob ordens judiciais. Protegia-se também o sigilo das correspondências e das comunicações, salvo autorização judicial segundo os casos previstos em lei e para fins de investigação criminal. Ainda, a imagem pessoal, a vida íntima e a familiar não poderiam ser divulgados sem autorização, não podendo haver atividades privadas de investigação e prestação de informações sobre esses temas, nem serviços estatais de informação sobre eles. No campo policial e militar, os serviços de informação eram permitidos apenas no que se referissem à delinquência ou a atividades violentas de subversão dos fundamentos constitucionais³⁰.

Após a apresentação do parecer do relator, intensas discussões seguiram-se. O Constituinte João Menezes, do PMDB-PA, evidenciou a falta de garantia da norma que proibia a divulgação, publicação ou invasão da vida íntima e familiar ou da imagem pessoal sem autorização do interessado. Não haveria garantia uma vez que a norma não define o que é invasão nem estabelece uma consequência para ela. Percebe-se, então, uma forte preocupação com a auto-aplicabilidade da lei. Menezes pontua, como exemplo da falta de garantia, que os pobres

têm sua miséria televisionada sem consentimento, não tendo garantido o direito à privacidade como os demais. Assim, afirma que “ou esse dispositivo tem uma sanção para que seja aplicada e cumprida, ou então ele está aqui como mera fantasia”³¹.

Em seguida, o relator José Paulo Bisol, do PMDB-RS, discordou da posição de que a ausência, no texto constitucional discutido, de sanções levaria a uma falta de aplicabilidade. Isto porque o dispositivo teria implicações penais, frente às tipificações já previstas no Código Penal e aquelas que porventura o legislador viesse a criar, além de poder ensejar a responsabilização civil frente à eventual violação. Sendo assim, teria aplicabilidade prática garantida³².

Ademais, outro centro de discussão sobre o mesmo dispositivo ocorreu devido à pontuação pelo Constituinte Samir Achôa, do PMDB-SP, de que a proibição da invasão da vida privada constituiria uma limitação à liberdade de imprensa, na medida em que noticiar um crime seria uma invasão à privacidade³³.

Em resposta, o relator Bisol afirmou que a privacidade é vinculada à própria dignidade da pessoa humana, sendo essencial para o desenvolvimento da personalidade. Além disso, em sua opinião não haveria cerceamento da liberdade de imprensa, pois quem quisesse poderia consentir, por exemplo, a aparecer na revista “Playboy”³⁴.

É notável que implícito à discordância dos parlamentares está a divergência sobre o caráter individual ou social da privacidade. A concepção de Achôa entende a privacidade como direito individual de menor peso frente a direitos com uma dimensão social evidente, como a liberdade de imprensa. Por outro lado, Bisol entende que a privacidade, ao se relacionar intrinsecamente com a personalidade, traz benefícios sociais, em visão semelhante à da norte-americana Ruth Gavison. Nessa linha, o relator afirma que “A privacidade é mais importante do que a propriedade, é muito mais

importante que o *status*, vinculada à dignidade da pessoa. [...] Esse texto é de uma importância humana, moral e jurídica extraordinária!”³⁵.

O debate sobre o viés social e político da privacidade reaparece também na fala do Constituinte José Genoíno, do PT-SP, que observou que a permissão dos serviços de informação estatais referentes à delinquência e à subversão violenta das bases constitucionais consistiria em uma legitimação de aparatos de vigilância, como o SNI, que violentam os direitos individuais³⁶. Evidencia-se, assim como nas reuniões da Subcomissão, forte influência dos temores em relação à perseguição política que poderia advir de uma ausência de proteção adequada da privacidade, em especial pelo ocorrido na ditadura militar e semelhante às preocupações europeias com o tema.

Após os debates, o anteprojeto da Comissão foi aprovado com alterações mínimas em relação ao novo substitutivo do relator Bisol, sendo consolidada a versão por ele defendida. Assim, apenas modificou-se a inviolabilidade da casa para inviolabilidade da moradia, sem que se realizassem mudanças quanto à auto aplicabilidade das disposições nem se alterando a permissão de serviços de informação sobre delinquência e subversão.

2.5. Comissão de Sistematização e Comissão de Redação: o direito à privacidade

Na Comissão de Sistematização - posteriormente alterada para ser denominada Comissão de Redação - reapareceu a discussão sobre a inviolabilidade do sigilo ser sujeita ou não ao controle jurisdicional, sendo que, na Subcomissão dos Direitos e Garantias Individuais, o convidado Orlando Coutinho já havia defendido

que não o fosse³⁷. Tal posição foi defendida na Comissão de Sistematização com a Emenda Brandão Monteiro, proposta pelo constituinte do PDT-RJ de mesmo nome³⁸.

Do lado discordante, o constituinte Gerson Peres do PDS-PA argumentou que o princípio da inviolabilidade do sigilo já era consagrado na tradição constitucional, mas deveria haver a possibilidade de quebra deste segundo ordem judicial, para possibilitar a adequada persecução penal a depender do caso³⁹. As mesmas razões alegou o constituinte Oswaldo Lima Filho, do PMDB-PE, para reforçar a necessidade do controle jurisdicional, sendo esta a corrente vitoriosa no Projeto de Constituição votado pela Comissão⁴⁰. Sobre a questão, o constituinte Adolfo Oliveira, do PFL-RJ, alinhado à corrente que defendia o controle jurisdicional, ressaltou ainda a completude do texto do relator no que tange à inclusão da inviolabilidade do processamento de dados, devido ao avanço tecnológico⁴¹.

Quanto aos demais temas, não houve discussões substanciais nem na Comissão de Sistematização nem na Comissão de Redação, optando-se, no projeto final desta última, por priorizar-se a sinteticidade⁴² e excluir-se a seção “Da privacidade”, reduzindo os dispositivos nela contidos a incisos esparsos do art. 5º da Constituição. Assim, foram fixados princípios genéricos a serem especificados por legislações infraconstitucionais posteriormente, em detrimento dos avanços e do tratamento abrangente que o tema havia recebido na Subcomissão dos Direitos e das Garantias Individuais, além da Comissão da Soberania e dos Direitos e Garantias do Homem e da Mulher. A privacidade foi, então, abordada nos seguintes incisos, *in verbis*, coincidindo os termos do projeto final promulgado com os do texto vigente atualmente:

Art. 5º Todos são iguais perante a lei, sem distinção de qualquer natureza, garantindo-se aos brasileiros e aos estrangeiros residentes no País a inviolabilidade do direito à vida, à liberdade, à igualdade, à segurança e à propriedade, nos termos seguintes:

X - são invioláveis a intimidade, a vida privada, a honra e a imagem das pessoas, assegurado o direito a indenização pelo dano material ou moral decorrente de sua violação;

XI - a casa é asilo inviolável do indivíduo, ninguém nela podendo penetrar sem consentimento do morador, salvo em caso de flagrante delito ou desastre, ou para prestar socorro, ou, durante o dia, por determinação judicial;

XII - é inviolável o sigilo da correspondência e das comunicações telegráficas, de dados e das comunicações telefônicas, salvo, no último caso, por ordem judicial, nas hipóteses e na forma que a lei estabelecer para fins de investigação criminal ou instrução processual penal;⁴³

Apesar da disparidade entre o texto final e os textos das anteriores subcomissões e comissões, resta saber como os dispositivos vigentes são interpretados atualmente para se assegurar o direito à privacidade no panorama constitucional no século XXI.

3. A privacidade como direito fundamental no século XXI

3.1. A proteção da intimidade e da vida privada como expressão de um direito maior à privacidade

Segundo Doneda, a opção pelo legislador constitucional por distinguir entre “intimidade” e “vida privada” no âmbito do art. 5º, X, CF/88, reflete uma indeterminação terminológica da época⁴⁴. Ademais, reflete também a influência doutrinária da teoria dos círculos concêntricos de Hubmann, também chamada de teoria das esferas⁴⁵.

Para tal teoria, elaborada em meados do século XX, haveria três níveis de proteção do direito geral à personalidade, no qual se incluiria a noção de privacidade. No núcleo mais interno, havia a esfera da intimidade ou do segredo, protegida com a maior intensidade, tanto contra o público quanto a outros indivíduos. Envolvendo-a, havia a esfera privada, a englobar relações cotidianas, e que deveria ser protegida do público em geral. Por fim, como esfera mais externa, havia a esfera pessoal, que se referiria à vida pública e cuja proteção se restringiria aos danos que adviessem das relações com os demais nesta esfera⁴⁶.

Seguindo a distinção feita pelo constituinte, a doutrina brasileira encontrou-se imersa em uma dubiedade terminológica, com esforços infrutíferos de diferenciar a “vida privada” da “intimidade”. Por exemplo, Silva⁴⁷ considera que a vida privada inclui a intimidade, assim

como Lôbo⁴⁸. Por outro lado, Ferraz Júnior as diferencia, estabelecendo que a intimidade é um estar sozinho sem implicações sociais e a vida privada envolve a proteção de formas de convivência nas quais se excluem terceiros de certas relações⁴⁹.

Destarte, Doneda entende que tal distinção não é relevante já que, apesar de cada um dos termos ter um campo semântico diferente, não tem valoração diferente. Ademais, a binariedade dos conceitos não encerra a pluralidade e a complexidade dos problemas relacionados à privacidade⁵⁰. Nesse sentido, é de se notar a pontuação feita pelo jurista norte-americano Daniel Solove de que não é possível conceituar a privacidade em torno de um núcleo único, uma vez que abrange uma ampla gama de situações relacionadas entre si por meio de suas diversas similaridades⁵¹.

Além da falta de utilidade de uma distinção entre ambos os conceitos para sua aplicação prática, é possível observar que, durante grande parte da Assembleia Nacional Constituinte, ambas as noções se situavam na categoria “Da privacidade”, evidenciando-se uma relação entre elas entrevista pelos constituintes. Assim, toma força a inclusão de ambos os conceitos dentro daquele da privacidade, como propõe Doneda⁵², posição também reconhecida jurisprudencialmente⁵³.

Quanto à privacidade, cabe destacar que, com o desenvolvimento das tecnologias da informação, além de sistemas de inteligência artificial, junto à disseminação do uso da *internet*, houve uma verdadeira transformação no conteúdo de tal direito. Se, no contexto da Constituinte, evidenciava-se já uma preocupação com o aspecto informacional da privacidade, em especial devido ao temor da evolução tecnológica, esta era ainda predominantemente entendida como um direito que implicava na exclusão de terceiros - fossem eles agentes privados ou estatais - da esfera privada ou íntima, como na teoria de Hubmann⁵⁴. No entanto, com o avanço

interpretativo, tal dimensão positiva vem sendo reforçada, ao se entender que, no contexto atual, um dos principais vetores de proteção à privacidade é a proteção dos dados pessoais⁵⁵.

A isto se soma um crescente posicionamento doutrinário que aponta a insuficiência de considerar a privacidade como um direito subjetivo para sua efetiva proteção, reconhecendo-se sua dimensão coletiva frente aos seus benefícios políticos e ao fomento às interações sociais que proporciona⁵⁶.

Ademais, devido ao aumento da capacidade inferencial dos sistemas de inteligência artificial aliados à massiva quantidade de dados da indústria de *Big Data*, a privacidade de um depende de que a dos outros seja minimamente preservada⁵⁷. Nesse sentido, Doneda pontua que a definição de Rodotà sobre o direito à privacidade reflete a evolução recente do tema, englobando tanto seu aspecto físico quanto o informacional e o coletivo⁵⁸. Segundo Rodotà, a privacidade é “o direito de manter o controle sobre as próprias informações e de determinar as modalidades de construção da própria esfera privada”⁵⁹.

3.2. O aspecto físico da privacidade: a inviolabilidade da casa

Ainda no século XX, o âmbito físico-espacial da privacidade já era reconhecido, nos EUA, pela doutrina do acesso de Gavison⁶⁰. No Brasil, a proteção constitucional da inviolabilidade da casa se deu com o art. 5º, XI, CF/88, além de ser esta considerada pela jurisprudência como parte do direito à intimidade, como no REsp nº1575681⁶¹.

Com a inviolabilidade da casa sendo a regra, as elaborações doutrinárias e jurisprudenciais sobre o tema giraram em torno do eixo das exceções constitucionais a ela, em especial o

flagrante delito, que permite a entrada no domicílio do indivíduo sem o seu consentimento e sem a necessidade de determinação judicial. Frente à necessidade de delimitação do tema para evitar o esvaziamento do direito fundamental em questão, o STF considerou, no RE nº 603.616, de repercussão geral reconhecida, que:

Recurso extraordinário representativo da controvérsia. Repercussão geral. [...]

6. Fixada a interpretação de que a **entrada forçada em domicílio sem mandado judicial só é lícita, mesmo em período noturno, quando amparada em fundadas razões, devidamente justificadas a posteriori, que indiquem que dentro da casa ocorre situação de flagrante delito**, sob pena de responsabilidade disciplinar, civil e penal do agente ou da autoridade e de nulidade dos atos praticados [...]⁶².

Tal interpretação, com a exigência de fundadas razões indicativas de um flagrante delito, é essencial para a consideração deste nos chamados crimes permanentes, como o tráfico de drogas⁶³. Em decisão monocrática em fevereiro de 2021, no Inq. nº 4781, novos contornos em relação ao flagrante delito por condutas cometidas nas redes sociais foram delineados, na medida em que o ministro Alexandre de Moraes considerou que, como o vídeo postado pelo deputado federal Daniel Silveira - no qual defendia o AI-5 e a substituição dos ministros do STF - permanecia *online*, caracterizava-se a perpetuação do delito e permitia-se sua prisão por crime inafiançável⁶⁴.

A decisão abre espaço para interpretações que legitimem a utilização de postagens de conteúdo em redes sociais e no ambiente virtual como justificativas para a configuração do flagrante delito, o que permitiria, inclusive, a invasão do domicílio do indivíduo sem o seu consentimento ou mandado judicial⁶⁵.

Assim, considerando-se o enorme fluxo de informações contemporaneamente, isto pode gerar um esvaziamento da regra da inviolabilidade da casa, discutida já durante a Comissão Afonso Arinos e posteriormente na Assembleia Nacional Constituinte, ameaçando sua adequada proteção constitucional.

3.3.

A inviolabilidade do sigilo de dados e o direito à privacidade

Com a positivação da inviolabilidade do sigilo de dados no art. 5º, XII, CF/88, uma inovação em relação às Constituições anteriores, novos desafios surgiram em relação aos impactos de tal proteção nos limites da atuação fiscalizadora do Estado⁶⁶. Nesse contexto, Tércio Sampaio Ferraz Júnior que, na época, era Procurador-Geral da Fazenda, elaborou parecer sobre o tema, que foi posteriormente publicado sob o título de “Sigilo de dados: o direito à privacidade e os limites à função fiscalizadora do Estado”.

Neste, Ferraz Júnior estabeleceu que a inviolabilidade do sigilo de dados tinha como conteúdo a faculdade de manter o sigilo para proteger interesses do Estado ou individuais, sendo que neste caso efetivar-se-ia o direito à privacidade, tido como uma “liberdade de negação”⁶⁷, semelhante à definição de Warren e Brandeis. Destarte, “não há um direito fundamental ao sigilo, e sim circunstâncias nas quais o sigilo é instrumental à proteção de um direito fundamental (à privacidade)”⁶⁸, como observam Queiroz e Ponce.

Ademais, além de notar a instrumentalidade do sigilo, o autor nota que há uma simetria no art. 5º, XII, e deduz que o objeto do direito à inviolabilidade do sigilo de dados era a comunicação destes, mas não os dados

em si, enquanto estivessem apenas armazenados⁶⁹. Por fim, quanto aos dados armazenados, o autor considera que, apesar de não serem protegidos pelo art. 5º, inciso XII, poderiam o ser sob o inciso X do mesmo artigo, quando a apreensão dos dados interferisse na privacidade do indivíduo e, conseqüentemente, na sua integridade moral⁷⁰. Assim, seria necessário sopesar, no caso concreto, se a fiscalização estatal geraria uma ingerência na esfera privada ou íntima do cidadão, casos nos quais não deveria ocorrer⁷¹.

A tese de Ferraz Júnior foi seletivamente incorporada na jurisprudência do STF, em especial no que tange à proteção do sigilo da comunicação de dados e não deles em si⁷². No Recurso Extraordinário 418.416/SC, Luciano Hang impetrou recurso extraordinário visando anular sua condenação por crimes tributários, aduzindo que se baseava em prova ilícita que violava o sigilo de dados⁷³. No julgamento de tal recurso, o STF considerou que:

EMENTA: [...] 3. Não há violação do art. 5º. XII, da Constituição que, conforme se acentuou na sentença, não se aplica ao caso, pois não houve "quebra de sigilo das comunicações de dados (interceptação das comunicações), mas sim apreensão de base física na qual se encontravam os dados, mediante prévia e fundamentada decisão judicial". 4. **A proteção a que se refere o art. 5º, XII, da Constituição, é da comunicação 'de dados' e não dos 'dados em si mesmos', ainda quando armazenados em computador.** [...] (RE 418416, Relator(a): SEPÚLVEDA PERTENCE, Tribunal Pleno, julgado em 10/05/2006, DJ 19-12-2006 PP-00037 EMENT VOL-02261-06 PP-01233) (grifos nossos)

É de se notar, primeiramente, que, apesar da distinção entre comunicação de dados e dados armazenados ser de fato realizada por Ferraz Júnior no supracitado artigo, o STF não considerou que o autor observou explicitamente que isso não significa uma ausência de proteção aos dados em si, o que ocorreria pelo art. 5º, X⁷⁴.

Ademais, Queiroz e Ponce notam que, enquanto o reconhecimento da importante relação entre a proteção de dados e a privacidade individual feita pelo autor se mantém atual, alguns pontos do artigo "Sigilo de dados" se mostram já superados. Entre eles: (i) a necessidade de um entendimento da privacidade também como liberdade positiva, com, por exemplo, o acesso e retificação dos dados pessoais; (ii) a superação da distinção entre dados em trânsito e armazenados como vetor para quantizar uma maior ou menor proteção a eles, até porque atualmente, com o aumento da capacidade de armazenamento de dados, estes podem ser mais sensíveis à privacidade do que aqueles⁷⁵.

A atualização doutrinária vem sendo seguida, também, pela jurisprudencial. No HC 168.052, o Relator, Min. Gilmar Mendes, considerou que com "o significativo desenvolvimento das tecnologias da comunicação, do tráfego de dados e dos aparelhos *smartphones*", houve uma mutação constitucional, estando também os dados armazenados protegidos pelo art. 5º, XII, além da proteção advinda do inciso X do mesmo artigo⁷⁶. Assim, observa-se que, com o advento das novas tecnologias, o aspecto informacional da privacidade torna-se central.

Como visto anteriormente, a amplitude da discussão sobre o direito à privacidade durante a Constituinte encerrou-se nos incisos X, XI e XII do art. 5º da CF/88. Contudo, frente às novas tecnologias, a legislação estabelecida pelo Constituinte originário fica ameaçada por ares de não-concretude. Nesse sentido, o *habeas data*, como instrumento de proteção do aspecto informacional da privacidade, torna-se central no debate sobre o tema no Brasil.

3.4. O instituto do habeas data

O termo objeto desta seção está em latim e significa “que tu tenhas os dados; as informações”⁷⁷. Aparece nas propostas apresentadas pelo jurista Márcio Thomaz Bastos ao falar na 10ª reunião em 28/04/87 no texto supra-referenciado: “Dar-se-á *habeas data* ao legítimo interessado para assegurar os direitos tutelados no art. [...]”⁷⁸. Vê-se que a ideia da proposta do juriconsulto era propiciar aos cidadãos ciência dos dados que houvessem sido coletados sobre eles.

Neste momento, cabe reflexão. Tenha-se em mente que o volume de dados coletados pelas empresas de informática e pelos órgãos estatais são de tal monta que a administração desses bancos de dados precisa ficar nas mãos de profissionais de TI amparados por inteligência artificial. Os profissionais trabalham em período integral e as máquinas vinte e quatro horas por dia. Pessoas físicas não dispõem desse tempo para cuidar de seus dados, não tendo capacidade econômica nem técnica para gerenciá-los e acompanhar sua coleta, na maior parte das vezes⁷⁹.

Outro aspecto é a insuficiência do consentimento do usuário para a coleta de dados, central na legislação de proteção de dados atualmente, tanto pela assimetria de informações e poder entre usuário e operadores de dados, quanto pelo fenômeno da fadiga do consentimento⁸⁰ e de que, por muitas vezes, serviços cotidianos são condicionados à aceitação da política de privacidade e dos termos de uso⁸¹.

A questão pendente é: o *habeas data* é viável ou é apenas letra morta? Enquanto tratava-se de número de Registro Geral, Cadastro de Pessoa Física, Certificado de Reservista e congêneres era possível para a pessoa natural acompanhar. No entanto, quem tem condição de saber quais as informações coletadas por celulares, *notebooks*, câmeras onipresentes, gravações digitais de tudo, desde uma simples aula de Direito até

sabe lá Deus o que? Atualmente, até os registros médicos são, muitas vezes, devido à profusão de dados e dos altos custos econômicos de mantê-los em arquivos físicos, armazenados em nuvem, havendo riscos para a segurança das informações⁸² e para a privacidade.

Novamente, o Direito é atropelado pela vida.

Na CF/1988, a abordagem da proteção de dados ficou difusa, como repisado, dependendo da proteção da intimidade e da vida privada assentada no art. 5º, X. No entanto, tal proteção indireta é insuficiente para garantir a adequada segurança e certeza jurídica, além de não englobar dados pessoais, ou seja, que identificam uma pessoa, mas que não são privados, como, por exemplo, o nome, além de dados de processos judiciais que não estão sob sigilo, entre outros⁸³. No entanto, segundo Laura Mendes, pode-se inferir um direito à proteção dos dados pessoais na Constituição Brasileira a partir de uma releitura do *habeas data* a partir de uma dimensão material, com amparo no art. 5º, X⁸⁴.

O *habeas-data* (sic) é previsto constitucionalmente no art. 5º, LXXII, in verbis:

LXXII - conceder-se-á "habeas-data":

a) para assegurar o conhecimento de informações relativas à pessoa do impetrante, constantes de registros ou bancos de dados de entidades governamentais ou de caráter público;

b) para a retificação de dados, quando não se prefira fazê-lo por processo sigiloso, judicial ou administrativo;⁸⁵

O instrumento foi uma inovação do Constituinte e inspirou diversos países latino-americanos, tendo o legislador optado por reconhecer garantias como o direito de acesso e retificação indiretamente, ao fixar um meio de ação processual para sua defesa.⁸⁶ Baseou-se

na doutrina da liberdade informática de Victor Frosini⁸⁷, caracterizada pelo direito de não tornar públicas informações de caráter pessoal e reservado (dimensão negativa), além da faculdade de exercer um direito de acesso e controle sobre os dados pessoais (dimensão positiva)⁸⁸. Tal faculdade implicaria no direito do *habeas data*⁸⁹. Como Tommaso Frosini pontua, o *habeas data* na sociedade contemporânea envolve não apenas uma defesa passiva da autonomia individual como uma participação ativa no circuito das informações⁹⁰.

Sobre tal observação, Mendes nota que, com o julgamento do RE nº 673.707, que tratava do acesso a dados armazenados na Receita Federal, o STF avançou no sentido de reconhecer tal sentido positivo – da autodeterminação informativa – no instrumento do *habeas data*⁹¹. Neste recurso extraordinário, reconheceu-se que “registro de dados deve ser entendido em seu sentido mais amplo, abrangendo tudo que diga respeito ao interessado, seja de modo direto ou indireto, causando-lhe dano ao seu direito de privacidade”⁹². Ademais, o min. Gilmar Mendes fundamentou, em seu voto, que “deste julgado, que pode ser, talvez, o marco inicial de uma vitalização do *habeas data*, numa percepção mais ampla, [...] há já uma reflexão, não no campo procedimental processual, mas também no campo do direito material.”⁹³

Segundo a autora, tal instrumento é essencial para o ecossistema de proteção de dados brasileiro, já que permite a proteção do titular de dados contra abusos do legislador. É de se notar que uma das bases legais da LGPD é a obrigação legal (art. 7º, II e art. 11, II, a, Lei nº 13.709/2018). Portanto, se o Legislativo aprovar uma lei que viole a privacidade, o cidadão estaria desprotegido caso a proteção de dados não fosse garantida constitucionalmente⁹⁴.

Por outro lado, Doneda constata que o *habeas data* é fruto de um contexto histórico específico, com a intenção de se reforçar as bases democráticas do país. Hoje, sua importância

seria eminentemente simbólica, mas não é o instrumento efetivo pelas seguintes razões: (i) por ter escopo limitado ao acesso e à retificação; (ii) por exigir o esgotamento da via administrativa; (iii) por não abranger os riscos de tratamentos desconhecidos pelo usuário⁹⁵. Por fim, como nota Dallari “verificamos que o legislador brasileiro, após a primazia da criação de um modelo que cativa por sua simplicidade, é hoje, em certa medida, seu prisioneiro”⁹⁶. Percebe-se então, que, apesar da CF/88 reconhecer o instituto expressamente, não o incluiu na forma preconizada por Márcio Thomaz Bastos.

Justificar-se-ia, assim, a inclusão, da proteção de dados no rol dos direitos fundamentais, dando guarida constitucional para tal direito, já previsto pela LGPD. Nesse contexto, destacam-se as ações constitucionais que discutiram a proteção de dados pessoais no STF, as ações diretas de inconstitucionalidade nºs 6388, 6389, 6390 e 6393, com tramitação conjunta.

Nelas, impugnava-se a Medida Provisória nº 954/2020, que determinava o compartilhamento dos dados pessoais dos usuários da rede de telefonia com o Instituto Brasileiro de Geografia e Estatística (IBGE), para fins de realização de pesquisas durante a pandemia de Covid-19. Em sede de medida cautelar posteriormente referendada, considerou-se que a autodeterminação informativa decorria da proteção às esferas de personalidade do indivíduo, pelo que o tratamento de dados pessoais deveria observar os limites das “cláusulas constitucionais assecuratórias da liberdade individual (art. 5º, *caput*), da privacidade e do livre desenvolvimento da personalidade (art. 5º, X e XII), sob pena de lesão a esses direitos”. Reconhecendo, assim, a partir do texto constitucional, um direito à proteção de dados pessoais, a Medida Provisória nº 954/2020 teve sua eficácia suspensa, já que não previa de forma adequada as condições e finalidades do tratamento, nem medidas técnicas ou administrativas de proteção dos dados pessoais.

Posteriormente a tal reconhecimento pelo STF, a proteção de dados pessoais foi positivada como direito fundamental pela Emenda nº 115/2022, com a inclusão do art. 5º, LXXIX, CF, que deu guarida constitucional a tal direito, inclusive nos meios digitais, nos termos da lei. Observa-se, portanto, que os moldes de um direito à privacidade e, em sua dimensão positiva, de um direito à proteção de dados pessoais, devem ser colmatados a partir de uma visão constitucionalizada, refletindo a evolução que o tema passou desde a Assembleia Nacional Constituinte até a contemporaneidade.

4. Conclusão

O artigo explora a evolução da proteção ao direito à privacidade no Brasil, a partir de suas raízes históricas na Assembleia Nacional Constituinte. Percebe-se, a partir de uma revisão exploratória sobre o tema, que a constitucionalização do direito à privacidade no Brasil refletiu influências internacionais, em especial a ideia de privacidade como uma proteção contra intrusões indesejadas originada nos Estados Unidos e a dimensão informacional sedimentada na Europa.

Ademais, no contexto brasileiro, a introdução do habeas data reflete uma compreensão da necessidade de se assegurar ao indivíduo mecanismos efetivos para conhecer quais dados sobre si estavam sendo tratados e, a partir disso, tutelar efetivamente o direito à privacidade em uma era de crescente coleta e processamento de dados. Por fim, a positivação da proteção de dados pessoais como direito fundamental pela Emenda Constitucional nº 115/2022 consolida a trajetória evolutiva do direito à privacidade no Brasil.

A análise da trajetória do direito à privacidade no ordenamento constitucional brasileiro reforça a necessidade de um debate que

considere os avanços internacionais sobre o tema, mas que assegure contribuições novas que se adequem à realidade social e tecnológica do Brasil. Dessa forma, evita-se a reprodução de fórmulas prontas, mas que pecam pela desconsideração das particularidades nacionais e pela baixa efetividade. Assim, uma visão constitucionalizada e informada historicamente a respeito da privacidade assegura que esta continue a ser uma prioridade no ordenamento brasileiro e que este consiga fazer face ao dinamismo tecnológico que, por vezes, avança a custo das esferas de individualidade e do interesse coletivo.

Notas finais

- | | |
|---|--|
| | groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others”. |
| 1 Dworkin, R. (1986). <i>Law's empire</i> . Cambridge: Harvard University Press. | 10 Gavison, R. (1980). Privacy and the Limits of Law. <i>The Yale law journal</i> , 89(3), 421-471. |
| 2 FORBRUKERRADET. (2020). Out of control: How consumers are exploited by the online advertising industry. Disponível em: https://fil.forbrukerradet.no/wp-content/uploads/2020/01/2020-01-14-out-of-control-final-version.pdf . | 11 Shaw, T. (2013). Privacy Law and History: WWII – Forward. IAPP. Disponível em: https://iapp.org/news/a/2013-03-01-privacy-law-and-history-wii-forward/ . |
| 3 Oliva, T. et al. O que é o direito ao esquecimento? InternetLab, 2017. Disponível em: https://www.internetlab.org.br/pt/privacidade-e-vigilancia/lespecial-o-que-e-o-direito-ao-esquecimento/ . | 12 Seltzer, W., & Anderson, M. (2001). The dark side of numbers: The role of population data systems in human rights abuses. <i>Social Research</i> , 481-513. |
| 4 Brayne, S. (2020) <i>Predict and surveil: Data, discretion, and the future of policing</i> . Oxford University Press, USA, 2020. | 13 Idem, <i>ibidem</i> . |
| 5 DETECTA monitora o Estado de SP com mais de três mil câmeras de vídeo. Governo do Estado de São Paulo, 2017. Disponível em: https://www.saopaulo.sp.gov.br/spnoticias/detecta-monitora-o-estado-de-sao-paulo-com-3-mil-cameras-de-video/ . | 14 Shaw, T. (2013). Privacy Law and History: WWII – Forward. IAPP. Disponível em: https://iapp.org/news/a/2013-03-01-privacy-law-and-history-wii-forward/ . |
| 6 KOBLE, Nicole. The complicated truth about China's social credit system. <i>Wired</i> , 2019. Disponível em: https://www.wired.co.uk/article/china-social-credit-system-explained . | 15 Idem, <i>ibidem</i> . |
| 7 Warren, S., & Brandeis, L. (1989). The right to privacy. In <i>Killing the Messenger</i> (pp. 1-21). Columbia University Press. | 16 Doneda, D. (2020). <i>Da privacidade à proteção de dados pessoais [livro eletrônico]: elementos da formação da Lei Geral de Proteção de Dados</i> . 2. Ed. São Paulo: Thomson Reuters, pp. 158-163. |
| 8 Gavison, R. (1980). Privacy and the Limits of Law. <i>The Yale law journal</i> , 89(3), 421-471. | 17 PORTUGAL. [Constituição (1976)]. <i>Constituição da República Portuguesa</i> . Disponível em: https://www.parlamento.pt/Legislacao/Paginas/ConstituicaoRepublicaPortuguesa.aspx . |
| 9 Westin, A. F. (1968). Privacy and freedom. <i>Washington and Lee Law Review</i> , 25(1), 166. No original: “Privacy is the claim of individuals, | 18 Lima, J. A. D. O., Passos, E., & Nicola, J. R. (2013). A gênese do texto da Constituição de 1988: volume I-quadros. In <i>A gênese do texto da Constituição de 1988: volume I</i> , p. xxiii. |

19 BRASIL. Anteprojeto Constitucional, elaborado pela Comissão Provisória de Estudos Constitucionais, instituída pelo Decreto nº 91.450, de 18 de julho de 1985. Diário Oficial da República Federativa do Brasil. Suplemento especial ao nº 185. Sexta-feira, 26 de setembro de 1986. Brasília, DF.

20 BRASIL. Assembleia Nacional Constituinte. Anais da Assembleia Constituinte. Atas das Subcomissões. Subcomissão dos Direitos e Garantias Individuais. In: Câmara dos Deputados. Disponível em: <http://imagem.camara.gov.br/Imagem/d/pdf/sup63anc21maio1987.pdf#page=23>.

21 Idem, ibidem.

22 Idem, ibidem.

23 BRASIL. Assembleia Nacional Constituinte. Anais da Assembleia Constituinte. Atas das Subcomissões. Subcomissão dos Direitos e Garantias Individuais. In: Câmara dos Deputados. Disponível em: <http://imagem.camara.gov.br/Imagem/d/pdf/sup66anc27mai1987.pdf#page=59>.

24 Idem, ibidem.

25 Ibidem, pp. 104-105.

26 Ibidem, p. 106.

27 Idem, ibidem.

28 Ibidem, p. 107.

29 BRASIL. Assembleia Nacional Constituinte. Anteprojeto aprovado pela Subcomissão em 23 de maio de 1987. Subcomissão dos Direitos e Garantias Individuais. In: Câmara dos Deputados. Disponível em: <https://www.camara.leg.br/internet/constituicao20anos/DocumentosAvulsos/vol-81.pdf>.

30 BRASIL. Assembleia Nacional Constituinte. Parecer e substitutivo (nova redação). Comissão da Soberania e dos Direitos e Garantias do Homem e da Mulher. In: Câmara dos Deputados. Disponível em: <https://www.camara.leg.br/internet/constituicao20anos/DocumentosAvulsos/vol-68.pdf>.

31 BRASIL. Assembleia Nacional Constituinte. Anais da Assembleia Constituinte. Atas das Comissões. Comissão da Soberania e dos Direitos e Garantias do Homem e da Mulher. In: Câmara dos Deputados. Disponível em: <https://imagem.camara.gov.br/Imagem/d/pdf/sup87anc02jul1987.pdf#page=2>.

32 Ibidem, p. 21.

33 Ibidem, p. 20.

34 Idem, ibidem.

35 Ibidem, p. 20.

36 BRASIL. Assembleia Nacional Constituinte. Anais da Assembleia Constituinte. Atas das Comissões. Comissão da Soberania e dos Direitos e Garantias do Homem e da Mulher. In: Câmara dos Deputados. Disponível em: <https://imagem.camara.gov.br/Imagem/d/pdf/sup84anc26jun1987.pdf#page=2>.

37 BRASIL. Assembleia Nacional Constituinte. Anais da Assembleia Constituinte. Atas das Subcomissões. Subcomissão dos Direitos e Garantias Individuais. In: Câmara dos Deputados. Disponível em: <http://imagem.camara.gov.br/Imagem/d/pdf/sup63anc21maio1987.pdf#page=23>.

38 BRASIL. Assembleia Nacional Constituinte. Anais da Assembleia Constituinte. Atas das Comissões. Comissão de Sistematização.

In: Câmara dos Deputados. Disponível em: <https://imagem.camara.gov.br/Imagem/d/pdf/sup171Can-c27jan1988VolumeI-II.pdf#page=>.

39 Idem, ibidem.

40 Idem, ibidem.

41 Lima, J. A. D. O., Passos, E., & Nicola, J. R. (2013). A gênese do texto da Constituição de 1988: volume I-quadros. In A gênese do texto da Constituição de 1988: volume I, p. 30.

42 BRASIL. Assembleia Nacional Constituinte. Parecer e substitutivo (nova redação). Comissão da Soberania e dos Direitos e Garantias do Homem e da Mulher. In: Câmara dos Deputados. Disponível em: <https://www.camara.leg.br/internet/constituicao20anos/DocumentosAvulsos/vol-68.pdf>.

43 BRASIL. Constituição da República Federativa do Brasil de 1988. Disponível em http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm.

44 Doneda, D. (2020). Da privacidade à proteção de dados pessoais [livro eletrônico]: elementos da formação da Lei Geral de Proteção de Dados. 2. Ed. São Paulo: Thomson Reuters, pp. 77-81.

45 Idem, ibidem.

46 Peixoto, E. L. C. (2017). O conteúdo do direito da privacidade no direito brasileiro contemporâneo.

47 Silva, J. (2009). Curso de direito constitucional positivo. 32. ed. São Paulo: Malheiros, p. 207.

48 Lôbo, P. (2012). Direito civil: parte geral. 3 ed. São Paulo: Saraiva, p. 143.

49 Ferraz Júnior, T. (1993). Sigilo de dados: o direito à privacidade e os limites à função fiscalizadora do Estado. Revista da Faculdade de Direito, Universidade de São Paulo, 88, 439-459.

50 Doneda, D. (2020). Da privacidade à proteção de dados pessoais [livro eletrônico]: elementos da formação da Lei Geral de Proteção de Dados. 2. Ed. São Paulo: Thomson Reuters, pp. 77-81.

51 Solove, D.J. (2008). Understanding privacy.

52 Doneda, D. (2020). Da privacidade à proteção de dados pessoais [livro eletrônico]: elementos da formação da Lei Geral de Proteção de Dados. 2. Ed. São Paulo: Thomson Reuters, pp. 77-81.

53 Idem, ibidem.

54 Idem, ibidem.

55 Idem, p. 43.

56 Hughes, K. (2015). The social value of privacy, the value of privacy to society and human rights discourse. Social dimensions of privacy: Interdisciplinary perspectives, 225, 225.

57 Rubinstein, I. (2012). Big data: The end of privacy or a new beginning?. International Data Privacy Law (2013 Forthcoming), NYU School of Law, Public Law Research Paper, (12-56).

58 Doneda, D. (2020). Da privacidade à proteção de dados pessoais [livro eletrônico]: elementos da formação da Lei Geral de Proteção de Dados. 2. Ed. São Paulo: Thomson Reuters, pp. 92-96.

59 Ibidem, p. 96.

60 Gavison, R. (1980). Privacy and the Limits of Law. The Yale law journal, 89(3), 421-471.	Estado: o que permanece e o que deve ser reconsiderado. Internet & Sociedade, São Paulo, n. 1, p. 64-90.
61 BRASIL. REsp 1574681/RS, Rel. Ministro ROGERIO SCHIETTI CRUZ, SEXTA TURMA, julgado em 20/04/2017, DJe 30/05/2017	69 Ferraz Júnior, T. (1993). Sigilo de dados: o direito à privacidade e os limites à função fiscalizadora do Estado. Revista da Faculdade de Direito, Universidade de São Paulo, 88, 439-459.
62 BRASIL. RE 603616, Relator(a): GILMAR MENDES, Tribunal Pleno, julgado em 05/11/2015, ACÓRDÃO ELETRÔNICO REPERCUSSÃO GERAL - MÉRITO DJe-093 DIVULG 09-05-2016 PUBLIC 10-05-2016	70 Idem, ibidem.
63 Sarlet, I. (2015). Decisão do STF sobre violação do domicílio indica posição prudencial. Conjur. Disponível em: https://www.conjur.com.br/2015-nov-13/direitos-fundamentais-decisao-stf-violacao-domicilio-indica-posicao-prudencial .	71 Idem, ibidem.
64 Zwicker, I. O risco à inviolabilidade de domicílio em razão do precedente Daniel Silveira. Disponível em: https://www.conjur.com.br/2021-fev-24/zwicker-inviolabilidade-domicilio-daniel-silveira .	72 Queiroz, R. M. R., & Ponce, P. P. (2020). Tércio Sampaio Ferraz Júnior e Sigilo de dados: o direito à privacidade e os limites à função fiscalizadora do Estado: o que permanece e o que deve ser reconsiderado. Internet & Sociedade, São Paulo, (1), 64-90.
65 Idem, ibidem.	73 Idem, ibidem.
66 Queiroz, R.; Ponce, P. Tércio Sampaio Ferraz Júnior e Sigilo de dados: o direito à privacidade e os limites à função fiscalizadora do Estado: o que permanece e o que deve ser reconsiderado. Internet & Sociedade, São Paulo, n. 1, p. 64-90.	74 Idem, ibidem.
67 Ferraz Júnior, T. (1993). Sigilo de dados: o direito à privacidade e os limites à função fiscalizadora do Estado. Revista da Faculdade de Direito, Universidade de São Paulo, 88, 439-459.	75 Idem, ibidem.
68 Queiroz, R.; Ponce, P. Tércio Sampaio Ferraz Júnior e Sigilo de dados: o direito à privacidade e os limites à função fiscalizadora do	76 BRASIL. Supremo Tribunal Federal (2ª Turma). Habeas Corpus 168052/SP. Paciente: Rodrigo Ricardo Laurindo. Impetrante: Arai Mendonça Brazão. Coator: Superior Tribunal de Justiça. Relator: Min. Gilmar Mendes, 02 de dez. de 2020. Disponível em: https://jurisprudencia.stf.jus.br/pages/search/sjur437471/false .
	77 Ferreira, A. B. D. H. (2009). Novo dicionário Aurélio da língua portuguesa.
	78 Ibid, p. 92.
	79 Mayer-Schönberger, V. (2010). Beyond Privacy, Beyond Rights—Toward a "Systems" Theory of Information Governance. California Law Review, 1853-1885.

80 Cate, F. H. (2006). The failure of fair information practice principles. Consumer protection in the age of the information economy.

81 FORBRUKERRADET. Deceived by design, 2018. Disponível em: <https://fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf>.

82 Cf. Ralston, W. (2021). They told their therapists everything, hackers leaked it all. Wired. Disponível em: <https://www.wired.com/story/vastaamo-psychotherapy-patients-hack-data-breach/>.

83 Mendes, L. S. F. (2018). Habeas data e autodeterminação informativa: os dois lados de uma mesma moeda. Revista Brasileira de Direitos Fundamentais & Justiça, 12(39), 185-216.

84 Idem, ibidem.

85 BRASIL. Constituição da República Federativa do Brasil de 1988. Disponível em http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm.

86 Doneda, D. (2020). Da privacidade à proteção de dados pessoais [livro eletrônico]: elementos da formação da Lei Geral de Proteção de Dados. 2. Ed. São Paulo: Thomson Reuters, pp. 272-285.

87 Idem, ibidem.

88 Frosini, T. E. (2008). La libertà informatica: brevi note sull'attualità di una teoria giuridica. Informatica e diritto, 17(1-2), 87-97.

89 Idem, ibidem.

90 Idem, ibidem.

91 Mendes, L. S. F. (2018). Habeas data e autodeterminação informativa: os dois lados de uma mesma moeda. Revista Brasileira de Direitos Fundamentais & Justiça, 12(39), 185-216.

92 BRASIL. Supremo Tribunal Federal. Recurso Extraordinário nº 673707/MG. Recorrente: RIGLIGMINAS Distribuidora LTDA. Recorrido: União. Relator: Min. Luiz Fux, 30 de set. de 2015. Disponível em: <https://jurisprudencia.stf.jus.br/pages/search/sjur322444/false>.

93 Idem, ibidem.

94 Mendes, L. S. F. (2018). Habeas data e autodeterminação informativa: os dois lados de uma mesma moeda. Revista Brasileira de Direitos Fundamentais & Justiça, 12(39), 185-216.

95 Doneda, D. (2020). Da privacidade à proteção de dados pessoais [livro eletrônico]: elementos da formação da Lei Geral de Proteção de Dados. 2. Ed. São Paulo: Thomson Reuters, pp. 272-285.

96 Dallari, D. (2002). O habeas data no sistema jurídico brasileiro. Revista da Faculdade de Direito, Universidade de São Paulo, 97, 239-253.

Referências bibliográficas

- BRASIL – ASSEMBLEIA NACIONAL CONSTITUINTE. Ata da 10ª reunião em 28/04/87. Publicada em 27/05/87 no Diário da Assembleia Nacional Constituinte (DANC) em 27/05/87 no Suplemento 66, p. 90), p. 105. Disponível em: sup66anc27mai1987.pdf (camara.gov.br).
- BRASIL. Anteprojeto Constitucional, elaborado pela Comissão Provisória de Estudos Constitucionais, instituída pelo Decreto nº 91.450, de 18 de julho de 1985. Diário Oficial da República Federativa do Brasil. Suplemento especial ao nº 185. Sexta-feira, 26 de setembro de 1986. Brasília, DF.
- BRASIL. Assembleia Nacional Constituinte. Anais da Assembleia Constituinte. Atas das Subcomissões. Subcomissão dos Direitos e Garantias Individuais. In: Câmara dos Deputados. Disponível em: <http://imagem.camara.gov.br/Imagem/d/pdf/sup63anc21maio1987.pdf#page=23>.
- BRASIL. Assembleia Nacional Constituinte. Anais da Assembleia Constituinte. Atas das Comissões. Comissão da Soberania e dos Direitos e Garantias do Homem e da Mulher. In: Câmara dos Deputados. Disponível em: <https://imagem.camara.gov.br/Imagem/d/pdf/sup87anc02jul1987.pdf#page=2>.
- BRASIL. Assembleia Nacional Constituinte. Anais da Assembleia Constituinte. Atas das Comissões. Comissão de Sistematização. In: Câmara dos Deputados. Disponível em: [https://imagem.camara.gov.br/Imagem/d/pdf/sup171Canc27jan1988VolumeI-II.pdf#page=.](https://imagem.camara.gov.br/Imagem/d/pdf/sup171Canc27jan1988VolumeI-II.pdf#page=)
- BRASIL. Assembleia Nacional Constituinte. Anteprojeto aprovado pela Subcomissão em 23 de maio de 1987. Subcomissão dos Direitos e Garantias Individuais. In: Câmara dos Deputados. Disponível em: <https://www.camara.leg.br/internet/constituicao20anos/DocumentosAvulsos/vol-81.pdf>.
- BRASIL. Assembleia Nacional Constituinte. Parecer e substitutivo (nova redação). Comissão da Soberania e dos Direitos e Garantias do Homem e da Mulher. In: Câmara dos Deputados. Disponível em: <https://www.camara.leg.br/internet/constituicao20anos/DocumentosAvulsos/vol-68.pdf>.
- BRASIL. Constituição da República Federativa do Brasil de 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm.
- BRASIL. RE 603616, Relator(a): GILMAR MENDES, Tribunal Pleno, julgado em 05/11/2015, ACÓRDÃO ELETRÔNICO REPERCUSSÃO GERAL - MÉRITO DJe-093 DIVULG 09-05-2016 PUBLIC 10-05-2016
- BRASIL. REsp 1574681/RS, Rel. Ministro ROGERIO SCHIETTI CRUZ, SEXTA TURMA, julgado em 20/04/2017, DJe 30/05/2017
- BRASIL. Supremo Tribunal Federal (2ª Turma). Habeas Corpus 168052/SP. Paciente: Rodrigo Ricardo Laurindo. Impetrante: Arai Mendonça Brazão. Coator: Superior Tribunal de Justiça. Relator: Min. Gilmar Mendes, 02 de dez. de 2020. Disponível em: <https://jurisprudencia.stf.jus.br/pages/search/sjur437471/false>.
- BRASIL. Supremo Tribunal Federal. Recurso Extraordinário nº 673707/MG. Recorrente: RIGLIGMINAS Distribuidora LTDA. Recorrido: União. Relator: Min. Luiz Fux, 30 de set. de 2015. Disponível em: <https://jurisprudencia.stf.jus.br/pages/search/sjur322444/false>.

- BRASIL. Supremo Tribunal Federal. ADI 6387 MC-Ref, Relator(a): ROSA WEBER, Tribunal Pleno, julgado em 07-05-2020, PROCESSO ELETRÔNICO
- Brayne, S. (2020) *Predict and surveil: Data, discretion, and the future of policing*. Oxford University Press, USA, 2020.
- Cate, F. H. (2006). *The failure of fair information practice principles. Consumer protection in the age of the information economy*.
- Dallari, D. (2002). *O habeas data no sistema jurídico brasileiro*. Revista da Faculdade de Direito, Universidade de São Paulo, 97, 239-253.
- DETECTA monitora o Estado de SP com mais de três mil câmeras de vídeo. Governo do Estado de São Paulo, 2017. Disponível em: <https://www.saopaulo.sp.gov.br/spnoticias/detecta-monitora-o-estado-de-sao-paulo-com-3-mil-cameras-de-video/>.
- Doneda, D. (2020). *Da privacidade à proteção de dados pessoais [livro eletrônico]: elementos da formação da Lei Geral de Proteção de Dados*. 2. Ed. São Paulo: Thomson Reuters.
- Dworkin, R. (1986). *Law's empire*. Cambridge: Harvard University Press.
- Ferraz Júnior, T. (1993). *Sigilo de dados: o direito à privacidade e os limites à função fiscalizadora do Estado*. Revista da Faculdade de Direito, Universidade de São Paulo, 88, 439-459.
- Ferreira, A. B. D. H. (2009). *Novo dicionário Aurélio da língua portuguesa*.
- FORBRUKERRADET. *Out of control: How consumers are exploited by the online advertising industry*. Disponível em: <https://fil.forbrukerradet.no/wp-content/uploads/2020/01/2020-01-14-out-of-control-final-version.pdf>.
- Frosini, T. E. (2008). *La libertà informatica: brevi note sull'attualità di una teoria giuridica*. *Informatica e diritto*, 17(1-2), 87-97.
- Gavison, R. (1980). *Privacy and the Limits of Law*. *The Yale law journal*, 89(3), 421-471.
- Hughes, K. (2015). *The social value of privacy, the value of privacy to society and human rights discourse. Social dimensions of privacy: Interdisciplinary perspectives*, 225, 225.
- KOBLE, Nicole. *The complicated truth about China's social credit system*. *Wired*, 2019. Disponível em: <https://www.wired.co.uk/article/china-social-credit-system-explained>.
- Lima, J. A. D. O., Passos, E., & Nicola, J. R. (2013). *A gênese do texto da Constituição de 1988: volume I-quadros*. In *A gênese do texto da Constituição de 1988: volume I*, p. 30.
- Lôbo, P. (2012). *Direito civil: parte geral*. 3 ed. São Paulo: Saraiva, p. 143.
- Mayer-Schönberger, V. (2010). *Beyond Privacy, Beyond Rights—Toward a "Systems" Theory of Information Governance*. *California Law Review*, 1853-1885.
- Mendes, L. S. F. (2018). *Habeas data e autodeterminação informativa: os dois lados de uma mesma moeda*. *Revista Brasileira de Direitos Fundamentais & Justiça*, 12(39), 185-216.
- Mendes, L. S., & Fonseca, G. C. S. D. (2020). *STF reconhece direito fundamental à proteção de dados*. *Revista de direito do consumidor [recurso eletrônico]*. São Paulo, (130).
- Oliva, T. et al. *O que é o direito ao esquecimento?* InternetLab, 2017. Disponível em: <https://www.internetlab.org.br/pt/privacidade-e-vigilancia/lespecial-o-que-e-o-direito-ao-esquecimento/>.

- Peixoto, E. L. C. (2017). O conteúdo do direito da privacidade no direito brasileiro contemporâneo.
- PORTUGAL. [Constituição (1976)]. Constituição da República Portuguesa. Disponível em: <https://www.parlamento.pt/Legislacao/Paginas/ConstituicaoRepublicaPortuguesa.aspx>.
- Queiroz, R.; Ponce, P. Tércio Sampaio Ferraz Júnior e Sigilo de dados: o direito à privacidade e os limites à função fiscalizadora do Estado: o que permanece e o que deve ser reconsiderado. *Internet & Sociedade*, São Paulo, n. 1, p. 64-90.
- Ralston, W. (2021). They told their therapists everything, hackers leaked it all. *Wired*. Disponível em: <https://www.wired.com/story/vastaamo-psychotherapy-patients-hack-data-breach/>.
- Rubinstein, I. (2012). Big data: The end of privacy or a new beginning?. *International Data Privacy Law* (2013 Forthcoming), NYU School of Law, Public Law Research Paper, (12-56).
- Sarlet, I. (2015). Decisão do STF sobre violação do domicílio indica posição prudencial. *Conjur*. Disponível em: <https://www.conjur.com.br/2015-nov-13/direitos-fundamentais-decisao-stf-violacao-domicilio-indica-posicao-prudencial>.
- Seltzer, W., & Anderson, M. (2001). The dark side of numbers: The role of population data systems in human rights abuses. *Social Research*, 481-513.
- Shaw, T. (2013). Privacy Law and History: WWII – Forward. IAPP. Disponível em: <https://iapp.org/news/a/2013-03-01-privacy-law-and-history-wwii-forward/>.
- Silva, J. (2009). Curso de direito constitucional positivo. 32. ed. São Paulo: Malheiros, p. 207.
- Solove, D. J. (2008). Understanding privacy.
- Warren, S., & Brandeis, L. (1989). The right to privacy. In *Killing the Messenger* (pp. 1-21). Columbia University Press.
- Westin, A. F. (1968). Privacy and freedom. *Washington and Lee Law Review*, 25(1), 166.
- Zwicker, I. O risco à inviolabilidade de domicílio em razão do precedente Daniel Silveira. Disponível em: <https://www.conjur.com.br/2021-fev-24/zwicker-inviolabilidade-domicilio-daniel-silveira>.

ARTIGO

Análise de podcast pela perspectiva metodológica da roleta interseccional

Carolina Salles Carvalho

<https://orcid.org/0000-0002-8657-5096>

sallescarolina@hotmail.com

Universidade de São Paulo (USP),
Centro de Pesquisa, Inovação
e Difusão em Neuromatemática
(CEPID NeuroMat),
São Paulo, SP, Brasil

João Alexandre Peschanski

<https://orcid.org/0000-0002-2352-1787>

japeschanski@casperlibero.edu.br

Faculdade Cásper Líbero (FCL),
São Paulo, SP, Brasil

Análise de podcast pela perspectiva metodológica da roleta interseccional

Palavras-chave

Comunicação Digital
Podcast
Roleta Interseccional
Pesquisa Empírica

Resumo

O artigo faz uma análise da primeira temporada do *podcast* Rádio Batente, adotando como alicerce teórico-metodológico a Roleta Interseccional, proposta por Fernanda Carrera. São identificados os eixos "Formação interseccional-discursiva", "Ethos interseccional" e "Negociações interseccionais" e são expostas as situações e estratégias identitárias no *podcast*. Sob a perspectiva da Comunicação, discutem-se tópicos como a questão da autoria em um produto marcado pela presença de múltiplas vozes e pela edição maciça, e busca-se entender as potencialidades e limites dessa proposta para esse tipo de mídia. Por fim, o trabalho busca compreender como os estereótipos marcam os profissionais e os discursos sobre si, bem como refletir sobre como a imersão propiciada pelo uso dos recursos sonoros e pela transformação dos jornalistas em personagens permitiram colocar em pauta a interação com o internauta-ouvinte, mesmo em uma mídia, *a priori*, unidirecional.

Podcast analysis under the intersectional roulette methodological perspective

Keywords

Digital Communication
Podcast
Intersectional Roulette
Empirical Research.

Abstract

This paper develops an analysis of the first season of the podcast “Rádio Batente”, broadcasted by Reporter Brasil NGO, adopting the intersectional roulette, proposed by Fernanda Carrera, as the theoretical framework. Three main axes have been identified in the podcast: “Intersectional discursive formation”, “Intersectional ethos” and “Intersectional negotiation” from which the multi-layered situations and identity strategies are presented. Under the perspective of Communication, the topic of authorship of a product characterised by multiple voices and massive editing, which are features of narrative podcasts, are discussed, seeking to understand the strengths and limits of Carrera’s proposal for this type of media. Finally, the paper seeks to understand how stereotypes affect workers and the discourses of the self as well as deepen the debate about how the immersion sensation promoted by audio resources and by the transformation of journalists into characters allowed to shed light on the interaction with the listener, even in a media that is, a priori, unidirectional.

1. Introdução

O artigo tem como proposta analisar a primeira temporada do *podcast* Rádio Batente (Jornadas), veiculado pela ONG Repórter Brasil, identificando potencialidades e limitações da metodologia da roleta interseccional, formulada para a compreensão científica de objetos comunicacionais. Tal investigação empírica adota a perspectiva da interseccionalidade e busca contribuir para a discussão sobre como os produtos midiáticos podem revelar, denunciar ou reverberar desigualdades e opressões silenciadas, colocando sob a luz dos holofotes as relações de poder que influenciam a vida em sociedades marcadas pela diversidade (Collins & Bilge, 2021).

A roleta interseccional é uma proposta teórico-metodológica que coloca em evidência as relações intersubjetivas e as negociações identitárias nos produtos midiáticos. Aplicar essa perspectiva em um produto de mídia novo e, com isso, expor a sua lógica de funcionamento e ampliar o seu uso, é, por si só, uma contribuição ao campo. Mas chama a atenção o fato de que as principais publicações sobre a roleta interseccional tenham investigado redes sociais em que há interação direta entre transmissor e receptor, especialmente o Instagram (Carrera, 2021a; 2021b); sendo assim, é oportuno questionar em que medida a proposta de Carrera aplica-se adequadamente ao estudo de um *podcast*? Quais são as potencialidades e os limites da roleta interseccional nesse caso?

Este artigo inicia-se com uma apresentação do marco teórico, seguida de uma exposição da proposta teórico-metodológica da roleta interseccional. Na sequência, há a análise do *podcast* Rádio Batente (Jornadas), a partir dos eixos principais da perspectiva de Carrera: Formação interseccional-discursiva, Ethos interseccional e Negociações interseccionais.

1.1 Marco teórico e recorte histórico inicial: primeiras aproximações

Como ferramenta analítica, a interseccionalidade considera que diferentes categorias são inter-relacionadas e moldam-se, mutuamente, adotando como marcadores a raça, a classe, o gênero, a orientação sexual, a nacionalidade, a etnia, a faixa etária, entre outros. “A análise interseccional privilegia a discussão de experiências e vivências que articulam categorias que, historicamente foram conceituadas isoladamente como, por exemplo, quando pensamos a o gênero sem darmos atenção aos marcadores raciais, ou de classe, que atravessam determinada experiência” (Díaz-Benítez & Mattos, 2019, p. 67).

Proposto na década de 1980, o conceito de interseccionalidade nasce da percepção das ativistas negras norte-americanas de que as suas necessidades relacionadas às questões laborais, à educação e à saúde eram ignoradas em movimentos sociais antirracistas, feministas e naqueles encabeçados por sindicatos atuantes em defesa da classe trabalhadora. “Cada um desses movimentos sociais privilegiou uma categoria de análise e de ação em detrimento de outras: por exemplo, raça no movimento em favor dos direitos civis; gênero no movimento feminista; classe no movimento sindical” (Collins & Bilge, 2021, p. 17).

1.2 Revisão de literatura: presença nas pesquisas brasileiras

Sifuentes (2022) fez um mapeamento das pesquisas em Comunicação que articulam gênero, raça e classe, apresentadas nos Encontros anuais da Compós – associação que reúne os Programas de Pós-Graduação em Comunicação de instituições de ensino superior públicas e privadas no Brasil – entre 2011 e 2020. Além de esmiuçar a presença dessa tríade, a autora também investigou se a interseccionalidade estava presente nos cruzamentos empíricos ou se foi utilizada como alicerce teórico em trabalhos que investigam as produções em plataformas digitais, aprofundando a discussão sobre a democratização das representações dessas três categorias nessas mídias.

De acordo com o levantamento, 26 trabalhos foram pautados na interseccionalidade, o que equivale a 1,5% do total, sendo que 22 deles articularam raça, gênero e classe (isoladamente, como uma dupla ou um trio de categorias), incluindo dois artigos de autoria de Carrera, pesquisadora que embasa o presente artigo, sendo que, nos trabalhos em questão, ela investigou as representações de raça e gênero nos algoritmos em bancos de imagens digitais (Sifuentes, 2022).

Dentre os pontos ressaltados por Sifuentes, no mapeamento realizado, está a ausência de trabalhos relacionados à raça, justamente em um momento histórico efervescente para o movimento negro do país, o que leva a autora a concluir que a elite intelectual que se debruça sobre a Comunicação tem deixado de abordar a temática. Outra constatação é a concentração de estudos sobre recepção e consumo, com poucas contribuições ao grupo de trabalho relacionado às sexualidades e aos gêneros. Por fim, o artigo aponta que há um número maior

de trabalhos no ano de 2020, em contraponto aos anos iniciais do levantamento, o que sinaliza um aumento recente de interesse dos profissionais de Comunicação pela interseccionalidade (Sifuentes, 2022).

Embora esses trabalhos ainda apareçam em pequeno número, verificamos uma clara tendência de crescimento e são claras as contribuições que esses estudos trazem, especialmente quando aliam discussões acerca das possibilidades – não concretizadas – de as plataformas digitais permitirem uma representação menos opressora desses grupos (Sifuentes, 2022, p. 24).

Uma busca por artigos em português no Google Acadêmico, por exemplo, utilizando-se as palavras-chave “Comunicação” e “Interseccionalidade” em qualquer ponto do texto, desde 2020, encontra um número, aproximadamente, de oito mil trabalhos. Quando essa busca pelas mesmas palavras-chave fica restrita aos títulos, as respostas reduzem-se a nove publicações, o que reforça o aumento de interesse por esse referencial nos últimos anos.

1.3

Interseccionalidade e Comunicação: especificidades do campo de estudo

Ao considerar a Comunicação como campo de investigação interseccional, tem-se a “interseccionalidade representacional”, que aponta a construção cultural e os sentidos marginalizados sobre os sujeitos, numa seara que abarca não só as produções midiáticas, mas que se faz presente nas relações sociais, inclusive acadêmicas: para Carrera, todos os produtos jornalísticos, publicitários e de mídia de massa – da fotografia aos modos de interação assíncrona propostos pelas tecnologias digitais – podem ser investigados a partir de uma metodologia própria.

(...) todas essas manifestações e seus aparatos são objetos de análise comuns em estudos da Comunicação e, também, estão sujeitos à reprodução e reforço de dinâmicas de opressão interseccionais. É sob a égide desta acepção que é proposto aqui um caminho metodológico inicial, denominado roleta interseccional; uma ferramenta discursivo-operacional que pretende identificar os rastros da interseccionalidade nas diversas expressões e experiências comunicacionais, sem perder de vista, é claro, sua motivação elementar: a reivindicação pela dignidade (Carrera, 2021a, p. 6).

Já ao considerar as especificidades da mídia *podcast*, é interessante ressaltar um ponto abordado por Carrera, quando ela esmiúça as particularidades de produtos audiovisuais.

Nesse sentido, possivelmente, a linguagem geolocalizada materializada pelos sotaques, por exemplo, pode ser extremamente iluminada em produtos audiovisuais, mas passar relativamente despercebida em imagens estáticas. É claro que o sujeito (representado na roleta pela circunferência central) é todo atravessado pela roleta e toda ela, além da junção com outros fatores, vai ajudar na sua constituição identitária. No entanto, a situação comunicacional não necessariamente mobiliza todos os atravessamentos (Carrera, 2021a, p. 12).

2. Metodologia

2.1 Da teoria ao empirismo: uma proposição investigativa

Carrera (2021a; 2021b) apresenta operadores analíticos e ferramentas epistemológicas viáveis no campo da Comunicação, buscando identificar como os rastros dessas hastes estão presentes nas interações cotidianas, na Comunicação midiática e nas representações discursivas. Baseada e completamente atrelada à iniciativa de Crenshaw (1989), que tinha como princípio a busca por igualdade em sistemas jurídicos, nessa seara, busca-se perceber como opressões interseccionais rasuram a subjetividade, os discursos, os produtos e os espaços comunicacionais e podem ser fundamentais para composição dos sujeitos e dos seus comportamentos em interação. Nesse sentido, não se negligencia aqui o *ethos* de justiça social, essencial a qualquer aplicação do conceito (Carrera, 2021a, p. 9).

A pesquisadora propõe uma roleta interseccional para direcionar as perguntas de pesquisa em Comunicação, sendo que a proposta é justamente questionar se determinada categoria é importante para o entendimento daquele objeto ou de que maneira deixaria rastros na materialidade comunicacional (Carrera, 2021a). Nesse contexto, tal instrumento é formado por hastes, que representam as avenidas identitárias fundamentais, brevemente descritas abaixo:

O marcador “gênero” é essencial para entender os avanços conquistados pelos estudos feministas além de abarcar concepções renovadas sobre masculinidades, cisgeneridade compulsória e cisheteronormatividade (Carrera, 2021a). Já a questão da “raça” é fundamental na metodologia interseccional por permitir a compreensão de como o racismo opera sobre os sujeitos, tanto pela sua representação e pelas práticas de violência quanto pela via do silenciamento, com a exclusão desse tema do debate na Comunicação. “Na metodologia da roleta interseccional, é essencial localizar de que forma a racialização é implicada no objeto, seja pelas escolhas discursivas, pelas marcações identitárias como pelos comportamentos sociais compreendidos” (Carrera, 2021a, p.5).

O marcador “classe” é importante pelo impacto que o marxismo tem nas ciências sociais, sendo oportuno reforçar que, geralmente, esse marcador e, por consequência, a desigualdade é resultante de um conjunto de fatores que ultrapassa a conjuntura econômica. Ao direcionar a problemática para a área da Comunicação, por exemplo, faz-se necessário compreender a influência da “classe” na construção dos sentidos, além de investigar como esse marcador estabelece limites interacionais para as apropriações por meio da linguagem e dos comportamentos sociais (Carrera, 2021a).

O marcador “geolocalização”, usualmente, é materializado por meio do corpo e no domínio da comunicação, pela oralidade, pelo modo como as interações acontecem e

pelo comportamento dos sujeitos, demarcando quem pode ou não existir. Na seara da Comunicação, é imprescindível apontar os efeitos da colonialidade sobre os sujeitos, que tendem a estigmatizar e estereotipar, evidenciando preconceitos linguísticos e a consequente hierarquização dos sujeitos, além de resvalar no risco da mídia representar as pessoas de maneira caricata, por exemplo (Carrera, 2021a).

O marcador “peso” coloca em pauta o corpo gordo, usualmente, atrelado ao fracasso estético, à marginalização e ao sentimento de inadequação social: “o sujeito gordo sente a exclusão social no plano estético, no âmbito do trânsito social e, inclusive, no acesso a direitos fundamentais, uma vez que há patologização do corpo gordo por meio da construção da doença da obesidade” (Carrera, 2021a, p.6). Nesse contexto, também é preciso considerar que homens e mulheres gordos ocupam lugares diferentes socialmente, tendo como referencial a assimetria em relação às expectativas e aos estigmas compartilhados socialmente. Na roleta interseccional, o termo peso designa parâmetros que classificam o corpo como magro ou gordo, extrapolando o entendimento médico-científico – atrelado ao conceito de Índice de Massa Corporal – para valorizar a percepção social do peso em relação ao sucesso ou ao fracasso (Carrera, 2021a).

Já o marcador “idade” reverbera como a sociedade costuma atrelar a velhice à percepção de decrepitude e a uma fase sombria, marcada por doenças e pelo temor à morte, além de ser vinculada ao isolamento social no fim da vida. Nesse cenário, é fundamental evidenciar que o processo de envelhecimento é atravessado por outras matrizes de opressão, com destaque para o marcador “classe”, já que a pobreza pode potencializar a experiência de marginalização e de abreviação da vida. Sendo assim, pode-se utilizar o termo “velhices”, no plural, considerando-se a interação entre diferentes marcadores interseccionais (Carrera, 2021a).

A “deficiência” também está presente como um marcador na proposta de Carrera. Nessa temática, a pesquisadora aponta que tal condição é atrelada à incapacidade inata ou à necessidade de superação, acompanhadas pelo sofrimento e pela pena, aprisionando o sujeito numa rede de estereótipos, que legitimam a exclusão. Ao se considerar a produção midiática, com destaque para a publicidade, nota-se a tendência de apresentar uma visão que homogeneiza as diferentes deficiências, desconsiderando as especificidades de cada condição, além de existir uma tendência a comercializar a piedade, reforçando uma abordagem capacitista sobre o assunto (Carrera, 2021a).

O marcador “sexualidade” busca superar a biologização que acompanha essa temática, evidenciando de que maneira as dimensões políticas sobre a sexualidade e as disputas de poder envolvem os caminhos do desejo, tornando-se pilares importantes na construção do sujeito sob investigação pela Comunicação. Nesse sentido, é crucial compreender os efeitos da heterossexualidade nas marcações comunicacionais, considerando-se a cultura heteronormativa e os discursos vigentes em relação a um corpo legitimado. Ao ampliar a discussão, coloca-se em pauta, também, os discursos direcionados às minorias sexuais (Carrera, 2021a).

A partir desse detalhamento, essa segunda etapa da análise da roleta interseccional deve ser realizada considerando-se os três eixos: “Formação interseccional-discursiva”, “*Ethos* interseccional” e “Negociações interseccionais”. Vale pontuar que o primeiro deles alicerça a sua proposta no conceito de formação discursiva, de autoria de Michel Foucault. A partir dessa referência, Carrera aponta para a existência um sistema normativo-social que rege as pessoas, determinando performances verbais e comportamentais, em um agrupamento de regras que delimitam os sujeitos como objetos de um discurso. Sendo assim, não seria possível uma existência *a priori*. “Na metodologia da

roleta interseccional, descrever as formações discursivas que regem cada eixo de opressão, portanto é identificar quais são os imperativos e quais são os silenciamentos aos quais os indivíduos por hastes iluminadas estão assujeitados” (Carrera, 2021a, p. 13).

O segundo eixo é o “*Ethos* interseccional”, que denominaria “(...) a silhueta imagética do sujeito construída no e pelo discurso, a partir das suas estratégias de enunciação” (Carrera, 2021a, p. 13), uma imagem que seria validada por interlocutores, a partir de coerções culturais e sociais vigentes em uma determinada época. Ressalta-se que tal imagem é intimamente ligada aos estereótipos presentes, sendo oportuno uma investigação aguçada para detectar situações normalizadas coletivamente justamente pela ausência de um olhar crítico que intervenha causando uma fissura no *status quo*: “(...) é relevante perceber não somente os estereótipos atribuídos a cada eixo que podem constranger as liberdades identitárias dos indivíduos, como também compreender as negociações, que acontecem nas interações e ressignificam os limites das suas existências” (Carrera, 2021a, p. 14).

Por fim, o eixo “Negociações interseccionais” se debruça sobre as interações sociais. Aqui, torna-se preponderante compreender que não existe uma essência que anteceda à interação com o outro, sendo que a própria constituição da identidade é entremeada pelo processo contínuo de se ajustar à realidade partilhada e de reagir às normas vigentes no entorno. De maneira sucinta, a formação da identidade só é possível a partir do contato social, o que nos leva a questionar: o quanto o quanto do senso de si é forjado a partir do outro? “Desse modo, dentro de cada formação interseccional-discursiva, diante de cada eixo de opressão, sujeitos negociam suas imagens de si, seus comportamentos e seus discursos, numa tentativa de gerenciar as impressões causadas e fazer valer a sua existência” (Carrera, 2021a, p. 15).

A partir dessa tríade, realiza-se uma análise dos cinco episódios com base na metodologia da roleta interseccional proposta por Carrera. Inicialmente, a própria autora publicou um estudo de caso comparativo para ilustrar a aplicabilidade de tal metodologia. No artigo científico, a pesquisadora analisou dois perfis de Instagram de influenciadoras digitais pautadas pelo movimento *Body Positive*, que tem como assinatura a “aceitação e valorização de corpos marginalizados, sobretudo aqueles que apresentam marcas visíveis da sua diferença – gordos, com deficiência, com cicatrizes ou marcas de envelhecimento, racionalizados” (Carrera, 2021b, p. 8).

Vale assinalar, nesse ponto, as especificidades midiáticas do Instagram e como são entrecortadas pelos três eixos da roleta interseccional. No tópico, “Formação interseccional-discursiva”, Carrera ressalta os temas abordados e assinala o fato de que, pela plataforma utilizada ser híbrida, é mais difícil fazer uma leitura que privilegie ora as imagens, ora o conteúdo audiovisual, separadamente. Evidentemente, uma particularidade que não encontra eco quando o objeto de investigação é um *podcast* (Carrera, 2021b).

Já ao se pensar no *ethos* interseccional, Carrera debruça-se sobre as escolhas enunciativas para a construção de um discurso sobre si, que pode ser “construído pelos conteúdos postados e pelas escolhas enunciativas materializadas nas imagens, na linguagem, nos produtos comunicacionais postos em circulação” (Carrera, 2021b, p. 14). É importante ressaltar que os dois perfis são contas pessoais, escritos em primeira pessoa. Como esse discurso de si pode ser avaliado em um *podcast*? Quem é o(a) protagonista? O jornalista que conduz uma narrativa ou o personagem principal, quando há um que se destaca? Não seriam o editor e o roteirista os verdadeiros detentores desse discurso já que podem conduzir a criação e a edição do texto, ainda que as suas vozes sejam “inaudíveis” no produto que chega ao público final?

Ao colocar em pauta esses questionamentos, pode-se reconhecer apontamentos que são decorrentes da escolha da roleta interseccional como método de análise, além de pontuar reverberações que também estão presentes na produção científica sobre *podcasts* narrativos que lançam mão do *storytelling* como estratégia para contar histórias. Nesse segundo caso, o papel do jornalista é revisitado porque a presença da subjetividade na descrição dos fatos tensiona a posição de destaque usualmente ocupada pelos sujeitos e assuntos que são foco da narrativa jornalística. Porém, mais do que hierarquizar o protagonismo dos personagens num *podcast* narrativo, busca-se ressaltar como esses papéis podem ser flexíveis.

A questão da autoria e como se relaciona com a análise interseccional da Comunicação segue um campo aberto de investigação, que vai além da expectativa de contribuição deste artigo. Portanto, a ambivalência sobre a autoria vinculada às escolhas enunciativas, influenciadas pela tecnologia da comunicação, não será aprofundada.

Por fim, ao esmiuçar o tópico “Negociações interseccionais”, Carrera coloca em primeiro plano a bidirecionalidade do Instagram como plataforma, analisando a interação das influenciadoras com seus seguidores, bem como a maneira como ressignificam as mensagens recebidas e as transformam em conteúdo. Aqui, materializa-se um desafio para um produto como o *podcast*, que é, à primeira vista, unidirecional e só permitiria essa troca se estivesse inserido em uma proposta transmidiática, o que não se enquadra nesta análise, interessada em compreender o produto singular e não o ecossistema de circulação do *podcast* em outras redes.

Uma leitura possível sobre a interação com o público seria a de que a própria experiência imersiva – definida como “capacidade de transposição da consciência para um outro ambiente seja imaginado ou sinteticamente criado” (Cordeiro & Costa, 2016, p.100) – também

apresenta-se como uma característica inerente aos *podcasts* narrativos pela potencialidade que têm de convidar o outro a imaginar, participando ativamente do processo.

É importante lembrar que o áudio, por si só, também poderia se equiparar à imersão trazida por plataformas multimídias que dispõem de tecnologias sofisticadas, como os aparatos que permitem que o usuário tenha uma visão em primeira pessoa e em 360º de determinado acontecimento. A combinação de um bom equipamento com uma apurada arquitetura de som pode transportar o ouvinte para a narrativa e o local desejado (Maia & Capistrando, 2019, p. 10).

Nesse cenário, tanto o uso do *storytelling* na escrita do roteiro quanto a existência de recursos tecnológicos de áudio mais sofisticados permitem a vivência de uma experiência multissensorial e profunda de cada história, sendo que, nessa proposta, o ouvinte pode formular imagens durante a experiência da escuta do *podcast* e, conseqüentemente, torna-se mais ativo na construção de uma versão singular do produto em questão, permeada, evidentemente, pelo próprio repertório de vida.

3. Resultados

3.1 Descrição

A Rádio Batente é uma produção da Rádio Novelo para a ONG Repórter Brasil. Fundada em 2001, tal organização tem o objetivo de “fomentar a reflexão e ação sobre a violação aos direitos fundamentais dos povos e trabalhadores no Brasil” (BRASIL, ONG Repórter, [s.d.]).

Nesse cenário, o podcast tem o trabalho como tema central, sendo que o artigo analisa os cinco episódios da primeira temporada. Nela, o ouvinte acompanha um dia na vida de um servidor público (ou trabalhador terceirizado), atuante em segmentos considerados essenciais.

São personagens, na ordem de publicação dos episódios: 1. Um professor de artes da rede pública municipal paulistana; 2. Uma assistente social que trabalha em uma casa de acolhimento para homens em situação de rua, em Campinas (SP); 3. Uma dupla de profissionais da saúde (uma enfermeira e um agente comunitário de saúde), vinculados à uma Unidade Básica de Saúde paulistana; 4. Uma cobradora de ônibus que trabalha em uma linha entre uma região periférica de São Paulo e a região central; 5. Um bombeiro que atende o centro da capital paulista.

Os episódios foram publicados entre os dias 1º e 29 de julho de 2020, com periodicidade semanal. Um episódio padrão segue de maneira linear o dia do profissional, sendo que um jornalista vai “a campo”, enquanto outro comenta as situações apresentadas. Vale pontuar, também, que os episódios foram gravados antes da pandemia, mas há uma atualização, feita por chamada telefônica, videochamada ou mensagem gravada, durante a quarentena, e que foi incluída nos episódios justamente para fazer um contraponto entre a rotina usual de trabalho e a jornada profissional no cenário de isolamento social.

3.2 Análise e discussão

3.2.1 Primeira etapa da metodologia da roleta interseccional

Para a análise da primeira temporada, foi feita a transcrição dos cinco episódios da primeira temporada do *podcast*, sendo que a investigação foi pautada nas hastes propostas pela roleta interseccional, considerando-se não apenas a presença das palavras que nomeiam os marcadores e/ou sinônimos, mas, principalmente, a análise do conteúdo de maneira mais aprofundada, a partir dos recortes conceituais de cada marcador proposto por Carrera. Como estratégia de codificação analítica, foi utilizada a minutagem, tendo as falas dos personagens como marco temporal. A tabela abaixo, mostra as hastes da roleta que foram iluminadas em cada episódio, iniciando a análise a partir da metodologia proposta por Carrera.

Tabela 1: Tabela relativa à primeira etapa da roleta interseccional

Fonte: Autores

Hastes/ Episódios	Gênero	Raça	Classe	Peso	Idade	Deficiência	Geolocalização	Sexualidade
1º EPISÓDIO O professor artista			•	•		•	•	
2º EPISÓDIO O trabalho da assistente social com moradores de rua	•	•	•					
3º EPISÓDIO A saúde bate na porta	•		•					
4º EPISÓDIO A cobradora simpatia	•		•				•	
5º EPISÓDIO Os corres de um bombeiro			•				•	

Nos cinco episódios, a haste transversal é a classe, uma espécie de espinha dorsal do programa. Essa constatação não surpreende, já que o foco do *podcast* é o trabalho. A haste de classe dá-se em dois sentidos: pela categorização social dos profissionais, mas, principalmente, quando se faz referência às populações atendidas e ao contexto de funcionamento dos serviços em destaque. O elemento de classe é, portanto, a posição que estrutura a posição social da pessoa entrevistada e a sua relação com outras pessoas.

Ao reconhecer a classe como marcador fundamental na construção das subjetividades e dos limites simbólicos que perpassam as análises comunicacionais, é preciso, no entanto, atestar que este marcador opera como estrutura fundamental atravessada e impregnada por outras diferenças. Isto é, há modos de exploração que ultrapassam a dimensão econômica e não podem ser explicados apenas a partir desta inscrição de desigualdade (Gonzalez, 2018). Para a Comunicação, entretanto, perceber a influência da classe na construção dos sentidos é compreendê-la como fundamental tanto para os limites interacionais que estabelece quanto para as apropriações dos sujeitos a partir da linguagem e dos comportamentos sociais (Carrera, 2021b, p. 5).

No primeiro episódio (O professor artista), as hastes de geolocalização, peso, deficiência e classe são iluminadas. O marcador de geolocalização é proeminente no ambiente escolar, já que entre 45% e 55% dos estudantes são filhos de imigrantes, majoritariamente, de origem boliviana, de acordo com dados disponibilizados no episódio. Ganha força o relato sobre o histórico de xenofobia, marcado pela violência

física e verbal contra esses alunos estrangeiros – chamados de “o boliva”, “o boliviano” – um quadro que foi revertido satisfatoriamente pela inserção da dança tradicional boliviana no espaço escolar, numa proposta em que a arte torna-se ferramenta de aproximação, abrindo caminho para que as diferenças possam ser acolhidas.

De antemão, é importante destacar que a xenofobia afeta a maior parte de grupos migrantes, mas, ainda assim, deve-se mencionar a existência de uma questão de interseccionalidade. Não se pode considerar que todos os grupos enfrentam a xenofobia do mesmo modo, ou seja, diferentes fatores devem ser levados em consideração ao analisar a xenofobia contra determinado grupo, já que características como origem geográfica, cultura, gênero, cor, etnia, classe social e religião afetam a recepção desses estrangeiros nos países de destino (Peixoto., et al., 2021, p. 7).

O uso pejorativo de “apelidos” no ambiente escolar pode ser análogo ao uso de expressões como “baiano”, “paraíba”, que acompanham muitos dos migrantes nordestinos que chegam ao sudeste/sul. Evidentemente, essa homogeneização que aparece com a supressão do nome próprio desumaniza e apaga a singularidade de cada criança.

No Brasil, pessoas negras estrangeiras podem acionar lugares de privilégio e vantagem a depender do cenário, pela lógica da colonialidade que direciona muitas práticas e discursos vigentes aqui. Além disso, não se pode subestimar os impactos dos colonialismos internos (QUIJANO, 2002), que transformam as experiências e as dinâmicas de opressão vivenciadas por pessoas negras nortistas e nordestinas quando estão localizadas

em estados que representam o lócus do poder econômico brasileiro. Os mesmos sujeitos, portanto, podem protagonizar situações comunicacionais diversas, tanto de opressão quanto de privilégio, a depender da conjuntura que se estabelece ao seu redor (Carrera, 2021b, p. 3).

Já a haste relacionada ao peso evidencia uma possível lacuna não abarcada pelos marcadores propostos por Carrera, considerando-se que tal marcador aparece atrelado à saúde mental.

Então, a gente está um pouquinho preocupada com ele [aluno], no sentido dele tá um pouco deprimido. Vocês comentaram no conselho que ele tem dificuldade de interação com os colegas, ele está obeso e, aí, a Rosane falou para mim que, no intervalo do (inaudível), que ele pegou a corda que eles brincam de...ele simulou um laço para colocar na cabeça “assim”. Aí ela falou: “nossa, por que você tá fazendo isso né?” Aí, ele falou assim “Ah, é porque eu me sinto muito sozinho, eu não tenho nada nesse mundo”. (Sonora Professora, personagem do episódio 01: 33’37- 34’08”).

Se por um lado, a sonora reverbera uma percepção de Carrera (2021b) sobre a representação midiática do corpo gordo, por outro a questão do peso, nessa situação, parece estar relegada a um segundo plano, evidenciando o sofrimento psíquico como uma problemática que não se enquadra em nenhum marcador específico da roleta interseccional, já que não deixa, *a priori*, um registro corporal, embora possa perpassar os demais marcadores.

Se não é relegado ao lugar do riso, o sujeito gordo está ligado simbolicamente ao mal-estar e ao isolamento social. Estas associações são construídas pelos discursos da oralidade, mas, também, reforçadas pelas imagens midiáticas, jornalísticas e publicitárias. É, portanto,

fundamental reconhecer que o peso adquire, na conjuntura sociocultural e histórica, um lugar significativo na produção das subjetividades, uma vez que constrange os limites possíveis de existência e desloca o sujeito da normatividade para a marginalização. (Carrera, 2021a, p.6)

Também é oportuno pontuar que, no primeiro episódio, o marcador da deficiência aparece por meio da informação de que existe um projeto pedagógico inclusivo na instituição de ensino. Por fim, a haste da classe fica evidente quando o protagonista aponta que parte significativa dos alunos se encontra em situação de vulnerabilidade socioeconômica, sendo que o professor nomeia alguns dos bairros de moradia das crianças, o que mostra como a própria ocupação da cidade repercute as disparidades sociais e econômicas.

No segundo episódio (O trabalho da assistente social com moradores de rua), as hastes classe, gênero e raça se iluminam. A questão de classe é contundente na delimitação do público-alvo do serviço: homens vindos da rua, marcados pela miséria socioeconômica e não pertencentes à engrenagem produtiva que, de certa forma, nos confere identidade social.

Já a questão de gênero está presente na própria delimitação do público acolhido pelo serviço em pauta, mas que também é explicitada quando a assistente social, protagonista do episódio, relata o manejo das situações de assédio em um ambiente, predominantemente, masculino. O fato dela ser negra não é verbalizado no programa, ainda que a informação fique evidente pela foto de capa do episódio. Esse seria o único momento em que a tríade gênero/raça, tão cara à interseccionalidade, ganha forma em todo o podcast, ainda que essa relação seja expressa nas entrelinhas.

Em alguns momentos eu tive que me impor. Em outros, que eu, acho que isso é muito mais o que acontece de, assim, eu sou a Gisele, assistente social do serviço. Só isso dá um limite. E não

é só com o morador né porque isso é estrutural (...). Então, assim, de pessoas que chegam em mim e diz assim ‘nossa, mas eu amo você’ e vem para pegar no meu cabelo e aí eu digo olha, então, peraí, vamos conversar, o limite é esse aqui, eu tô aqui, eu sou profissional, a nossa relação não passa disso’. Fica mais fácil, do que você mascarar (Sonora Assistente social, protagonista do episódio 02: 30’33” - 31’47”).

Nesse contexto, uma reflexão interessante pode ser feita para a compreensão que entrelaça gênero e raça.

(...) tanto racismos quanto sexismos podem ser entendidos como fenômenos complexos de poder que operam no contexto de atribuição de diferenças categoriais. Mesmo que não seja sempre necessariamente assim, eles frequentemente funcionam por meio de referências a características corporais e, portanto, por meio de referências a supostas certezas biológicas. É por isso que atribuições de diferença de cunho racista ou sexista são geralmente atribuições de diferenças naturalizadas que exigem validade atemporal ou pelo menos por longos períodos de tempo. Nesse sentido também as formas racistas e sexistas de poder são diferentes daquelas que operam vinculadas a relações de classe ou de produção (Kerner, 2012, p. 36).

Aliás, um dos pontos que precisa ser evidenciado é a quase ausência de referência à raça nos episódios como um todo, um fator que ganha significação considerando-se a própria origem da interseccionalidade como campo do saber. Ainda que seja possível identificar traços negros nas fotos das capas dos protagonistas de cada episódio, é no discurso que essa falta se torna perceptível. Há uma ausência sobre

o tema tanto quando é feita uma definição do público-alvo dos serviços, como quando os personagens falam de si.

No cenário apresentado, algumas questões podem ser colocadas: a especificidade do meio podcast, em tese, também poderia “enfraquecer” o marcador da raça pela falta de referências visuais nesse objeto de estudo. Já a ênfase da Rádio Batente no universo do trabalho poderia, de antemão, supervalorizar a haste classe” desde a sua concepção. Diante dos pontos levantados, fica a pergunta: a análise interseccional é válida quando a haste da raça não é intensamente iluminada? Essa é uma questão abordada em referência bibliográfica recente.

Colocar as marginalizações no centro do debate é incluir, conseqüentemente, todas as pessoas, entendendo que a abertura da porta para sujeitos em situação de vulnerabilidade facilita a entrada tanto para estes indivíduos quanto para aqueles que não apresentam desvantagem alguma. Fica evidente, então, que os estudos interseccionais, mesmo vinculados à base da sua construção teórica, não precisam ou devem estar restritos a análises em gênero e raça, ao contrário, podem articular outras avenidas identitárias que direcionam os sujeitos para experiências complexas na vida social e política (Carrera *et al.*, 2022, p. 16).

Ainda ao refletir sobre como o conceito da interseccionalidade pode alçar voos distantes do ponto de partida conceitual, ressalta-se:

O processo de explorar e examinar experiências complexas, portanto, é entender que análises da vida social devem resistir à generalização simplificada, embora sua completa negação não seja o caminho ideal, sobretudo quando se pensa em ações políticas estratégicas.

Entende-se que interseccionalidade surge como uma abordagem “intracategórica” (MCCALL, 2005, p. 1773), ao basear seu argumento nos desvios das categorias de gênero e raça, mas termina sendo um modelo teórico “intercategórico”, uma vez que adota categorias em um primeiro momento, justamente para documentar suas desigualdades e perceber as existências que se entrecruzam, mas não se restringe a nenhuma categoria particular (Carrera *et al.*, 2022, p.16)

Nesse escopo de indagações, também é oportuno questionar se a expectativa de uma correlação direta entre vulnerabilidade econômica e a raça negra, por exemplo, não seria um preconceito do pesquisador ou, ainda, um desdobramento natural diante do fato de o Brasil ser um país historicamente alicerçado em relações escravocratas, sendo que esse legado repercute nas estatísticas sociodemográficas. Nesse ponto, mais do que respostas, é importante pensar que cada objeto não só interroga os limites da própria metodologia, mas também implica no questionamento e possível reconhecimento de qualquer viés que esteja atrelado ao pesquisador.

No terceiro episódio (A saúde bate na porta), as hastes classe e gênero são iluminadas. Ao abordar a haste classe, fica evidente como as dificuldades financeiras impactam o cuidado com a saúde, extrapolando uma percepção de que essa seria uma responsabilidade exclusivamente relacionada às escolhas e hábitos do sujeito. Ainda que o episódio mostre situações individuais, é oportuno colocar em pauta o fato de que, mesmo que existam fatores comportamentais em jogo no cuidado em saúde, é preciso ter condições dignas de vida para fazer escolhas saudáveis.

As desigualdades no estado de saúde estão de modo geral fortemente atreladas à organização social e tendem a refletir o grau de inequidade existente em cada sociedade. O acesso e a utilização dos serviços refletem também essas diferenças, mas podem assumir feições diversas, dependendo da forma de organização dos sistemas de saúde. Há sistemas que potencializam as desigualdades existentes na organização social e outros que procuram compensar, pelo menos em parte, os resultados danosos da organização social sobre os grupos socialmente mais vulneráveis (Barata, 2020, p. 20).

Também é oportuno ressaltar que a ausência do Estado como garantidor de serviços fundamentais reforça a importância da solidariedade na comunidade, criando uma rede informal de cuidado, por exemplo, diante do impasse entre precisar trabalhar e, concomitantemente, permanecer em casa para cuidar de um parente que precisa de tutela ou perdeu a autonomia para a realização de atividades cotidianas e de autocuidado (crianças, idosos acamados, entre outras condições marcadas pela vulnerabilidade).

Nesse episódio, a haste do gênero também é iluminada quando a jornalista chama a atenção do ouvinte para o fato de que existe apenas um homem, o agente de saúde, na equipe da Unidade Básica de Saúde. Ao longo do episódio, a comunicadora também nota que, nas visitas domiciliares, são sempre as mulheres que recebem os profissionais. Nesse contexto, pergunta-se: seria o cuidar um verbo relacionado ao feminino? Ao analisar o trabalho doméstico remunerado e o trabalho na área da saúde, uma das referências aponta:

Ambas as dimensões do trabalho de cuidado (doméstico e externo ao trabalho doméstico) têm duas características em comum: são atividades exercidas em sua ampla maioria por mulheres, na mesma medida em que são atividades geralmente desvalorizadas em um cotidiano de desigualdades estruturais e precariedades múltiplas, que se acentuam a depender de outros indicadores interseccionais de classe e raça. Em que pese a dinâmica de alterações sociais profundas e também nos modos de produção capitalista, há um continuísmo em naturalizar a responsabilidade à mulher no trabalho de cuidar, por primazia ou exclusividade (Borges, 2020, p. 300).

No quarto episódio (A cobradora simpatia), as hastes de gênero, classe e geolocalização são iluminadas. Ao abordar o primeiro marcador, é nítido como o machismo está entranhado nas relações sociais, inclusive no ambiente de trabalho: a cobradora de ônibus ilustra como o problema marca a relação com os colegas atuantes na área de transportes – um contexto majoritariamente masculino – e também a interação com os passageiros. No primeiro caso, a profissional relata que, no passado, havia uma crença generalizada de que as trabalhadoras do setor estariam ali por serem amantes de alguém importante na hierarquia da empresa ou por terem feito o “teste do sofá”.

Aponta, ainda, que já ouviu motoristas homens de outras linhas perguntarem ao condutor da linha em que trabalha se ela já estaria no “nome dele”, numa alusão à posse. Além disso, a personagem cita passageiros que não se aproximam muito por acharem que ela “seria do motorista”. Em uma reivindicação coletiva das trabalhadoras do setor para que terminassem seus turnos até às 22h e, assim, pudessem retornar para casa em um horário seguro, chegou a ouvir dos colegas: “Ah, mas vocês não

lutam pela igualdade de direitos? Cadê a igualdade de gênero?”.

Nesse mesmo episódio, as hastes classe e geolocalização podem ser analisadas conjuntamente. O ouvinte acompanha um dia na linha de ônibus que faz o trajeto entre o Jardim Ângela e o Terminal Santo Amaro, na zona sul paulistana. Embora não haja uma explicação sobre o perfil sociodemográfico do bairro, fica evidente, pelo tempo do trajeto, que o ônibus parte de uma região periférica para uma região mais central de São Paulo. Em uma das viagens, em que o ônibus é ocupado, majoritariamente, por mulheres, a cobradora salienta que boa parte delas são empregadas domésticas e que, em muitos casos, são as principais mantenedoras da família. Mais uma vez, pode-se pensar como a ocupação da cidade segue uma lógica excludente e como o ambiente doméstico está associado à mulher, no âmbito pessoal e profissional.

O quinto episódio (Os corres de um bombeiro) é o mais desafiador a ser analisado. De imediato, o que determina as ocorrências atendidas é a localização geográfica, já que existe um perímetro a ser coberto pela base em que o profissional está alocado. Há, ainda, uma informação sobre tentativas de suicídio na região central paulistana em decorrência da presença de viadutos.

Pode-se fazer, também, uma inferência sobre classe quando o personagem protagonista salienta que o carro dos bombeiros pode ficar preso na emergência de um hospital público porque a instituição não tem macas próprias para a transferência imediata do paciente. Será que isso aconteceria em um hospital particular? Difícil de imaginar, mas o conteúdo do episódio em si não permite fazer afirmações taxativas, considerando-se que esse serviço é uma responsabilidade exclusiva do poder público.

3.2.2 Segunda etapa da metodologia da roleta interseccional

Após a identificação e a descrição das hastes iluminadas, parte-se para a discussão dos três eixos já apresentados:

3.2.2.1 Formação interseccional-discursiva

A formação interseccional-discursiva coloca em evidência a conjuntura sociocultural e histórica do objeto em análise (Carrera, 2021b). No podcast Rádio Batente, a própria escolha dos serviços e dos protagonistas dos episódios – trabalhadores da rede pública – aponta para a preocupação de colocar em pauta experiências bem-sucedidas de assistência e cuidado a grupos sociais usualmente marginalizados. Nesse contexto, o projeto tem um caráter de denúncia e, concomitantemente, mostra-se como uma ferramenta de resistência e empoderamento.

A denúncia, entretanto, em nada tem um caráter panfletário. Do ponto de vista editorial, a seleção criteriosa dos serviços permitiu dar voz a personagens que refletem continuamente sobre a atuação profissional e que reconhecem a diversidade como uma condição a ser abraçada. Profissionais que são comprometidos em desenvolver a autonomia e a participação dos públicos atendidos nos processos decisórios que impactam a vida desses sujeitos.

Denúncia, também, por desnudar como a falta de políticas públicas ou de investimentos fragiliza áreas prioritárias, o que ficou ainda mais evidente pelo contraponto feito entre o cotidiano antes e durante a pandemia da COVID-19, colocando em primeiro plano tanto

o impacto da crise sanitária nas populações atendidas como na configuração do trabalho de quem atuou na linha de frente. Nesse ponto, é fundamental evidenciar a pandemia como um evento disparador universal, circunscrito no tempo e no espaço, pois permite assinalar as assimetrias sociais.

A crise da Covid-19 escancara a realidade de um sistema econômico, que, na prática, é fundado por um modelo extrativista e predatório dos recursos do planeta e que produz e sustenta desigualdades em múltiplas dimensões e concentração de riquezas em escalas abissais. Mas não apenas isso, é um sistema que alimenta e retroalimenta desigualdades estruturais em relações assimétricas de poder naturalizadas e, portanto, invisibilizadas pela sociedade (Borges, 2020, p. 267).

É preciso pontuar, ainda, uma dualidade ou uma contradição na estruturação narrativa singularizada e quase heróica adotada nos episódios. Nela, há um herói e/ou uma heroína (os profissionais que são protagonistas em cada episódio), que centralizam valores positivos norteadores das jornadas de trabalho. Uma vivência que pode ser análoga à própria Jornada do Herói, proposta por Joseph Campbell para entendimento das narrativas, a partir da mitologia e dos arquétipos junguianos.

Em princípio, apesar da variedade infinita, a história de um herói é sempre uma jornada. Um herói abandona seu ambiente confortável e comum para se aventurar em um mundo desafiador e desconhecido. Pode ser uma jornada ao exterior, a um lugar de verdade (...). No entanto, existem muitas histórias que conduzem o herói a uma jornada interior(...). Em qualquer boa história, o herói cresce e se transforma,

empreendendo uma jornada de um modo de ser para outro: do desespero à esperança, da fraqueza à força, da tolice à sabedoria, do amor ao ódio e vice-versa. São essas jornadas emocionais que prendem o público (Vogler, 2015, p. 45).

Em contrapartida a essa jornada solitária e mítica, cada profissional dá voz a outros tantos servidores públicos anônimos que, numa atuação coletiva, têm vossas práticas bem alicerçadas na Constituição brasileira de 1988 e nas diretrizes que sustentam o SUAS (Sistema Único de Assistência Social) e o SUS (Sistema Único de Saúde), por exemplo. De certa forma, esses dois sistemas são os verdadeiros protagonistas do segundo e do terceiro episódios, respectivamente, pois, a partir das trajetórias singulares dos profissionais, é possível conhecer os serviços e compreender os alicerces teóricos e práticos que os sustentam a partir do que a lei brasileira estabelece.

Nesse contexto, há de se considerar o aspecto multifacetado do conceito de resistência que, por um lado, significa estabelecer uma rede de sociabilidade mais adequada no microcosmo. E, por outro, implica em compreender um panorama amplificado que coloca em voga a disputa entre o modelo neoliberal, que encolhe o papel governamental na garantia de direitos, e o modelo de Bem-Estar Social, que assegura o protagonismo governamental nessa seara.

Como exemplos de iniciativa em pequena escala, pode-se destacar um trecho no primeiro episódio (O professor artista). Nele, o docente comenta a escolha da escola em adotar a música como “sinal” que avisa sobre o começo e o final do dia letivo e/ou marca os intervalos entre as aulas, em detrimento de um apito estridente, usado usualmente. Pontua, também, a restrição que ele tem em formar filas como maneira de organizar o deslocamento das crianças pela instituição:

(Versão instrumental de música dos Rolling Stones toca ao fundo da sonora, como “sinal”).

Aqui foi uma coisa muito bacana que desde que eu cheguei que já não existia, aquele sinal de fábrica, sabe? Que já vai treinando os nossos alunos para essa realidade que é o chão da fábrica. Então aqui nós usamos música na hora do sinal e eu acho muito mais agradável. Eu sou um professor que eu não faço fila, por não acreditar na fila, e não dizer que ela não seja necessária, mas aqui eu faço um acordo com os alunos, eu falo “o fato de não ter fila não quer dizer que você vai sair gritando, pulando, se jogando no chão (...) até essa desconstrução é difícil (...) porque eles já estão condicionados (...)” (Sonora Professor de artes - protagonista do episódio 01: 18’18”-19’).

Ao se pensar no microcosmo, também é oportuno citar o segundo episódio (O trabalho da assistente social com moradores de rua). Pode-se acompanhar uma instituição que apresenta uma gestão horizontalizada e participativa, pautada pelo diálogo e alicerçada em três pilares: a educação popular, de Paulo Freire; a Comunicação Não-Violenta, proposta por Marshel Rosenberg; e o conceito de justiça restaurativa, em detrimento da justiça retributiva, que tem um viés punitivo, como detalhado no próprio episódio do *podcast* (Casteli, T.; Suzuki, N., 2020).

Esse arcabouço teórico foi materializado, por exemplo, na realização semanal de uma assembleia, ocasião em que os moradores e os profissionais podem decidir coletivamente as regras de convívio na casa, o que inclui discutir sobre o horário-limite de chegada na casa, o horário de funcionamento de TV, a escolha do cardápio servido aos moradores e as condições para receber visitas vindas de fora. Vivências que, outrora, aconteciam de forma impositiva por parte da direção do serviço, sendo que os próprios profissionais teriam repensado a forma de estabelecer regras.

Tá, mas aí, o que coloca no lugar?” O que coloca no lugar é o diálogo. E se o problema era chegar tarde da noite, qual é esse problema? Então, vamos dialogar sobre isso (...). Aí, a gente começou a perceber que o foco não era a gente querer que ela [pessoa] cumprisse regras, a gente queria que ela pudesse olhar para o outro e olhar para si. E que o abrigo pudesse propiciar isso. E aí, quando a gente começou a substituir as regras pelo diálogo, a gente começou a perceber que a própria casa passava a ser transformadora (Sonora Psicólogo, personagem do episódio 02: 17’48”- 18’23”)

Como contraponto, para um questionamento amplo dos modelos socioeconômicos, o episódio relacionado à saúde faz um registro apropriado sobre a presença/ausência do estado durante a pandemia da COVID-19, tendo como cenário a favela paulistana de Heliópolis. Nesse contexto, ficou evidente que a orientação de “ficar em casa” e manter o maior distanciamento possível entre as pessoas era incompatível com a necessidade de trabalhar, numa realidade marcada pela grande circulação de pessoas e pela abertura do comércio informal diante da falta de fiscalização.

Também merece destaque a parceria entre os equipamentos de saúde e as ONGs para a realização do trabalho de conscientização, apontando, concomitantemente, para agilidade da sociedade civil em se articular para o enfrentamento da situação e para a morosidade do Estado em viabilizar algum respaldo socioeconômico em grande escala diante de um contexto crítico.

3.2.2.2 Ethos interseccional

Além das questões de autoria, pontuadas anteriormente, uma possibilidade para conhecer a imagem de si e do *ethos* discursivo pertinente ao *podcast*, seria compreender que os protagonistas dos episódios, de certa forma, são portadores das respectivas profissões e, mais especificamente, dos serviços em que estão inseridos.

Nos episódios, os jornalistas buscam compreender a função social atrelada ao cargo dos protagonistas. Em alguns momentos, essa resposta fica restrita à descrição de tarefas relacionadas aos cargos ocupados. Em outros, há uma reflexão por parte do profissional e/ou do usuário do serviço sobre o papel daquele trabalhador e de como ele impacta a comunidade que atende. Isso implica em reconhecer não só as experiências individuais, como, também, vislumbrar as representações sociais inscritas no imaginário coletivo.

Os estudos sobre a representação social têm sido crescentes desde sua publicação, na tentativa de explicar e compreender fenômenos sociais, por meio de situações comuns do cotidiano, que quando passados de uns para outros se tornam verdades. São conhecimentos naturais dos indivíduos que são transmitidos de geração para geração, são opiniões e ideias de pessoas de um grupo sobre determinado fenômeno ou acontecimento. A representação é um conhecimento real, formado pelo senso comum entre um grupo de indivíduos, o conhecimento se torna uma representação a partir do momento que se torna uma verdade para todos os membros que compõem o grupo (Jodelet, 2001) (Sousa & Souza, 2021, p.1).

A partir desse preâmbulo, alguns trechos dos episódios podem ser destacados por colocarem em pauta a percepção dos personagens sobre as suas profissões.

A outra coisa que as aulas de artes ajuda (sic) muito. (...) aqui, ele tem que aprender que se ele misturar azul com amarelo vai dar uma terceira cor que pode ser o verde, mas se ele colocar uma outra... então, é a descoberta, é a autonomia, a coordenação fina, é saber que não está sozinho no mundo. Então, é uma aula que eu acho extremamente necessária e qualquer pessoa que queira comprometer ou tirar esse tipo de linguagem de uma criança, pode-se considerar crime, enfim. (Sonora Professor de artes, protagonista do episódio 01: 17'04" - 17'48").

Mas a função do assistente social aqui no abrigo, ela é uma função muito importante porque auxilia os moradores a conseguirem acessar o que é de direito, que muitas vezes é desconhecido por eles né (...). Então, às vezes, a pessoa tem uma doença mental, ela tá sem retaguarda familiar, ela não tem retaguarda financeira de ninguém. Quando ela vem para o abrigo, assim, o assistente social tem essa função de saber antes quais são os direitos que a pessoa tem e, no caso, inserir. Às vezes, eu escuto 'Ah, a assistente social tem que ser boazinha, tem que dar cesta básica', E, na verdade, é assim, o trabalho do assistente social, ele tá muito pautado nos direitos mesmo né que a pessoa tem e que, às vezes, ela desconhece, não acessa a informação. Então, muita gente, às vezes, não sabe qual o caminho que precisa percorrer para ter acesso ao direito. E o assistente social faz essa ponte e, também, acho que problematiza a condição que ele tá, auxilia ele a entender a condição social que ele tá porque aí muda a visão de mundo né" (Sonora Assistente social, protagonista do episódio 02: 20'19"-21'-19").

3.2.2.3

Negociações interseccionais

Ao colocar a imersão como ponto de partida para a inclusão do ouvinte na experiência narrada, é oportuno ressaltar que os episódios lançam mão de sonoridades externas de maneira maciça e não só durante as entrevistas pontuais, mas também mantendo a captação contínua nos deslocamentos, nas conversas entre membros da equipe técnica, nas interrupções feitas por curiosos e, ainda, com o intuito de captar o som ambiente, um conjunto de situações que imprimiu uma certa familiaridade às situações em destaque.

Vale pontuar, ainda, que a postura dos jornalistas reforça a percepção de que são testemunhas, emprestando os seus sentidos ao ouvinte que não está de corpo presente. Adota-se, assim, uma certa ingenuidade durante a cobertura, o que os libera da condição de sujeitos do suposto saber e os deixa livres para indagar cada detalhe, evidenciando que não pertencem ao universo mostrado, aproximando-os dos ouvintes que também estão adentrando num universo desconhecido. Há, assim, uma similaridade com outros *podcasts* narrativos, que, em muitos momentos, deslocam o repórter do centro da narrativa.

Esse reposicionamento inverte os papéis do repórter, do entrevistado e do ouvinte, intercalando os seus papéis, e tornando-os, ao mesmo tempo, narradores e protagonistas das histórias contadas. Por meio delas, podemos presentificar o passado, acessar o presente e ousar novos voos na criação de narrativas radiofônicas que permitem deslocar o senso comum e mobilizar o sentido de comunidade que, afinal, é o que move nosso desejo

de, incansavelmente, contar e recontar histórias (Vicente & Soares, 2020, p. 18).

Outra questão que aproxima a vivência dos jornalistas da experiência do ouvinte é a forma como se colocam: usam a primeira pessoa do singular, dividem as suas percepções sobre as situações que vivenciam, compartilham experiências anteriores em contextos similares aos mostrados no *podcast*, além de exporem as suas vidas particulares ao perceberem que essa escolha poderia enriquecer o enredo. Embora a proposta dos *podcasts* narrativos implique em uma jornada pouco afeita à improvisação, considerando-se desde o trabalho de pesquisa na pré-produção até a edição, há, nessa primeira temporada do podcast, espaço para o ineditismo dos encontros e para a surpresa diante dos imprevistos.

Nesse contexto, existe, também, um vislumbre de como as negociações para as realizações das captações aconteceram, um material que poderia ser excluído do produto que chega ao ouvinte. Em uma analogia, seria como se o ouvinte pudesse espiar por detrás da cortina enquanto o elenco se prepara para entrar em um palco no teatro. Isso acontece, por exemplo, no segundo episódio, que relata o cotidiano dos antigos moradores de rua: a informação de que eles assinaram um termo de consentimento e que a participação no *podcast* passou por votação na assembleia realizada no abrigo foram verbalizadas no programa. Nesse ponto, o fato de, no geral, somente o produto final ser acessível, sem que se conheça os meandros da produção, não permite que se faça apontamentos mais incisivos sobre esse tópico.

4. Conclusões

A análise empírica empreendida a partir da roleta interseccional mostrou-se adequada, considerando que a metodologia apresentou as ferramentas necessárias para a investigação do objeto comunicacional escolhido, servindo como uma bússola que direciona sem, entretanto, impedir o livre navegar do pesquisador.

É possível assinalar, também, que as possíveis limitações – disparidades entre o Instagram e o *podcast* como mídias e o fato de os eixos de análise serem, majoritariamente, pautados no aspecto bidirecional da interação – tensionam os limites dessa análise, mas paradoxalmente, servem como força motriz para os questionamentos em relação ao aspecto interacional em um *podcast* narrativo como este em pauta.

Vale pontuar, ainda, que a própria interseccionalidade veio ao encontro da proposta desse *podcast*, que denuncia vulnerabilidades sociais, mas também empodera e resiste ao dar voz a profissionais e usuários de serviços que seguem no trabalho cotidiano e anônimo de fomentar a autonomia em contextos sociais onde o silenciamento costuma imperar.

Por fim, é necessário que novas investigações empíricas aconteçam no intuito de dialogar com o exercício realizado nesse texto, uma iniciativa que ganhou fôlego justamente quando Carrera se propôs a ultrapassar a descrição da diferença, instrumentalizando a investigação empírica por meio de uma contribuição original e alicerçada na escrita de tantas outras mulheres que adotaram a interseccionalidade, sedimentando esse campo de investigação teórico e da *práxis*.

Referências bibliográficas

- BARATA, R. B. (2020). *Como e por que as desigualdades sociais fazem mal à saúde*. Fiocruz.
- BORGES, M. J. R. (2020). O vírus e o invisível: a desigualdade de gênero e o trabalho de cuidado. *Revista Tribunal Regional do Trabalho da 3a Região* (Edição especial), 265–310.
- BRASIL, ONG Repórter. ([s.d.]). *Repórter Brasil, 20 anos*. <https://reporterbrasil.org.br/quem-somos/>
- CARRERA, F. (2021a). Para além da descrição da diferença: apontamentos sobre o método da roleta interseccional para estudos em Comunicação. *Liinc em Revista*, 17(2), 1-19. DOI: 10.18617/liinc.v17i2.5715
- CARRERA, F. (2021b). Roleta interseccional: proposta metodológica para análises em Comunicação. *E-Compós*, 24, 1–22. DOI: 10.30962/ec.2198
- CARRERA, F., Fernandes, P., & Vieira, E. (2022). Interseccionalidade e plataformas digitais: dimensões teórico-metodológicas de pesquisas em Comunicação. *revista Fronteiras – estudos midiáticos*, 24(1), 14–21. DOI: 10.4013/fem.2022.241.02
- CASTELI, T., & SUZUKI, N. (Anfitriões). (2020, julho). Jornadas: o professor artista (nº1) [Podcast áudio]. In *Rádio Batente*. Rádio Novelo. https://open.spotify.com/episode/0bN6bsvjAYjDoHwsqCZnTa?si=a0Z0_TfZQS-iaKJSb_roWg
- CASTELI, T., & SUZUKI, N. (Anfitriões). (2020, julho). Jornadas: o trabalho da assistente social com moradores de rua (nº2) [Podcast áudio]. In *Rádio Batente*. Rádio Novelo. <https://open.spotify.com/episode/4PgA9RcS01YHUerTnXwemO?si=I8n1CaCmQhGFGvMJBjNcOg>
- CASTELI, T., & SUZUKI, N. (Anfitriões). (2020, julho). Jornadas: a saúde bate na porta (nº3) [Podcast áudio]. In *Rádio Batente*. Rádio Novelo. <https://open.spotify.com/episode/2bzclu8v5U7B5AdyMYIxrP?si=zQUUhE-6SJSI-mkR-OKSXg>
- CASTELI, T., & SUZUKI, N. (Anfitriões). (2020, julho). Jornadas: a cobradora simpatia (nº4) [Podcast áudio]. In *Rádio Batente*. Rádio Novelo. https://open.spotify.com/episode/301FjRWSsDkapKfFAB6fqz?si=AO6yLUPEQ_GHw-vEpkP2Fag
- CASTELI, T., & SUZUKI, N. (Anfitriões). (2020, julho). Jornadas: os corres de um bombeiro (nº5) [Podcast áudio]. In *Rádio Batente*. Rádio Novelo. https://open.spotify.com/episode/7ikIeu7LTQJ3cCq3PMUyg?si=EPwAW_znT76eBvmLs7Pd7w
- COLLINS, P. H., & BILGE, S. (2021). *Interseccionalidade*. Boitempo.
- CORDEIRO, W., & COSTA, R. (2016). Jornalismo imersivo: perspectivas para os novos formatos. *Leituras do Jornalismo*, 2(6), 99–116.
- DIAZ-BENITEZ, M. E., & MATTOS, A. (2019). Interseccionalidade: zonas de problematização e questões metodológicas. In SIQUEIRA, I. R de S.; MAGALHÃES, B.; CALDAS, M. & MATOS, F. (ed.). *Metodologia e Relações Internacionais: debates contemporâneos: vol II* (pp. 67-94). PUC-Rio. [http://www.editora.puc-rio.br/media/Methodologias_e_RI_volume-2%20\(1\).pdf](http://www.editora.puc-rio.br/media/Methodologias_e_RI_volume-2%20(1).pdf)

- KERNER, I. (2012). Tudo é interseccional?: Sobre a relação entre racismo e sexismo. *Novos Estudos CEBRAP*, (93), 45-58. DOI: 10.1590/S0101-33002012000200005
- MAIA, L., & CAPISTRANO, I. (2019, Junho). Para mergulhar nas histórias, use fones de ouvido: Uma Análise do Programa “Vozes: histórias e reflexões”, da Central Brasileira de Notícias (CBN). In *XXI Congresso de Ciências da Comunicação na Região Nordeste* (1-14). Intercom – Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, São Luís, Maranhão.
- PEIXOTO, A.R, CAVALCANTI, C.M, & SENA, K. B. P (2021). A presença dos refugiados do século XXI no Brasil: uma leitura da xenofobia a partir do discurso da crise econômica. *Revista Direito, Estado e Sociedade*. Ahead of print. DOI: 10.17808/des.0.1493
- SIFUENTES, L. (2022). Perspectivas interseccionais de gênero, classe e raça: um mapeamento de estudos de Comunicação. *Mídia e Cotidiano*, 16(3), 7–27. DOI: 10.22409/rmc.v16i3.54630
- SOUSA, K. N. de, & SOUZA, P. C. (2021). Representação social: uma revisão teórica da abordagem. *Research, Society and Development*, 10(6), 1–12. DOI: 10.33448/rsd-v10i6.15881
- VICENTE, E., & SOARES, R. (2020, Junho). Rádio Ambulante e a tradição do podcast narrativo no radiojornalismo. In *XXIX Encontro Anual da Compós* (pp. 1-19). Compós – Associação Nacional dos Programas de Pós-Graduação em Comunicação, Campo Grande, Mato Grosso do Sul.
- VOGLER, C. (2015). *A jornada do escritor: estrutura mítica para escritores*. Aleph.

ARTIGO

Queering ai: uma epistemologia queer para a ia com uma perspectiva de maioria global

Umut Pajaro Velasquez

umutpajaro@gmail.com

Bacharel em Comunicação e mestre em Estudos Culturais. Atualmente, trabalha como pesquisador independente em questões relacionadas aos direitos digitais, à ética e à governança da IA.

QUEERING AI: Uma epistemologia queer para a ia com uma perspectiva de maioria global

Palavras-chave

IA
Gênero
Raça
Queer
Epistemologia

Resumo

A maneira como definimos "inteligência" e epistemologia¹ é uma resposta a uma forma de conhecimento. A IA, como um conjunto de novas tecnologias, reproduz isso, e a teoria queer, juntamente com a teoria racial crítica, mostrou como exclui os outros, incluindo aqueles tradicionalmente considerados femininos² (Adam, 1998), e também mostrou que essa inteligência tem sua equivalência com o conhecimento do homem branco³ (Ali, 2019). Assim, a IA se envolve em uma exclusão ou repressão sociotécnica mais ampla do conhecimento das mulheres e reifica uma conceituação de "inteligência" baseada em gênero e raça. Por outro lado, a IA oferece uma oportunidade de mudar as suposições sobre a epistemologia de gênero⁴ (Collett e Dillon, 2019). Por exemplo, as narrativas de inteligência "dura" e "suave" são frequentemente classificadas como masculinas

e femininas, respectivamente. Adrian Weller⁵ (2019) aponta que essa inteligência "dura", que engloba a lógica e a racionalidade, é mais fácil de ser reproduzida na forma tecnológica, o que reforça a ideia de que tudo é "inteligência", mas a inteligência "suave" e a solução criativa de problemas, a empatia, a negociação e a persuasão, qualidades que historicamente têm sido identificadas mais com as mulheres e têm sido fomentadas nelas, podem se tornar privilegiadas por serem difíceis de codificar⁶ (O'Connor, 2019). Quer se pense que a IA depende da incorporação de uma epistemologia masculina, quer se pense que a IA promete dar uma vantagem à epistemologia feminista, a IA está perpetuando e reforçando estereótipos binários de gênero, deixando para trás aqueles que estão fora dos conceitos normativos e binários de gênero. Por isso, propomos uma epistemologia queer para a IA com uma perspectiva de Maioria Global a fim de criar uma estrutura que permita uma linguagem comum em todo o ciclo de vida da IA.

QUEERING AI: A queer epistemology for ai with a global majority perspective

Keywords

AI
Gender
Race
Queer
Epistemology

Abstract

The way we define "intelligence" and epistemology⁷ response to a way of knowing. AI as a set of new technologies reproduces this, and queer theory alongside critical race theory have shown how excludes others, including those traditionally considered feminines⁸, and also shown that this intelligence has its equivalence to white man's knowledge⁹. AI thus engages in a broader sociotechnical exclusion or repression of women's knowledge and reifies a gender- and race-based conceptualization of "intelligence."

On the other hand, AI offers an opportunity to change assumptions about gendered epistemology¹⁰. For example, narratives of "hard" and "soft" intelligence are often classified as male and female, respectively. Adrian Weller¹¹ points out that this "hard" intelligence, which encompasses logic and rationality, is easier to

reproduce in technological form, which reinforces the idea that it is all "intelligence", but "soft" intelligence, and creative problem-solving, empathy, negotiation, and persuasion, qualities that have historically been identified more with women and have been fostered in them, can become privileged as it is difficult to codify¹².

Whether AI is thought to depend on embodying a male epistemology, or whether AI promises to give a feminist epistemology a head start, AI is perpetuating and reinforcing binary gender stereotypes, leaving behind those who are outside the normative and binary concepts of gender. Hence, we propose a queer epistemology for AI with a Global Majority perspective in order to create a framework that allow a common language in all AI life-cycle.

1. Introduction: the foundation of the ai life-cycle: the 4ds, design, development, deployment, and detection of biases

To have an effective queer epistemological framework, we need a methodology with a queer perspective, we need to answer how queers, trans vision, and narratives will be included in our way to design, deploy, distribute, and detect biases in AI? And more importantly, what is done to make these AI inclusive for queer, BIPOC people from the majority world¹³, just to mention? How are we taking into consideration all the biases faced every day just for living in a cis-male white society?

If this framework is going to be used by technologists, final users, policymakers, sociologists, activists, and everyone involved with making real the practical aspects of Artificial Intelligence, we must consider using it in all stages of the foundation of any AI: **design**¹⁴ (first stage, what kind of AI we want to do, and who is going to be beneficiary or user from it), **development**¹⁵ (coding and modelling the technology we want to do and how articulate with an ethical and regulatory framework and first stage of tests with the possible beneficiaries or users), **deployment**¹⁶ (testing and putting in contact to his final users in small, medium, or large scale in real life conditions according to necessities), and **detection of biases** (addressing and auditing the biases found when the technology has been deployed and create the necessary amendments in terms of regulations, code, and others related to make the AI less harmful and more beneficial)¹⁷ (4Ds). We must have into account the following aspects

according to the technology we are going to bring to the people: first theoretical aspects that respendo to which theories are part of it, and why?; second, methodological aspects which means how to apply these theories to practices and actions?, and, finally methods and tools, or in other words, which ones and how to use them according to the stage?

As we are asking designers, developers, technologists, and policymakers to be more fair, accountable, transparent, and ethical in setting the foundation for AI and making them less harmful for queer people, these aspects must include non-traditional theories to feed the datasets that are going to make it work¹⁸, new approaches to traditional social sciences methods and tools as interviews, focus groups, participatory design, and new combinations of methodological approaches from social sciences, humanities, art, mathematics, natural sciences, engineering, and others¹⁹.

Also, this is going to demand that all of us involved in this process think outside the box and bring queer²⁰, decolonial²¹, abya yalá theories²², MLP²³, NLP²⁴, neurosciences²⁵, trans theories from the Global Majority²⁶, body theories²⁷, embody theories²⁸, cultural studies²⁹ as a methodology, interviews, focus groups, participatory design, documentary, or case analysis, research on the field, law-making regulatory processes³⁰ and others. Altogether, creating an epistemological framework according to the current expectation of respect for human rights from AI, with this we are saying all human rights including the ones from queer people no matter where they come from or are.

Probably, this is not yet a definitive solution to making these technologies more beneficial for queer communities. Also, is understandable that reach this issue just from a theoretical framework could create biased results, that is why also in the practice of detection of biases we need to include the language that can transform this specific harmful stereotype into regulations,

formal definitions, and actions that can benefit not only queer people but all of us. After all, as someone once said: "everything exists in the law, and nothing stands outside of it."

2. Methodology: Making AI frameworks with a queer perspective

There is a need for research that analyses AI and that will impact gender equality, gender diversity, and/or queerness. So far, there has been little attention to interpreting these laws, policies, and theories through gender and queer lens, or indeed research into how these structures could be exploited to strive for gender equality and diversity.

Research could explore existing and emerging frameworks concerning AI, gender, and queerness. Specifically, such exploration could include, but would by no means be restricted to, policies, laws, research, ongoing frameworks, and social aspects surrounding two particular areas that are key in the 4Ds stages of an AI: data and privacy, and technological design.

As demonstrated in the research context, these areas are already being considered to AI more generally speaking, but would benefit from additional gender-based and queer -based perspective. These could be analysed through two mechanisms: (1) gender, race, queer theories, decolonial theories (2) a series of interviews with technologists, experts, and policymakers.

Firstly, a theoretical analysis could be used to consider how policy, legislation, and the use of queer theories from the Global Majority can facilitate AI to work for gender equality, and social equality more broadly speaking³¹. Secondly, the interviews would function as a way to gain mutual understanding between policymakers and technologists regarding

definitions of gender, and how vulnerable gender groups would be impacted by certain structural changes. Whittlestone et al.³² outline that knowledge of technological capabilities should inform our understanding of the ethical tensions, which will be "crucial for policymakers and regulators working on the governance of AI-based technologies". Collaboration between experts, policymakers, and technologists would enable the formation of frameworks that tackles the main issues in a thorough, accurate, and realistic manner.

Additionally, a set of guidelines for ongoing developments would outline certain standards to be upheld when designing and implementing, which both, directly and indirectly, impact issues surrounding AI and gender diversity. The practical element of these standards is of the utmost importance. They are not the same as an ethical framework that cannot be directly applied, but rather they would be specific, context-related, and therefore straightforward for policymakers, technologists, and others to implement.

Overall, there is a need for research that assesses how emerging and future epistemological frameworks³³ (Villani, 2018) are failing to establish gender equality, and how they could be altered to strive for social justice. In a few points, this should point to the following: first, to analyze current and emerging law and policy, ethical, and academic frameworks that impact the intersection of AI and gender diversity, especially those that are from the Majority World. Following by to outline specific recommendations for alterations to laws and policies, academic frameworks surrounding AI and gender diversity, as well as a set of research-based guidelines for ongoing developments. These would rigorously promote the enhancement of social justice and gender equality and diversity.

After that we need to collaborate with existing research projects, initiatives, policymakers, experts, users, designers, developers, and

technologists. This research aims to contribute to understanding what it means to be inclusive, especially concerning questions of gender diversity and queerness beyond the perspective of the Global North.

And finally it is also important to harness an intersectional approach and consider how these are impacting and shaping gender, as well as race, ethnicity, sexuality, social class, disability, and so on. This will aid the pursuit of shaping structures in a way that considers not only one aspect of identity that could be detrimentally impacted by AI but multiple, in a few words create mechanisms embedded into the different frameworks that can allow anybody that detects a bias can share that information to improve the technology already deployed, developed and designed.

3. Methods for research

3.1 Interviews

Interviews would be conducted with researchers, designers, developers, technologists, and policymakers who are working in the relevant field and users who self-identify as queer, gender-diverse, women from the Global Majority with some or no expertise about the topic but are exposed to these technologies in their everyday lives. When interviewing, it would be useful to understand how AI functions in discriminatory or inclusive ways:

Data. In addition to understanding how data is used in ways that are both visible and invisible to the public eye, and how this could be abused, interviews would cover how viable it would be to regulate or address inside an

epistemological framework such large amounts of data used by the diverse AI.

Technological design. Interviews would focus on the process of design, seeking insight into decision-making and which processes, theories, laws, and policies influence these design decisions.

Interviews with policymakers would allow the research to understand processes, definitions, tensions, and trade-offs that are being employed in current policy documents. Overall, interviews would enable the recommendations to be as specific and realistic as possible, especially for users who could bring a set of expectations that would allow understanding better of the ideal AI when it comes to gender and queer issues.

3.2 Theoretical Analysis

Some principles of feminist theory had been used in the past to analyse AI and to shape ethics surrounding these technologies³⁴. Feminist legal theory, for example, has been employed to analyse technical issues such as privacy, surveillance, and cyberstalking³⁵.

Feminist, decolonial, and the majority of world theories will be employed to analyse gendered and queer aspects regarding AI. Mary Hawkesworth³⁶ outlines how feminist scholarship seeks to reshape the dominant paradigms so that women's and gender-expansive people's needs, interests, and concerns can be understood and considered in this process. Canada, Norway, and Sweden have all adopted gender, queer and feminist-informed approaches to their foreign policies, for example. Aggestam, Rosamund, and Kronsell³⁷ draw upon feminist theory and ethics of care to theorize feminist foreign policy. This use of gender theory could be replicated to shape practices and theories surrounding AI.

In addition, anti-essentialist theories could be harnessed and used for analysis. In Feminist Legal Theory, Levit and Verchick³⁸ outline how during the mid to late-1980s, several legal theorists complained about the essentialist nature of feminist legal theory. In ‘Race and Essentialism in Feminist Legal Theory’³⁹, Angela P. Harris argues that feminist legal theory relies on gender essentialism. This is the notion that a unitary, essential women’s experience can be isolated and described independently of race, socio-economical class, sexual orientation, and other realities of experience. The result of this is:

“[...]Not only that some voices are silenced to privilege others...but that the voices that are silenced turn out to be the same voices silenced by the mainstream legal voice of “we the people” – among them, the voices of black women.” Harris,(1990) .

This research would draw on relevant theories relating to race, gender, ethnicity, disability, sexuality, and so on to analyse existing and emerging frameworks from an intersectional perspective. This will help to ensure that a broad range of standpoints are considered when it comes to shaping AI with a more inclusive queer perspective. For example, this could include the use of critical race feminist theory, which looks at how traditional power relationships are maintained, as well as post-modern gender theory, and scholars who apply queer or transgender theory to technology and law. Such research could also consider how narrative analysis might enhance traditional methodologies when you include perspectives from the Global Majority that usually are excluded and considered just as users and not designers in the process of developing and deployment of these kinds of technologies.

3.3 Proposed analysis

Harnessing this theoretical work alongside the interviews, this kind of way of research would examine relevant theories, narratives, laws, and policies to examine their impact on issues of gender equality, gender diversity, and queerness with a global majority perspective. This would especially be concerned with theory, social needs, legislation, and policy surrounding the key areas identified here: data and privacy and technological design.

Broadly speaking, this would isolate any content or wording which relates to how technology can facilitate inequality of power, discrimination, or social injustice. Within this analysis, it could focus on: 1) how these structures impact gender, racial and ethnic minorities; (2) how these theories, narratives, and legislation use language and terminology to assume essentialist views of gender in intersection with race and ethnicity; (3) the loopholes which could allow for potential inequality of power or discrimination; (4) the subtext or sub-narrative in these pieces of legislation, or ethical frameworks, including their assumptions of what is meant by gender, race, and ethnicity; and finally, how these structures could be altered to endorse social justice and equality better.

We will face some challenges at the time using this approach to research AI from a gender and queer perspective, with also a glimpse of critical race and decolonial theories in mind. Some of them are ensuring that technical and legal definitions of bias, equality, and fairness match up with what is valued more broadly in society, especially the ones in the Majority of the World; developments, deployments, and laws and policies on AI are still at the embryonic stage, which could make the process slightly staggered. However, this could also be an opportunity, especially as many policies and

technologies are not yet ossified. Such research will need to keep abreast of emerging developments and work to create access to and inform policies, and technologies in development; and also it will be important for researchers to consider how they will address the trade-offs in terms of moral and ethical guidelines⁴⁰.

This work hopes to contribute to the ongoing development of theories, policies, and law, but in general create a common language, an epistemological framework, surrounding AI, gender diversity, and queerness within the inclusion of the Global Majority, and therefore will be influential in shaping their content and impact. It has been established that the concepts that use and the way we embedded them into technologies affect our behaviour⁴¹. Structural changes implemented could contribute to shifting behaviour surrounding gender equality and the intersectional nature of it would enable us to consider many different standpoints, working for widespread social justice and redistribution of power.

4. F.A.T.E: A complementary analysis

Another approach addresses the detection of biases more obliquely, with accountability measures designed to identify discrimination in the processing of personal data. Numerous organizations and companies as well as several researchers propose such accountability. Therefore, having the difficulties of foreseeing AI technologies outcomes as well as reverse-engineering algorithmic decisions, no single measure can be completely effective in avoiding perverse effects. Thus, where algorithmic decisions are consequential, it makes sense to combine measures that should be taken to work together.

Advance measures such as **fairness, accountability, transparency, and ethics (F.A.T.E)**, combined with the retrospective checks of audits and human review of decisions (detection of biases), could help identify and address unfair results. A combination of these measures can complement each other and add up to more than the sum of the parts. This also would strengthen existing remedies for actionable discrimination by providing documentary evidence that could be used in litigation, creating new laws and policies, and frameworks, and developing a deeper understanding of the social implication of the different AI technologies and how we could use those results to improve them or not longer use them.

5. Conclusions

We think this proposal at the beginning could be problematic and considered hard to achieve, but it is exactly in that aspect where this trans-disciplinary approach could offer a more holistic way to understand, embody, and code the experiences of queer, trans, marginalized people into AI and other new technologies where data is the main source.

We argue that any AI design, development, deployment, and detection of biases framework that aspires to be fair, accountable, transparent, and ethical must incorporate queer, decolonial, trans, and other theories from the Global Majority into their 4Ds. Not only that, but we additionally explain the importance of justice and enfranchising, shifting power to the disempowered core values of any accountable and responsible AI system. Creating these AI necessities starting by funding, supporting, and empowering grass-roots work and advocacy to discuss if and how gender, sexuality, and other aspects of queer identity should be used in datasets and AI systems and how risks

should be addressed in order to cause less or not harms along these lines, and don't forget about the final users so the field of AI fosters from diversity and inclusion to credibly and effectively develop reliable AI. Based on this, we want to create an epistemological framework with a queer perspective for AI, as it was explained, and start to analyse the benefits it can bring from its social, mathematical, technical, and practical regulatory aspects.

Referências bibliográficas

- ADAM, A. *Artificial knowing: Gender and the thinking machine*. London: Routledge. 1998.
- ADAM, A. (2005). *Gender, Ethics and Information Technology*. London: Palgrave MACM, New York NY, USA. Illan. 2005.
- AGGESTAM K, BERGMAN-ROSAMOND A and KRONSELL A. *Theorising feminist foreign policy*. *International Relations* [online] 32(4), 1–17. 2018. Available at: <https://journals.sagepub.com/doi/pdf/10.1177/0047117818811892>
- ASARO, P. *Robots and Responsibility from a Legal Perspective*. In: *Proceedings of the IEEE International Conference on Robotics and Automation*. Rome: IEEE. 2007.
- BUTLER, Judith. *Gender Trouble: feminism and the subversion of identity*. New York: Routledge. 1990.
- BUTLER, Judith. *Critically queer* (pp. 11-29). Routledge. 2020.
- COLLETT, Clementine and DILLON, Sarah. *AI and Gender: Four Proposals for Future Research*. Cambridge: The Leverhulme Center for the Future of Intelligence. 2019.
- COLLINS, P. H. *Black Sexual Politics: African Americans, Gender, and the New Racism*. New York: Routledge. 2005.
- COWLS, Josh and KING, Thomas and TADDEO, Mariarosaria and FLORIDI, Luciano, *Designing AI for Social Good: Seven Essential Factors*. 2019. Available at SSRN: <https://ssrn.com/abstract=3388669> or <http://dx.doi.org/10.2139/ssrn.3388669>
- DIGNUM, V. *Responsible Artificial Intelligence: Designing AI for Human Values*. ITU Journal: ICT Discoveries, Special Issue No.1, pp. 1-8. 2017.
- ELLIOT, P. and ROEN, K. *Transgenderism and the Question of Embodiment: Promising Queer Politics?* *GLQ: A Journal of Lesbian and Gay Studies*, 4(2), pp. 231-261. 1998.
- EPSTEIN, S. *A queer encounter: Sociology and the study of sexuality*. *Sociological Theory*, 188-202. 1994.
- ERDÉLYI, O. and GOLDSMITH, J. *Regulating Artificial Intelligence: Proposal for a Global Solution*. In: AIES '18 Proceedings of the 2018 AAAI/ACM, NEW YORK NY, USA. Conference on AI, Ethics, and Society. 2018. Available at: http://www.aies-conference.com/wp-content/papers/main/AIES_2018_paper_13.pdf
- FOMICHOV, V. *Semantics-Oriented Natural Language Processing: Mathematical Models and Algorithms* (Vol. 27). Springer Science & Business Media. 2009.
- FOUCAULT, Michael. *Discipline and punish: The birth of the prison*. London: Penguin Books. 1991.
- GROSSBERG, Lawrence. *Cultural Studies and Deleuze-Guattari, Part 1*. *Cultural Studies* (28): 1-28. 2014
- GROSZ, E. *Volatile bodies: Toward a corporeal feminism*. Bloomington: Indiana University Press. 1994.
- HARAWAY, Dorothy. *Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980s*. *Socialist Review*, 80, pp.65-108. 1985.
- HARAWAY, Dorothy. *Modest_Witness@Second_Millennium.FemaleMan[C]_Meets_OncoMou seTM: Feminism and technoscience*. *Journal of the History of Biology*, 30(3) pp. 494-497. 1997.
- HARRIS, A. *Race and Essentialism in Feminist Legal Theory*. *Stanford Law Review*, 42(3), pp. 581-616. 1990.
- HAWKESWORTH, M. *Policy studies within a feminist frame*. *Policy Sciences*, 27, pp.97-118. 1994.

- JAGOSE, A. *Queer Theory*. United States: NYU Press. 1997.
- LEVIT, N. and VERCHICK, R. *Feminist legal theory: A primer*. New York: New York University Press. 2006.
- M. FINDLAY and J. SEAH, *An Ecosystem Approach to Ethical AI and Data Use: Experimental Reflections*, 2020 IEEE / ITU International Conference on Artificial Intelligence for Good (AI4G), 2020, pp. 192-197, doi: 10.1109/AI4G50087.2020.9311069.
- MONASTERIOS, G. *Abya Yala en Internet políticas comunicativas y representaciones de identidad de organizaciones indígenas en el ciberespacio*. Políticas de identidades y diferencias sociales en tiempos de globalización, 303-330. 2003.
- MUSTAFA, A. "White Crisis" and/as "Existential Risk", or *The Entangled Apocalypticism of Artificial Intelligence*. *Zygon: Journal of Religion and Science*, 54(1), pp. 207-224. 2019.
- O'CONNOR, S. *The robot-proof skills that give women an edge in the age of AI*. *Financial Times*. 2019. Available at: <https://www.ft.com/content/06afd24a-2dfb-11e9-ba00-0251022>
- OYĒWŪMÍ, O. *Invention of Women: Making an African Sense of Western Gender Discourses*. Minneapolis; London: University of Minnesota Press. 1997.
- SALES, L. *Algorithms, Artificial Intelligence and the Law*. *Judicial Review*, 25(1), 46-66. 2020.
- SAVAGE, N. *How AI and neuroscience drive each other forwards*. *Nature*, 571(7766), S15+. 2019. Available at: <https://link.gale.com/apps/doc/A594456957/AONE?u=anon~18f351ba&sid=google Scholar&xid=719bedf7>
- SHAHIDUL Alam. *Majority World: Challenging the West's Rhetoric of Democracy*, *Amerasia Journal*, 34:1, 88-98, 2008, DOI: 10.17953/amer.34.1.13176027k4q614v5
- SHIRAEV, E. B., & LEVY, D. A. *Cross-cultural psychology: Critical thinking and contemporary applications*. Routledge. 2020.
- SILVERSTONE, R. *Domesticating domestication. Reflections on the life of Domestication of media and technology*, 229. 2005.
- SMITH, C.J.. *Designing Trustworthy AI: A Human-Machine Teaming Framework to Guide Development*. ArXiv, abs/1910.03515. 2019.
- SPIVAK, G. *Can the Subaltern Speak?*, in Nelson, C. and Grossberg, L. (eds.) *Marxism and the Interpretation of Culture*. Urbana: University of Illinois Press, pp. 271-313. 1988.
- SPIVAK, G. *In Other Worlds: Essays in Cultural Politics*. New York: Routledge. 1988.
- STANLEY, E. and SMITH, N. *Captive genders: Trans embodiment and the prison industrial complex*. Edinburgh: AK Press. 2011.
- STONE, S. *The Empire Strikes Back: A Postranssexual Manifesto*. *Camera Obscura*, 10(2 29), pp. 150-176. 1992.
- STRYKER, S. *Transgender history*. Berkeley, CA: Seal Press. 2008.
- THEODOROU, A., DIGNUM, V. *Towards ethical and socio-legal governance in AI*. *Nat Mach Intell* 2, 10–12. 2020. Available at: <https://doi.org/10.1038/s42256-019-0136-y>
- VILLANI, C. *For a Meaningful Artificial Intelligence: Towards a French and European Strategy*. Mission assigned by the Prime Minister Édouard Philippe. 2018. Available at: https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf
- WAJCMAN, J. **The Feminisation of Work in the Information Age?** In: M. Frank Fox, D. Johnson and S. Rosser, eds., *Women, Gender, and Technology*. Champaign, Ill.: University of Illinois Press, pp. 80-97. 2006.

- WARNER, M. (Ed.). *Fear of a queer planet: Queer politics and social theory* (Vol. 6). U of Minnesota Press. 1993.
- WAYAR, Marlene. *Travesti. Una teoría lo suficientemente buena*. Fotografías de Lina M. Etchesuri. Ilustrada por Nina Kunan. 2ª reimpressão. CABA: Muchas Nueces. 2019.
- WELLER, Adam. *Transparency: Motivations and Challenges*. In: Samek, W., Montavon, G., Vedaldi, A., Hansen, L., Müller, KR. (eds) *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*. Lecture Notes in Computer Science(), vol 11700. Springer, Cham. 2019. Available at: <https://doi.org/10.1007/978-3-030-28954-62>
- WHITTLESTONE, J. et al. *Ethical and societal implications of algorithms, data, and artificial intelligence: a roadmap for research*. London: Nuffield Foundation. 2019
- XI, Zhexu. *How Can Humans Drive the Development of Ethical Artificial Intelligence?*. Doi: 10.1007/978-3-030-73103-8_66. 2021.

Notas finais

1 A teoria do conhecimento, especialmente com relação a seus métodos, validade e alcance, e a distinção entre crença justificada e opinião. A maneira como vemos e nos movemos no mundo.

2 Adam, A. Op. Cit.

3 Ali, Mustafa Op. Cit.

4 Collett, Clementine and Dillon, Sarah Op. Cit.

5 Weller, A. Op. Cit.

6 O'Connor, S. Op. Cit.

7 the theory of knowledge, especially with regard to its methods, validity, and scope, and the distinction between justified belief and opinion. The way we see and move through the world.

8 Adam, A. (1998). *Artificial knowing: Gender and the thinking machine*. London: Routledge.

9 Ali, Mustafa (2019). "White Crisis" and/as "Existential Risk", or The Entangled Apocalypticism of Artificial Intelligence. *Zygon: Journal of Religion and Science*, 54(1), pp.207-224

10 Collett, Clementine and Dillon, Sarah (2019). *AI and Gender: Four Proposals for Future Research*. Cambridge: The Leverhulme Center for the Future of Intelligence.

11 Weller, A. (2019). Transparency: Motivations and Challenges. In: Samek, W., Montavon, G., Vedaldi, A., Hansen, L., Müller, KR. (eds) *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*. Lecture Notes in Computer Science(), vol 11700. Springer, Cham. <https://doi.org/10.1007/978-3-030-28954-62>

12 O'Connor, S. (2019). The robot-proof skills that give women an edge in the age of AI. [online] *Financial Times*. Available at: <https://www.ft.com/content/06afd24a-2dfb-11e9-ba00-0251022>

13 Shahidul Alam (2008) Majority World: Challenging the West's Rhetoric of Democracy, *Amerasia Journal*, 34:1, 88-98, DOI: 10.17953/amer.34.1.l3176027k4q614v5

14 Cows, Josh and King, Thomas and Taddeo, Mariarosaria and Floridi, Luciano, *Designing AI for Social Good: Seven Essential Factors* (May 15, 2019). Available at SSRN: <https://ssrn.com/abstract=3388669> or <http://dx.doi.org/10.2139/ssrn.3388669>

15 Smith, C.J. (2019). *Designing Trustworthy AI: A Human-Machine Teaming Framework to Guide Development*. ArXiv, abs/1910.03515.

16 Xi, Zhexu. (2021). How Can Humans Drive the Development of Ethical Artificial Intelligence?. 10.1007/978-3-030-73103-8_66.

17 Theodorou, A., Dignum, V. Towards ethical and socio-legal governance in AI. *Nat Mach Intell* 2, 10–12 (2020). <https://doi.org/10.1038/s42256-019-0136-y>

18 Dignum, V. (2017). *Responsible artificial intelligence: designing AI for human values*.

19 M. Findlay and J. Seah, "An Ecosystem Approach to Ethical AI and Data Use: Experimental Reflections," 2020 IEEE / ITU International Conference on Artificial Intelligence for Good (AI4G), 2020, pp. 192-197, doi: 10.1109/AI4G50087.2020.9311069.

20 Butler, J. (2020). *Critically queer* (pp. 11-29). Routledge. Epstein, S. (1994). A queer encounter: Sociology and the study of sexuality. *Sociological Theory*, 188-202. Jagose, A. (1997). *Queer Theory*. United States: NYU Press. Shiraev, E. B., & Levy, D. A. (2020). *Cross-cultural psychology: Critical thinking and contemporary applications*. Routledge. Stone, S. (1992) *The Empire Strikes Back: A Posttranssexual Manifesto*. *Camera Obscura*, 10(2 29), pp. 150-176. Stryker, S. (2008). *Transgender history*. Berkeley, CA: Seal Press. Warner, M. (Ed.). (1993). *Fear of a queer planet: Queer politics and social theory* (Vol. 6). U of Minnesota Press.

21 Collins, P. H. (2005). *Black Sexual Politics: African Americans, Gender, and the New Racism*. New York: Routledge. Oyěwùmí, O. (1997) *Invention of Women: Making an African Sense of Western Gender Discourses*. Minneapolis; London: University of Minnesota Press. Spivak, G. (1988) 'Can the Subaltern Speak?', in Nelson, C. and Grossberg, L. (eds.) *Marxism and the Interpretation of Culture*. Urbana: University of Illinois Press, pp. 271-313. Spivak, G. (1988) In *Other Worlds: Essays in Cultural Politics*. New York: Routledge.

22 Monasterios, G. (2003). *Abya Yala en Internet políticas comunicativas y representaciones de identidad de organizaciones indígenas en el ciberespacio. Políticas de identidades y diferencias sociales en tiempos de globalización*, 303-330.

23 Butler, J. (1990). *Gender Trouble: feminism and the subversion of identity*. New York: Routledge. Foucault, M. (1991 [1975]). *Discipline and punish: The birth of the prison*. London: Penguin Books. Grosz, E. (1994). *Volatile bodies: Toward a corporeal feminism*. Bloomington: Indiana University Press.

24 Fomichov, V. (2009). *Semantics-Oriented Natural Language Processing: Mathematical Models and Algorithms* (Vol. 27). Springer Science & Business Media.

25 Savage, N. (2019). How AI and neuroscience drive each other forwards. *Nature*, 571(7766), S15+. <https://link.gale.com/apps/doc/A594456957/AONE?u=anon~18f351ba&sid=googleScholar&xid=719bedf7>

26 Wayar, M. Wayar, Marlene (2019). *Travesti. Una teoría lo suficientemente buena. Fotografías de Lina M. Etchesuri. Ilustrada por Nina Kunan. 2ª reimpresión*. CABA: Muchas Nueces.

27 Haraway, D. (1985). *Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980s*. *Socialist Review*, 80, pp.65-108. Haraway, D. (1997). *Modest_Witness@Second_Millennium.FemaleMan[C]_Meets_OncoMou seTM: Feminism and technoscience*. *Journal of the History of Biology*, 30(3) pp. 494-497. Wajcman, J. (2006) *The Feminisation of Work in the Information Age?* In: M. Frank Fox, D. Johnson and S. Rosser, eds., *Women, Gender, and Technology*. Champaign, Ill.: University of Illinois Press, pp. 80-97.

- 28 Elliot, P. and Roen, K. (1998) Transgenderism and the Question of Embodiment: Promising Queer Politics? *GLQ: A Journal of Lesbian and Gay Studies*, 4(2), pp. 231-261. Stanley, E. and Smith, N. (2011). *Captive genders: Trans embodiment and the prison industrial complex*. Edinburgh: AK Press.
- 29 Grossberg, Lawrence. 2014. Cultural Studies and Deleuze-Guattari, Part 1. *Cultural Studies* (28): 1-28.
- 30 Sales, L. (2020). Algorithms, Artificial Intelligence and the Law. *Judicial Review*, 25(1), 46-66.
- 31 Erdélyi, O. and Goldsmith, J. (2018). Regulating Artificial Intelligence: Proposal for a Global Solution. In: AIES '18 Proceedings of the 2018 AAAI/ACM, NEW YORK NY, USA. Conference on AI, Ethics, and Society. [online] Available at: http://www.aies-conference.com/wp-content/papers/main/AIES_2018_paper_13.pdf
- 32 Whittlestone, J. et al. (2019). Ethical and societal implications of algorithms, data, and artificial intelligence: a roadmap for research. London: Nuffield Foundation.
- 33 Villani, C. (2018). For a Meaningful Artificial Intelligence: Towards a French and European Strategy. Mission assigned by the Prime Minister Édouard Philippe [online] Available at: https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf
- 34 Asaro, P. (2007). Robots and Responsibility from a Legal Perspective. In: Proceedings of the IEEE International Conference on Robotics and Automation. Rome: IEEE.
- 35 Adam, A. (2005). *Gender, Ethics and Information Technology*. London: Palgrave MACM, New York NY, USA. illan
- 36 Hawkesworth, M. (1994). Policy studies within a feminist frame. *Policy Sciences*, 27, pp.97-118
- 37 Aggestam K, Bergman-Rosamond A and Kronsell A (2018) Theorising feminist foreign policy. *International Relations* [online] 32(4), 1–17. Available at: <https://journals.sagepub.com/doi/pdf/10.1177/0047117818811892>
- 38 Levit, N. and Verchick, R. (2006). *Feminist legal theory: A primer*. New York: New York University Press
- 39 Harris, A. (1990). Race and Essentialism in Feminist Legal Theory. *Stanford Law Review*, 42(3), pp. 581-616.
- 40 Dignum, V. (2017). Responsible Artificial Intelligence: Designing AI for Human Values. *ITU Journal: ICT Discoveries*, Special Issue No.1, pp. 1-8
- 41 Silverstone, R. (2005). Domesticating domestication. *Reflections on the life of Domestication of media and technology*, 229.

ARTIGO

A Plataformização das Escolas Brasileiras: impactos à educação e proteção de dados de crianças e adolescentes

Anderson Batistaandersonrohe2@gmail.com

Doutorando em Tecnologias da Inteligência e Design Digital (TIDD) pela PUC-SP. Mestre em Políticas Internacionais pela PUC-Rio. Pesquisador no Grupo de Pesquisa de Direito e Inovação (GEDI) da Universidade Presbiteriana Mackenzie.

Eduardo Arientoeduardo.ariante@mackenzie.br

Pesquisadora do Grupo de Pesquisa de Direito e Inovação (GEDI) da Universidade Presbiteriana Mackenzie. Graduada em Direito na Universidade Presbiteriana Mackenzie. Bolsista pelo MackPesquisa no projeto de iniciação tecnológica (PIBITI) no tema de Direitos de Crianças e Adolescentes diante de Tecnologias Persuasivas e monetização do YouTube.

Natália Ribeironataliagribeiro2@gmail.com

Doutor e Mestre em Direito pela Universidade de São Paulo; Professor da Faculdade de Direito da Universidade Presbiteriana Mackenzie e Coordenador do Núcleo de Inovação Tecnológica (NIT) da mesma instituição; Líder do Grupo de Pesquisa de Direito e Inovação (GEDI); Pesquisador do Núcleo Jurídico do Observatório da Inovação e Competitividade (NJ-OIC), do IEA-USP; Advogado em São Paulo-SP

A Plataformização das Escolas Brasileiras: impactos à educação e proteção de dados de crianças e adolescentes

Palavras-chave

Crianças e Adolescentes

Inteligência Artificial

Plataformização da Educação

Privacidade

Proteção de Dados

Resumo

O artigo versa sobre a plataformização das escolas brasileiras, os riscos e benefícios dela decorrentes para a garantia de uma educação com igualdade e qualidade de acesso, sobretudo no que diz respeito à coleta de dados feita por empresas de tecnologia. O debate é imprescindível uma vez que os alunos são, em sua maioria, crianças e adolescentes que por se encontrarem em peculiar fase de desenvolvimento sociocognitivo são considerados grupos vulneráveis e, por isso, sujeitos à proteção integral e prioridade absoluta garantida por lei. Nesse contexto, objetiva-se identificar os potenciais impactos da plataformização por meio do uso de novas tecnologias de informação e comunicação na educação, inclusive em relação ao aprendizado dos alunos. Como método de pesquisa, realizar-se-á uma análise comparativa de materiais bibliográficos e documentais acerca do uso

crescente de plataformas digitais e Inteligência Artificial na educação, da coleta excessiva de dados pessoais e metadados no ambiente escolar por empresas de tecnologia voltadas à educação, assim como as diretrizes pedagógicas e regulamentações vigentes no Brasil quanto à proteção de dados durante a infância a adolescência, no tocante ao seu melhor interesse. Para tanto, e como resultado a ser alcançado, espera-se efetuar recomendações para o uso e desenvolvimento mais seguro, ético e responsável desses novos aparatos tecnológicos nas escolas, na intenção de potencializar benefícios e mitigar danos.

The Platformization of Brazilian Schools: impacts for education and data protection of children and teenagers

Keywords

Artificial Intelligence

Children and Teenagers

Data Protection

Platformization of Education

Privacy

Abstract

The article addresses the platformization of Brazilian schools, as well as the risks and benefits arising from ensuring education with equal access and quality, especially regarding the data collection conducted by technology companies aimed at providing educational services. The debate is essential since the most part of students are children and adolescents, who, being in a unique phase of socio-cognitive development, are considered vulnerable groups and, therefore, subject to comprehensive protection and absolute priority guaranteed by law. In this context, the objective is to identify the potential positive and negative impacts of platformization using new information and communication technologies in education, including the student learning. As a research method, a comparative analysis of bibliographic and documentary materials will be conducted concerning the increasing use of

digital platforms and Artificial Intelligence in education; the excessive collection of personal data and metadata in the school environment by technology companies applied in education; besides the current pedagogical guidelines and regulations in Brazil regarding data protection during childhood and adolescence, focusing on their best interests. Consequently, the aim is to develop recommendations for the safer, ethical, responsible use and development of these new technological devices in schools, intending to enhance benefits and mitigate harm.

1. Introdução

O advento das chamadas tecnologias emergentes, sobretudo da Inteligência Artificial (IA), causou uma série de mudanças sociais e comportamentais na vida humana. Em função da pandemia de Covid-19, alguns desses novos aparatos tecnológicos tiveram sua implementação acelerada (DPB, 2023). Dentre esses mecanismos, o uso de plataformas digitais voltadas aos serviços educacionais foi essencial para manter a continuidade do ensino, permitindo que este continuasse de forma remota. Porém, mesmo com o retorno ao ensino de forma presencial, não houve descontinuidade do processo de aplicação de novas Tecnologias Digitais de Informação e Comunicação (TDIC) na educação, desde o ensino básico até o ensino superior. Pelo contrário, sistemas movidos por Inteligência Artificial são hoje uma realidade cada vez mais comum em secretarias de educação, escolas públicas e privadas.

Novas metodologias de ensino têm empregado a Inteligência Artificial como promessa de eficiência e inovação em várias frentes, servindo de instrumento de apoio à gestão escolar. Essas funcionalidades podem auxiliar os processos de avaliação, desempenho e aprendizagem dos alunos, por meio da análise comportamental e necessidades personalizadas dos alunos. Além disso, sistemas de IA podem aprimorar a eficiência de agentes pedagógicos através de sistemas instrucionais e de diálogo personalizados; da análise da escrita dos alunos e criação de ambientes mais inteligentes em jogos, *chatbots* de suporte ao aluno (Holmes, Bialik, & Fadel, 2019), entre outros. Por outro lado, ainda não há comprovação científica suficiente em muitas dessas iniciativas, sobre quanto à efetividade dessas tecnologias, como é o caso do monitoramento da atenção dos alunos nas salas de aula por programas de reconhecimento de emoções (Sé, 2020).

Contudo, mesmo diante de um contexto de mudanças, tanto positivas quanto negativas, o cenário atual é de escalada para uma plataforma da educação e dos sistemas de ensino. Para efeito de contextualização, o conceito de plataforma tem mudado ao longo do tempo, mas, sinteticamente, são estruturas ou ambientes virtuais facilitadores de interações e compartilhamento de informações que abrangem desde sites, buscadores e redes sociais a aplicativos e serviços *online* de transporte, negócios e transações bancárias, entre outros. Já “Plataformização” ou “sociedade de plataforma” (Van Dijck, Poell, & De Waal, 2018) são termos derivados da primeira para descrever o processo de introdução de “plataformas digitais em diferentes setores econômicos e esferas da vida humana” (Poell, Nieborg, & Van Dijck, 2019, p. 2). Isto é, de como os fluxos econômicos e de convivência social são hoje modulados e redefinidos por um ecossistema global de plataformas digitais *online* (Poell, Nieborg, & Van Dijck, 2019, p. 5).

Então, o que há de errado com a plataforma? Grande parte dessas infraestruturas são ofertadas por *Big Techs* como Amazon, Google, Meta, Microsoft, e empresas do setor educacional como Knewton e Carnegie Learning, mas nem sempre com o devido dever de cuidado. Assim, a forma como são projetadas é vista de maneira negativa, já que seu “design viciante” tem provocado a dependência do usuário, principalmente o público infante juvenil, um dos mais vulneráveis ao tempo prolongado de exposição às telas, aos jogos *online*, às reproduções automáticas de vídeos e rolagem infinita do *feed* de notícias. Justamente pela falta de autocontrole dos jovens na permanência dentro dessas plataformas (Fernandes, 2024).

Ademais, entre as redes de ensino pesquisadas pelo Comitê Gestor da Internet no Brasil (CGI.br, 2022) não foram encontrados indícios de criação de *softwares* e códigos de plataformas educacionais próprios das escolas brasileiras,

muito menos um controle consistente sobre termos de uso e política de dados. Isso significa que a coleta de dados pessoais dos usuários feita por essas plataformas encontra-se hegemonicamente sob responsabilidade das próprias *Big Techs* (em sua maioria estrangeiras).

Ao entrar na questão da coleta de dados feita pelas *EdTechs* – empresas de tecnologia voltadas ao oferecimento de serviços educacionais –, sobretudo por meio da aplicação de seus sistemas de IA na educação básica e superior, constata-se que estão envolvidos dados principalmente de crianças e adolescentes, de acordo com relatório divulgado pelo CETIC. Particularmente no ambiente digital brasileiro, uma vez que “em 2022, 92% da população com idade entre 9 e 17 anos era usuária de Internet no país (aproximadamente 24,4 milhões de crianças e adolescentes)”¹. Essa informação, por si só, é preocupante, já que este público carece de proteção especial, garantida legalmente pelo dever de cuidado e de garantia do melhor interesse com absoluta prioridade.

Ocorre que, diante da vulnerabilidade dessa faixa etária, coloca-se em risco a privacidade e proteção de dados pelo mero oferecimento desses sistemas educacionais plataformizados, tanto gratuitos quanto pagos. Foi essa a conclusão do *Human Rights Watch* (HRW, 2022), organização internacional não governamental com foco na defesa dos direitos humanos, ao publicar o relatório *How Dare They Peep into My Private Life*² (Como Eles Ousam Espiar Minha Vida Privada, em português). Tal estudo investigou violações dos direitos das crianças por parte de governos que apoiaram a aprendizagem online durante a pandemia de Covid-19, constatando-se que a maior parte das *EdTechs* analisadas forneceu dados pessoais dos alunos a empresas de tecnologia e de publicidade sem o consentimento das crianças ou de seus pais.

Não somente a privacidade e a proteção dos dados pessoais de crianças e adolescentes são postas em risco diante da implementação da

plataformização da educação nas escolas, mas também o seu aprendizado. Assim, a aplicação de forma desmedida de sistemas de Inteligência Artificial na educação pode apresentar limitações relevantes que afetam não apenas os direitos desses indivíduos, como seu próprio desenvolvimento cognitivo e social. O uso excessivo dessas tecnologias pode afetar sobretudo jovens que ainda carecem de maior amadurecimento e senso crítico (Da Hora, 2023), “em razão de seu peculiar estágio de desenvolvimento biopsicossocial” (Instituto Alana, 2024).

As plataformas digitais, inclusive mecanismos de IA utilizados na educação, propiciam algumas vantagens por meio da possibilidade de *feedback* imediato aos alunos e personalização do aprendizado. Além disso, podem oferecer uma vasta gama de recursos educacionais digitais capazes de enriquecer o ensino e facilitar a aprendizagem interativa, preparando os estudantes para um mundo cada vez mais digitalizado. Por outro lado, podem trazer desvantagens diante da ausência do professor, o que poderia impactar negativamente a capacidade de análise crítica, a interação social e o desenvolvimento socioemocional dos alunos (Da Hora, 2023). Logo, poucas pessoas entendem que a Inteligência Artificial será disruptiva, mas que a educação estará na linha de frente dessa revolução³.

Diante deste cenário ambivalente que atinge em cheio a realidade da educação brasileira, tanto no setor público quanto no privado, o presente artigo tem como escopo analisar os impactos da plataformização na educação, sobretudo para os direitos fundamentais de crianças e adolescentes no ambiente escolar. As justificativas do artigo são, as seguintes: a) são grupos vulneráveis, sujeitos à proteção integral e prioridade absoluta; b) a plataformização da educação pode servir de aliada, mas também coloca em risco direitos fundamentais à privacidade e proteção de dados durante a infância e juventude; c) o uso ainda precoce e desmedido

das IAs nas escolas pode provocar danos ao aprendizado e ao desenvolvimento cognitivo e social dos alunos.

Para tanto, apoia-se em análise bibliográfica e documental, por meio de casos exemplificativos da tentativa brasileira de lidar com as matérias que se interrelacionam com a temática das IAs nas escolas. Ao final, efetuam-se recomendações, a título de conclusão, de como mitigar danos relacionados à proteção de dados e privacidade de crianças e adolescentes sem, contudo, inibir a inovação tecnológica para a área da educação.

2. Plataformização da educação e suas implicações para o público em geral

A educação não ficou de fora diante da inserção de uma série de mecanismos digitais na sociedade e vida humana. Primeiramente, durante a pandemia da Covid-19, o fenômeno da plataforma da educação se intensificou pela necessidade de adaptação do ensino presencial para o remoto, o que contribuiu para o aumento da visibilidade desse efeito. Concomitante, houve o surgimento de novas inquietações nas áreas acadêmica, pedagógica, jurídica e das organizações defensoras dos direitos humanos.

Em um segundo momento, mesmo com o retorno às aulas de forma presencial, o uso de mecanismos digitais e o envolvimento de plataformas como as *EdTechs*, as aplicações de Inteligência Artificial na educação não só continuaram presentes, como também se intensificaram. Entre as razões da permanência desse novo recurso na educação para além da situação emergencial pandêmica estão: a necessidade de trazer inovação para a educação e servir de apoio aos professores para transmitir

seus conhecimentos; organizar conteúdos, facilitar o contato entre pais e alunos⁴; e prestar auxílio aos estudantes⁵.

Já quando se pensa em benefícios, a inserção de aparatos tecnológicos envolvendo Inteligência Artificial nas escolas funciona como: a) recurso de inovação, ou seja, a tecnologia permite realizar tarefas de difícil efetivação, como o uso de aplicativos de realidade aumentada que buscam auxiliar o ensino de alunos com dislexia, e de aplicativos de boards colaborativos em tempo real; b) recurso de modificação, visto que a IA permite o “re-design” das tarefas que já eram realizadas antes do advento do digital, a exemplo das plataformas de gravação de filmes no ambiente online para compartilhamento de tarefas; c) recurso de substituição de tarefas que permitem a troca do físico pelo virtual, tais como as notas digitais que substituem cadernos e calculadoras⁶.

Pelo lado positivo, ressalta-se ainda, para os estudantes, a existência de *softwares* que possibilitam a melhora da escrita, da análise de dados e tradução de textos. São opções mais baratas na aprendizagem de idiomas e plataformas 3D que tornam o estudo das disciplinas mais interessante. Já para os professores, destacam-se o auxílio de plataformas para detecção de plágio, controle de notas e frequência; aplicativos para elaboração de aulas mais interativas e simuladores de questões.

Por outro lado, é necessário considerar os riscos associados ao uso dessas tecnologias. A aplicação dessas estruturas tecnológicas de forma desmedida, desregulada e não supervisionada pode trazer impactos negativos tanto para professores, quanto para os pais e alunos. Portanto, se o propósito de introduzir tecnologia nas escolas for o de realmente inovar, a sua aplicação não deveria se distanciar do objetivo principal das escolas em promover o ensino, a educação e desenvolvimento da pessoa, o preparo para o exercício da cidadania e qualificação para o trabalho⁷.

O desafio começa com a capacitação e formação dos professores para possibilitar o uso desses novos recursos digitais de forma conjunta e articulada. Porém, na maior parte das vezes, essa situação ideal nem sempre ocorre. Começa a partir daí um *problema* com o possível aumento na dependência do setor educacional em relação ao rol de serviços ofertados pelas grandes plataformas digitais. Ao menos em tese, essa assimetria técnica e informacional pode colocar em risco a autonomia das escolas públicas e privadas de ensino. Para além do ponto de vista pedagógico, essa provável concentração monopolística de poder pelas *EdTechs* também acarreta discussões jurídicas acerca da autonomia das redes de ensino, já que parece ir contra ao pluralismo de ideias da Lei de Diretrizes e Bases da Educação Nacional⁸.

A título de exemplo, em agosto de 2023, houve uma tentativa do governo do estado de São Paulo de fornecer apenas material didático digital aos alunos do Ensino Fundamental II e do Ensino Médio⁹. Acadêmicos e especialistas em educação teceram críticas contundentes à decisão do governador, mormente pelo material ter sido desenvolvido não por educadores, e sim pelos próprios técnicos estaduais. Dada a má repercussão, posteriormente o governador recuou e decidiu oferecer livros também impressos aos alunos¹⁰, especialmente pelo fato de que não é possível impor esse sistema tão rapidamente.

Não se trata, portanto, de atribuir somente aspectos negativos às tecnologias digitais na educação, porém é necessário haver um processo de adequação paulatina à nova realidade¹¹, principalmente por se tratar do desigual cenário socioeconômico brasileiro. Assim, a inclusão desses novos mecanismos no ensino, havendo consenso entre educadores e gestores públicos, precisaria ocorrer de forma gradual, por meio de adaptações na infraestrutura das escolas, na capacitação dos docentes e nas próprias propostas pedagógicas, de forma analítica, crítica, democrática e participativa.

3. A plataformização da educação em várias frentes

3.1 na infraestrutura das escolas

No Brasil, a plataformização da educação encontra dois desafios relevantes: desigualdade de acesso à educação e violação dos princípios orientadores de justiça e equidade de acesso à tecnologia, inclusive à Inteligência Artificial^{12 13}. De acordo com a análise quantitativa dos dados publicados na pesquisa TIC - Educação 2022¹⁴, realizada pelo Comitê Gestor da Internet no Brasil (CGI.br, 2022), apenas 58% das escolas cadastradas no *Censo Escolar da Educação Básica 2021 de Ensino Fundamental e Médio* possuem computador e acesso à *internet* para uso dos alunos em atividades de ensino e de aprendizagem. Isso demonstra que ainda há um longo caminho a ser percorrido para a efetiva inserção de mecanismos tecnológicos para a educação nas escolas brasileiras.

Ademais, constataram-se outras desigualdades marcantes. Enquanto 74% das escolas localizadas na região urbana possuem acesso a essas ferramentas, apenas 30% das escolas localizadas na região rural têm essa disponibilidade. No que diz respeito à dependência administrativa, 73% das escolas particulares têm o acesso a computadores e à *internet*, ao passo que 82% das escolas estaduais possuem esses instrumentos e apenas 43% das escolas municipais garantem esse acesso. Com relação à localização também há uma desproporção, já que 73% das escolas da capital dispõem de computadores e acesso à *internet*, enquanto 55% das escolas no interior dispõem de tais recursos.¹⁵

A partir dos dados coletados naquela pesquisa, é perceptível que a digitalização de atividades educativas e pedagógicas no atual cenário brasileiro enfrenta dificuldades em garantir igualdade à educação.

3.2 na capacitação dos docentes e digitalização do ensino

O relatório *TIC - Educação 2022* (CGI.br, 2022) indicou que apenas 56% dos professores reportaram terem participado de algum tipo de formação continuada sobre o uso de tecnologias digitais em atividades de ensino e de aprendizagem nos doze meses anteriores à pesquisa. Além disso, 75% dos professores entrevistados afirmaram que a falta de um curso específico dificultava a adoção de tecnologias digitais nas atividades educacionais com os alunos. Logo, a carência de formação de educadores para o uso de meios digitais nas escolas brasileiras é uma questão que necessita de maior investimento por parte do Estado e de uma articulação coordenada para incluir prefeitura, secretaria da educação e União, porém sem excluir a iniciativa privada das escolas particulares. Isso antes mesmo de se pensar em implementar mecanismos digitais na educação.

O panorama de uma “educação digital” como direito fundamental, por sua vez, ampara-se na garantia do art. 26 do Marco Civil da Internet¹⁶, ao estabelecer que o cumprimento do dever constitucional da educação inclui a capacitação integrada a outras práticas educacionais para o uso seguro, consciente e responsável da internet. Essa capacitação inclui educação para o uso responsável da *web* pelos alunos, o que depende da capacitação plena dos docentes nesse aspecto.

3.3 nas propostas pedagógicas

Os propósitos dos sistemas de ensino originalmente giravam em torno da criação de trabalhadores, da transmissão de conhecimento religioso e das habilidades básicas de alfabetização e numeracia (isto é, a capacidade de usar números e habilidades matemáticas para contar e resolver problemas da vida cotidiana). Porém, à medida que a estrutura da sociedade mudou, a educação começou a desempenhar outras funções práticas, sociais e emocionais. Agora, a educação é tanto uma forma de moldar os alunos para atender às necessidades da sociedade, quanto um meio pelo qual os alunos podem se capacitar para melhor atender às suas próprias necessidades¹⁷.

No artigo 205 da Constituição Federal de 1988, a educação tem por objetivos o pleno desenvolvimento da pessoa, seu preparo para o exercício da cidadania e sua qualificação para o trabalho. Todavia, na atual perspectiva de plataforma da educação e inserção de mecanismos de Inteligência Artificial nas escolas, há risco de que o excesso de telas e de imersão nos meios digitais esteja causando danos futuros ao desenvolvimento cognitivo e à sociabilidade dos alunos. No estado do Paraná, por exemplo, foi empregado um controverso sistema de biometria facial e IA para fiscalizar o comportamento dos alunos¹⁸. Esses danos são significativos pois, quando pensados no ambiente da educação infantil até o ensino médio, a problemática se agudiza, uma vez que é um público singular, composto em sua maioria por crianças e adolescentes, dada a fase peculiar de desenvolvimento sociocognitivo em que se encontram.

Na primeira infância – período que vai desde a concepção do bebê até os 6 anos de idade¹⁹ – são formadas 90% das conexões cerebrais²⁰. Contudo, apesar dessa fase da vida envolver

um enorme potencial para o desenvolvimento, também os caracteriza como seres hiper vulneráveis que necessitam de proteção integral. Em termos jurídicos, a doutrina da proteção integral foi inserida no Brasil na Constituição Federal de 1988²¹, tendo como pressuposto a compreensão de que crianças e adolescentes vivenciam peculiar fase de desenvolvimento que justifica a proteção em caráter especial e diferenciado em relação aos outros entes da sociedade.

Os potenciais danos ao desenvolvimento cognitivo e ao aprendizado dos alunos na fase de inserção dos meios digitais no ambiente escolar se iniciam com os desafios causados à habilidade de leitura, o que é fundamental para uma educação de qualidade e formação digna do cidadão. O fato, por exemplo, de se realizar a leitura cada vez mais por meio de telas, substituindo o papel físico, gerou como consequência prática o ato de apenas "passar os olhos" de forma superficial em vários dos textos e conteúdos *online*. Essa mudança comportamental, explica a neurocientista Maryanne Wolf (2019), pode estar arruinando a capacidade humana de entender argumentos complexos, de fazer uma análise crítica do que se lê e até mesmo de dificuldade em gerar empatia por pontos de vista diferentes aos seus.

Como explica Wolf (2019), o circuito da leitura é moldado pela forma como se lê e o tempo gasto pelo leitor. Assim, pelo fato de a exposição exacerbada aos mecanismos digitais influenciar uma leitura menos aprofundada e mais superficial, há o risco de perda da habilidade de se entender argumentos complexos. Neste cenário, o circuito da leitura no cérebro não vai alocar tempo suficiente para o processamento cognitivo que, por sua vez, é necessário para processamento crítico.

Tais evidências foram percebidas de forma mais clara por meio da meta-análise feita por Pablo Delgado, Cristina Vargas e Ladislao Salmerón, da Universidade de Valência, assim como Rakefet Ackerman, do Instituto de

Tecnologia de Tel Aviv. Tais pesquisadores analisaram estudos publicados entre 2000 e 2017, a partir de dados de 171 mil pessoas, para comparar a compreensão de leitura dos participantes no meio digital e papel. Eles concluíram que o melhor entendimento do conteúdo é feito por meio de textos em papel, principalmente quando o leitor possui um tempo pré-estabelecido para concluir a leitura. O estudo identificou, ainda, a "inferioridade da tela", tendo em vista que a leitura digital aparentemente não é positiva para a habilidade de compreensão dos leitores, uma vez que o processamento de informações é mais superficial nos meios virtuais (Delgado, Vargas, Ackerman, & Salmerón, 2018).

Ao voltar-se para o público mais jovem, a preocupação é intensificada, pois pessoas nessa idade ainda estão desenvolvendo a habilidade da leitura. O uso excessivo de telas acarreta o aumento da impaciência cognitiva, o que não favorece a leitura crítica. Tampouco a compreensão de qualquer fenômeno social complexo que exija uma leitura mais cuidadosa do leitor (Wolf, 2019). Tendo em vista que a inserção de mecanismos digitais e de plataformação da educação não são um cenário futuro, e sim presente, não se pode buscar como solução uma oposição maniqueísta entre os mecanismos digitais e os físicos. A adoção de um em detrimento do outro seria, então, inviável diante da atual transformação digital.

Como sugere Wolf (2019), os alunos e leitores em formação devem entender o que está acontecendo com suas próprias capacidades cognitivas, bem como o propósito do que se está lendo. Ao depararem com um texto complexo, que necessite de um olhar sob diferentes perspectivas e, principalmente, em que necessitem discernir o verdadeiro valor da informação, é relevante refletir se aquele meio vai promover o processamento mais lento e profundo de uma análise crítica. Ademais, crianças e adolescentes em fase de alfabetização precisam de maior concentração naquilo que

estão aprendendo a ler, por se tratar de uma fase crucial na formação do pensamento. Logo, é necessário ponderar em quais momentos e atividades pedagógicas o meio digital pode ser inserido, principalmente perante a necessidade da educação digital, garantida pelo art. 26 do Marco Civil da Internet^{22 23}.

3.4 Ponderação entre riscos e benefícios da plataformização da educação

A exposição direta e indiscriminada dos alunos aos mecanismos tecnológicos digitais pode gerar mais danos do que benefícios às suas capacidades de desenvolvimento cognitivas, principalmente diante de uma variedade excessiva e simultânea de informações. A cognição envolve mecanismos individuais e distribuídos, inerentes às atividades em colaboração que envolvem percepção, atenção, memória, linguagem e raciocínio, atividades cujas origens são culturais²⁴. Logo, é tanto um ato quanto um processo de adquirir conhecimento, através dos quais o ser humano interage com seus semelhantes e com o meio em que vive sem perder sua identidade²⁵.

A possível alteração na cognição das crianças e adolescentes em face da exposição tecnológica exacerbada ocorre, então, porque a potencialidade tecnológica se relaciona com a capacidade de excitar o indivíduo e de estimular seus sistemas auditivo, visual e emocional²⁶. Desse modo, é a preocupação com a cognição que fica mais em evidência diante da proposta de plataformização da educação. A utilização da tecnologia de forma indiscriminada potencializa os transtornos de atenção, os transtornos obsessivos, de ansiedade e de problemas com a linguagem e a comunicação, fenômenos que afetam diretamente a aprendizagem²⁷. Assim,

os usuários regulares de internet têm aumentado a atividade nas regiões pré-frontais do cérebro envolvido na tomada de decisões e resolução de problemas (...) A atividade cerebral é mantida a um nível tão superficial que impede a retenção de informação. Ao manter constantemente ativas as funções executivas do córtex cerebral a sobrecarga cognitiva aparece: a informação passa na frente dos nossos olhos, mas não é mantida (Cánovas, 2015).

A sociabilidade, que é uma das funções da educação e habilidade essencial a ser desenvolvida pelo aluno, pode então ser prejudicada diante de um possível cenário de uso exclusivo de ferramentas digitais e de Inteligência Artificial na educação, uma vez que a redução do tempo de contato físico com os outros estudantes ou mesmo com os professores pode privar os alunos da oportunidade de aprimorar suas habilidades sociais fundamentais, tais como trabalho em equipe, comunicação eficaz e resolução de conflitos²⁸.

No âmbito nacional, a Constituição Federal de 1988 traz, no artigo 227²⁹, a garantia de que crianças e adolescentes possuam a proteção do Estado, da família e sociedade com absoluta prioridade, incluindo direitos como educação. Destaca-se, ainda, que:

o apoio e a proteção à infância e juventude devem figurar, obrigatoriamente, entre as prioridades dos governantes. Essa exigência constitucional demonstra o reconhecimento da necessidade de cuidar de modo especial das pessoas que, por sua fragilidade natural ou por estarem numa fase em que se completa sua formação, correm maiores riscos. A par disso, é importante assinalar que não ficou por conta de cada governante decidir se dará ou não apoio prioritário às crianças e

adolescentes. Reconhecendo-se que eles são extremamente importantes para o futuro de qualquer povo, estabeleceu-se como obrigação legal de todos os governantes dispensar-lhes cuidados especiais. Essa exigência também se aplica à família, à comunidade e à sociedade. Cada uma dessas entidades, no âmbito de suas respectivas atribuições (Dallari, 2022, p. 553).

Nesse contexto, tal garantia deve ser assegurada com absoluta prioridade nas escolas públicas e privadas, compostas em sua maioria por crianças e adolescentes. Logo, diante dos riscos advindos da implementação desorientada da plataformização da educação e do uso de mecanismos de ferramentas digitais no ensino, é necessária a responsabilização não só do Estado, mas da sociedade civil como um todo, o que inclui as empresas envolvidas que fornecem seus serviços para o ensino nas escolas.

4. Proteção de dados de crianças e adolescentes na educação

Além das implicações mencionadas, o emprego da plataformização e de mecanismos de Inteligência Artificial na educação - advindos sobretudo do uso crescente do chat GPT para pesquisa e geração de imagens e textos - de algum modo acarretam a coleta massiva de dados educacionais de crianças e adolescentes. No auxílio ao aprendizado, escolas estaduais, municipais ou particulares utilizam-se, portanto, não só dos serviços ofertados por grandes empresas de tecnologia (*Big Techs*), como também daquelas voltadas diretamente à educação (*EdTechs*). Assim, empresas como Google, Meta e Microsoft buscam integrar suas ações,

desenvolvendo serviços para a área educacional, seja por meio de extensões e conexões vinculados às suas plataformas, seja pelo serviço de armazenamento em nuvem, bibliotecas e sistemas comunicacionais. É assim que, por meio dessas plataformas, passam a dispor de expressiva quantidade não anonimizada de dados pessoais coletados dos estudantes.

Diante de uma economia movida cada vez mais por dados, muitas dessas empresas ofertam seus serviços de maneira “gratuita às escolas”. Porém, o “pagamento” vem por meio da coleta de dados pessoais e metadados relacionados aos conteúdos gerados, às informações compartilhadas e armazenadas nessas plataformas (Gonsales & Amiel, 2020). Na realidade, o monitoramento e a análise de dados dos usuários ensejam uma capitalização digital, nomeada por Shoshana Zuboff (2021) de “Capitalismo de Vigilância”. Assim, algumas *EdTechs* compartilham dados pessoais com empresas de tecnologia para publicidade (*AdTechs*) que, por sua vez, analisam e processam esses dados por meio de algoritmos para melhor compreenderem as preferências dos consumidores e personalizarem suas tarefas.

Não obstante, o tratamento de dados pessoais de crianças e adolescentes deve levar em conta o princípio do melhor interesse desses titulares, reconhecendo a necessidade da tutela especial. Este princípio tem como base expressa o artigo 227 da Constituição Federal de 1988³⁰, garantindo que esses indivíduos possuam a proteção do Estado, família e sociedade, com absoluta prioridade na garantia de seus direitos, e em observância ao disposto no Estatuto da Criança e do Adolescente³¹.

Ademais, a Lei Geral de Proteção de Dados (LGPD)³² destaca, em seção específica, sobre o tratamento de dados pessoais desse público em particular, no *caput* do artigo 14³³. Tal ação deve ser realizada em seu melhor interesse e vir acompanhada do consentimento específico e em destaque de pelo menos um dos pais ou

do responsável legal, salvo se a coleta for necessária para contatá-los ou mesmo para a proteção do menor de idade.

O tratamento de dados pessoais de crianças e adolescentes, por sua vez, está sujeito à interpretação jurídica e doutrinária em relação à base legal aplicável diante do texto normativo. O Enunciado nº 1 publicado pela Autoridade Nacional de Proteção de Dados (ANPD)³⁴ resalta que o tratamento de dados de crianças e adolescentes pode ser realizado com amparo nos artigos 7º e 11 da LGPD³⁵, já que incluem um rol mais abrangente para as hipóteses que autorizam o tratamento de dados, desde que observado o seu melhor interesse, a ser avaliado no caso concreto.

A partir desse entendimento, o tratamento de dados de crianças e adolescentes pode se realizar sem maiores restrições da base legal utilizada, desde que prevalecente o princípio do melhor interesse. A eventual responsabilização por um dano ocorrido, portanto, não recairia apenas aos pais e responsáveis ao fornecer o consentimento, mas também às empresas. Na prática, entretanto, as *EdTechs* e *Big Techs* contratadas por governos e escolas para o oferecimento de serviços educacionais podem estar sobrepondo seus interesses comerciais ao invés de priorizar o melhor interesse de crianças e adolescentes, sobretudo ao utilizar os dados pessoais coletados para além da finalidade educacional.

O relatório *How Dare They Peep into my Private Life? (Como Eles Ousam Espiar Minha Vida Privada)*, da Human Rights Watch (HRW, 2022) investigou *EdTechs* apoiadas por governos para a educação infantil durante a pandemia, com base nas análises técnica e política divulgadas de 163 empresas *EdTechs*, em países incluindo o Brasil. Assim, foi possível concluir que o apoio dos governos à maioria destas plataformas de aprendizagem online coloca em risco ou viola diretamente a privacidade de crianças e adolescentes, entre outros direitos para fins não

relacionados com a educação. Além disso, observou-se que 145 empresas *EdTech* estavam enviando ou concedendo diretamente acesso a dados pessoais de crianças a 196 empresas terceirizadas, esmagadoramente *AdTechs*.

Em síntese, algumas empresas de tecnologia voltadas à educação encaminham os dados dos seus usuários para serem utilizados no direcionamento publicitário, a partir da personalização de anúncios baseado nos seus hábitos e comportamentos. Estes dados pessoais são extraídos de ambientes educativos digitais sem consentimento dos pais ou dos alunos e/ou mesmo sem a ciência dos titulares, responsáveis e professores de que estão sendo tratados para além das finalidades educacionais (HRW, 2022).

Como explica Gonsales e Amiel (2020), poucos gestores, mães, pais e responsáveis estão cientes de que estudantes e educadores têm seus dados coletados e utilizados indevidamente. Tais empresas utilizam principalmente tecnologias de Inteligência Artificial para processar, analisar e identificar maneiras de monetizar os dados pessoais de usuários de forma automatizada, já que estes são hoje um dos principais ativos da economia digital³⁶.

Outro impasse observado na análise do *Human Rights Watch* (2022) foi que alguns dos governos que forneceram educação *online* a crianças por meio de *EdTechs* tornaram obrigatória a utilização de seus serviços por estudantes e professores, o que impossibilitou que estes optassem por alternativas ao acesso à educação. O relatório *Problems with Data Governance in UK Schools: The Cases of Google Classroom and ClassDojo* (Problemas com a Governança de Dados nas Escolas do Reino Unido: os Casos da Google Classroom e ClassDojo), da 5Rights Foundation (5RF, 2022), constatou que diversas escolas britânicas que utilizavam a Google Classroom no ensino enfrentavam problemas significativos com a coleta excessiva de dados. A plataforma coletava mais informações do que o necessário para a

funcionalidade educacional, incluindo dados de navegação e comportamento *online* dos alunos, utilizados para fins de rastreamento e, em alguns casos, para publicidade personalizada. Em resposta, as autoridades regulatórias do Reino Unido realizaram investigações e tomaram ações contra a Google, pressionando a empresa a melhorar suas práticas de privacidade e segurança de dados. A conformidade com o GDPR foi uma preocupação central nessas ações³⁷.

Já na Indonésia, seu Ministério da Educação e Cultura recomendou o aplicativo “Ruangguru”³⁸, que propõe soluções para todas as necessidades de aprendizagem, desenvolvido pela *startup* de mesmo nome, que possui parceria com 326 governos municipais e distritais daquele país. A empresa possuía, em 2020, 22 milhões de alunos como usuários. Observou-se, no entanto, que o aplicativo coletava dados pessoais dos estudantes e os compartilhava com o Facebook, por meio do domínio *graph.facebook.com*. Por sua vez, o Facebook (atual Meta) utilizava esses dados de crianças e adolescentes para o direcionamento de publicidade comportamental.

No Brasil, em agosto de 2023, houve restrição da autodeterminação informativa³⁹ no que se refere ao direito individual de controlar e proteger dados pessoais de pais, professores e alunos no âmbito da rede estadual de ensino do estado de São Paulo. Os usuários desse serviço depararam com um aplicativo instalado em seus aparelhos de celular pessoais, sem o devido consentimento. O aplicativo “*Minha Escola SP*” tinha como finalidade permitir o acompanhamento de informações escolares por pais e estudantes, como as notas do boletim e as faltas às aulas. Posteriormente, a Secretaria da Educação do Estado de São Paulo (Seduc-SP) informou que ocorreu uma falha durante um teste promovido pela área técnica. Todavia, a Secretaria não revelou que o aplicativo foi instalado em aparelhos de sistema operacional Android do qual a Google é proprietária, e com

quem a pasta possui convênio para uso de plataformas educacionais⁴⁰.

Pela análise dos casos mencionados, o tratamento inadequado de dados pessoais para fins educacionais viola os princípios da transparência e da adequação, conforme rege o artigo 6º da Lei Geral de Proteção de Dados (LGPD)⁴¹. Isso ocorre porque os titulares, pais, responsáveis e professores não foram devidamente informados sobre o uso real que é feito com os seus dados, uma vez que essas informações são frequentemente utilizadas além das finalidades previamente comunicadas, mas que deveriam ser restritas à prestação de serviços educacionais. Além disso, por envolverem dados pessoais de estudantes menores de idade, essa prática vai de encontro ao princípio do melhor interesse, previsto no artigo 14 da LGPD⁴², pois as empresas de tecnologia estão, em nome do lucro, sobrepondo seus interesses comerciais.

5. Recomendações para uso de ferramentas digitais na educação

Os casos elencados não apenas evidenciam violação à privacidade e proteção de dados de alunos e professores, como também a excessiva e desnecessária exposição às telas, diante da crescente imersão nos meios digitais para fins educacionais. Diante desse cenário, o uso de Inteligência Artificial nas escolas e a contratação de serviços de *EdTechs* precisam se adequar ao tratamento ético e responsável de dados pessoais, principalmente quando seus titulares são crianças e adolescentes.

Para atender ao princípio do melhor interesse de crianças e adolescentes, as empresas deveriam atentar para que o tratamento dos dados seja feito apenas ao estritamente necessário. Significa dizer que esses dados poderão

ser utilizados para a finalidade específica de promover os serviços e produtos relacionados com aprendizado e educação, mas que não serão utilizados em um contexto de “*behavioral surplus*”⁴³; ou seja, aquelas informações coletadas dos seus titulares que não possuam uso imediato. Percebe-se, assim, que há uma grande chance de os dados serem comercializados ilegalmente, uma vez que:

grande parte dessas tecnologias funciona a partir da coleta massiva de dados dos estudantes para traçar seus hábitos e interesses e, então, direcionar publicidade personalizada a eles. Em outras palavras, um número significativo de tecnologias voltadas ao ensino-aprendizagem se desvia dessa função e se torna um meio para a exploração comercial de crianças e adolescentes (Data Privacy Brasil, 2023, p.12)

Trata-se de dados coletados além do que o necessário para a sua finalidade original, cuja prática foi expressamente vedada pelo art. 14, § 4º, da LGPD⁴⁴, visto que o tratamento adequado de dados pessoais está condicionado ao princípio da necessidade. É preciso, então, que as empresas busquem um ponto de equilíbrio entre atender seus interesses comerciais, priorizar esse público hiper vulnerável, bem como compreender as necessidades e demandas do contexto local (Instituto Alana, 2024).

Essa adequação está ligada à base principiológica para o tratamento de dados pessoais, prevista no artigo 6º, I e III da LGPD, que apenas admite – inclusive para adultos – a coleta dos dados absolutamente indispensáveis para os propósitos legítimos a que devem ser destinados. No contexto deste estudo, o tratamento dos dados pessoais deve, portanto, observar os limites das finalidades e necessidades educacionais e pedagógicas, visando ao objetivo final do aprendizado e a formação social dos estudantes.

Ademais, o princípio da transparência tem sido alvo de violação por parte das EdTechs, em face do tratamento de dados pessoais para além das finalidades esperadas para uma plataforma que tem por objetivo auxiliar na educação. A transparência é princípio trazido pela LGPD (art. 6º, inciso VI) como forma de garantir que os titulares terão informações claras, precisas e facilmente acessíveis sobre a realização do tratamento de seus dados e os respectivos agentes de tratamento.

Para além do tratamento de dados pessoais, a transparência e a explicabilidade também foram um dos princípios estabelecidos pela OCDE (2019) para a utilização de Inteligência Artificial de forma inovadora e confiável. Assim, determina que os desenvolvedores da IA devem fornecer informações significativas sobre a compreensão geral dos sistemas e, inclusive, permitir que as pessoas afetadas por um sistema de IA compreendam os resultados produzidos.

Apesar desses princípios não serem vinculantes, e sim recomendações, o Projeto de Lei que busca regulamentar o uso da Inteligência Artificial no Brasil também traz a transparência, explicabilidade e auditabilidade como princípios norteadores para os sistemas de IA (*ex vi* art. 3º, inciso VI, PL 2.338/ 2023). Por isso devem ser instituídos desde a concepção de produtos e serviços aplicados na educação, principalmente por estarem lidando com dados pessoais de crianças e adolescentes.

Quanto à atribuição de responsabilidade, diante do tratamento indevido de dados dos usuários, as *Edtechs* podem responder solidariamente com os governos. Tal solidariedade ocorre em função da parceria⁴⁵ entre governo e *EdTechs*, ou da relação entre controlador e operador⁴⁶ com as empresas, a depender de análise do caso em concreto. Em qualquer cenário, nenhuma das disposições de compartilhamento de dados exime a responsabilidade dos agentes perante o uso indevido ou abusivo de dados pessoais dos titulares. Torna-se

necessário, antes mesmo de realizar a contratação de EdTechs, que se analise a adequação da Política de Privacidade da empresa, assim como a sua transparência no tratamento dos dados pessoais. Visto que:

o potencial do uso de tecnologias no ensino-aprendizagem, mas sem pregar o discurso tecnosolucionista da inevitabilidade do seu uso como única forma de progresso, é crucial também reconhecer seus riscos. Portanto, acreditamos ser fundamental pensar o processo de introdução de recursos tecnológicos na sala de aula de forma cuidadosa, crítica e, sobretudo, participativa. [...] (Data Privacy Brasil, 2023, p.12).

Por fim, a título de precaução, no intuito de mitigar o tratamento inadequado de dados pessoais dos alunos, tanto governantes quanto as escolas responsáveis pela contratação de EdTechs deveriam: a) realizar o chamado Relatório de Impacto à Proteção de Dados Pessoais (RIPD) - documentação que contém a descrição do tratamento de dados pessoais que podem gerar alto risco à garantia dos princípios gerais de proteção de dados, às liberdades civis e aos direitos fundamentais do titular de dados. Sobretudo antes da inserção dos serviços e produtos dessas empresas nas escolas, a fim de analisar e identificar possíveis fatores de risco envolvidos no tratamento; b) bem como a realização de auditorias periódicas para verificar a adequação dessas empresas, tanto perante o tratamento adequado de dados pessoais, quanto direcionadas aos riscos de aplicação de Inteligência Artificial.

6. Conclusão

Ao longo do trabalho foi demonstrada uma série de riscos em potencial que crianças e adolescentes estão sendo hoje submetidos, em razão da opacidade quanto ao uso e destino de informações pessoais, da coleta indiscriminada e do abuso no tratamento de dados pessoais, assim como da assimetria de poder existente nas relações estabelecidas no ambiente digital. Por outro lado, pouco adianta tentar proibir ou evitar o uso de mecanismos tecnológicos no cenário sociotécnico atual. A digitalização de hábitos e costumes já está inserida em praticamente todos os setores da vida, sobretudo em uma área significativa para a formação do cidadão: a educação.

A plataformização da educação e o uso crescente de ferramentas digitais como métodos de apoio ao ensino, monitoramento e avaliação de desempenho por reconhecimento de emoções não são um cenário futuro, e sim presente. Portanto, o atual impasse entre inovação e regulação não se resolve com a mera proibição desses novos aparatos tecnológicos, ou mesmo retrocedendo para um tempo em que não havia o virtual nas salas de aula.

A proposta do artigo foi a de buscar uma nova composição entre necessidades coletivas dos alunos, gestores da educação, profissionais do ensino e interesses de empresas privadas que oferecem serviços nesse setor. As ferramentas digitais podem ser muito úteis para complementar experiências e imersões, contudo sem prejudicar a sociabilidade, a concentração e a privacidade de pessoas em fase de desenvolvimento.

Logo, são exigidas medidas éticas e mais responsáveis das EdTechs e plataformas que tratam dados pessoais. Deve-se então observar os princípios do melhor interesse de crianças e adolescentes, a finalidade estritamente necessária para a qual os dados educacionais foram

coletados, e a transparência perante o que é feito com os dados pessoais. Assim, diante de uma eventual responsabilidade compartilhada entre famílias, Estado, empresas e a sociedade civil como um todo, algumas medidas precaucionais podem ser adotadas, como é o caso do Relatório de Impacto à Proteção de Dados Pessoais – RIPDD, para efeito de redução dos danos no cenário de uma educação digitalizada.

Por fim, no tocante às implicações relacionadas ao ensino e às práticas pedagógicas no cenário da inserção de mecanismos digitais, incluindo a IA na educação, é imprescindível um ambiente mais democrático de diálogo, com participação mais efetiva de alunos e professores. Desse modo, a interação entre inovações tecnológicas e o objetivo precípua das escolas em promover o ensino e o desenvolvimento pessoal para o exercício da cidadania não podem mais andar em direções opostas, no intuito de assegurar a garantia e o respeito aos direitos de crianças e adolescentes como agentes finais das redes públicas e privadas de educação.

Referências bibliográficas

- Agência Nacional de Proteção de Dados – ANPD (2023). *Enunciado CD/ANPD nº 1, de 22 de maio de 2023*. Diário Oficial da União. Recuperado de <https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-divulga-enunciado-sobre-o-tratamento-de-dados-pessoais-de-criancas-e-adolescentes/Enunciado1ANPD.pdf/view>
- Assembleia Legislativa do Estado de São Paulo - ALESP (2023, agosto). *Seduc instala aplicativo sem consentimento em celulares pessoais de professores e alunos*. Recuperado de <https://www.al.sp.gov.br/noticia/?10/08/2023/seduc-instala-aplicativo-sem-consentimento-em-celulares-pessoais-de-professores-e-alunos>
- Audi, A. (2023). *Reconhecimento facial no Paraná impõe monitoramento de emoções em escolas*. Apública. Recuperado de <https://apublica.org/2023/10/reconhecimento-facial-no-parana-impoe-monitoramento-de-emocoes-em-escolas/>
- Barbosa, R., & Blikstein, P. (2019). *Robótica educacional: experiências inovadoras na educação brasileira*. Penso Editora.
- Cánovas, G. (2015). *Cariño, he conectado a los niños*. Bilbao: Ediciones Mensajero.
- Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação – Cetic.br (2022). *Resumo Executivo TIC Kids Online Brasil 2022*. Recuperado de <https://cetic.br/pesquisa/kids-online/>
- Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação - Cetic.br (2022). *Pesquisa sobre o acesso, o uso e a apropriação de tecnologias digitais em escolas públicas e particulares brasileiras que ofereçam Ensino Fundamental e Médio: TIC Educação 2022 – Coletiva de Imprensa*. Recuperado de https://cetic.br/media/analises/tic_educacao_2022_coletiva%20de%20imprensa.pdf
- CNN Brasil (2023, agosto). *Governo de SP recua e diz que vai oferecer livros também impressos para alunos*. Recuperado de <https://www.cnnbrasil.com.br/nacional/governo-de-sp-recua-e-diz-que-vai-oferecer-livros-tambem-impressos-para-alunos/>
- Comitê Gestor da Internet no Brasil - CGI.br (2022). *Educação em um cenário de plataformação e de economia dos dados: problemas e conceitos*. Evangelista, R. A. (coord). Recuperado de https://cgi.br/media/docs/publicacoes/1/20220929112852/educacao_em_um_cenario_de_plataformizacao_e_de_economia_de_dados_problemas_e_conceitos.pdf
- Comitê Gestor da Internet no Brasil - CGI.br (2022). *Pesquisa sobre o acesso, o uso e a apropriação das tecnologias de informação e comunicação (TIC) em escolas públicas e privadas de educação básica brasileira: TIC Educação 2022*. São Paulo: CGI.br. Recuperado de <https://cetic.br/pt/pesquisa/educacao/indicadores/>
- Constituição da República Federativa do Brasil (1988). Recuperado de https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm
- Data Privacy Brasil – DPB (2023). *IA na sala de aula: modelos de participação para a comunidade escolar*. Recuperado de <https://conteudos.dataprivacy.com.br/cartilha-ia-na-sala-de-aula>
- Da Hora, N. (2023). *Inteligência Artificial e Educação no Brasil: potencialidades e limitações*. MIT Technology Review. Recuperado de <https://mittechreview.com.br/inteligencia-artificial-e-educacao-no-brasil-potencialidades-e-limitacoes/>
- Dallari, D. In Cury, M., Garrido, P. A., & Marçura, J. N. (2002). *Estatuto da criança e do adolescente anotado*. In *Estatuto da criança e do adolescente anotado*, p. 553.

- Decreto n. 99.710/1990. Convenção sobre os Direitos da Criança. Recuperado de https://www.planalto.gov.br/ccivil_03/decreto/1990-1994/d99710.htm
- Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational research review*, 25, pp. 23-38. Recuperado de https://www.researchgate.net/publication/330854760_Don't_throw_away_your_printed_books_A_meta-analysis_on_the_effects_of_reading_media_on_reading_comprehension
- De Paiva, N. M. N., & Costa, J. (2015). A influência da tecnologia na infância: desenvolvimento ou ameaça. *Psicologia*, pt, 1, pp. 1-13.
- Fernandes, Sofia. (2024). *O que é design viciante e por que a UE quer limitar seu uso?* Folha Uol. Recuperado de <https://www1.folha.uol.com.br/cotidiano/2024/06/o-que-e-design-viciante-e-por-que-a-ue-quer-limitar-seu-uso.shtml>
- Folha de São Paulo (2023, julho). *SP abre mão de verba para material didático e usará só livro digital a partir do 6º ano*, Recuperado de <https://www1.folha.uol.com.br/educacao/2023/07/sp-fica-fora-de-programa-nacional-de-livro-didatico-e-ira-usar-so-conteudo-proprio-em-escolas.shtml>
- Fundo das Nações Unidas para a Infância - UNICEF (2006). *Relatório anual do Fundo das Nações Unidas para a infância: crianças de até 6 anos, o direito à sobrevivência e ao desenvolvimento*. Recuperado de https://crianca.mppr.mp.br/arquivos/File/publi/unicef_sowc/sit_inf_brasil_2006_completo.pdf
- Gawer, A. R., & Srnicek, N. (2021). *Online platforms: Economic and societal effects*. Panel for the Future of Science and Technology (STOA) European Parliament.
- Gonsales, P., & Amiel, T. (2020). Educação na contemporaneidade: entre dados e direitos. *Panorama Setorial da Internet*. São Paulo, (3), pp. 1-7. Recuperado de https://cetic.br/media/docs/publicacoes/6/20201110120042/panorama_setorial_anoxii_n_3_inteligencia_artificial_educacao_infancia.pdf
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in education: promises and implications for teaching and learning*. Center for Curriculum Redesign. Recuperado de https://www.researchgate.net/publication/332180327_Artificial_Intelligence_in_Education_Promise_and_Implications_for_Teaching_and_Learning
- Hoogeveen, M. (1997). Toward a theory of the effectiveness of multimedia systems. *International journal of human-computer interaction*, 9(2), pp.151-168.
- Hooper, L., Livingstone, S., & Pothong, K. (2022). *Problems with data governance in UK schools: The cases of Google Classroom and ClassDojo*. Digital Futures Commission, 5Rights Foundation, London, UK.
- Human Rights Watch - HRW (2022). "How Dare They Peep into My Private Life?" *Children's Rights Violations by Governments that Endorsed Online Learning During the Covid-19 Pandemic*. Nova Iorque: HRW.Org. Recuperado de <https://www.hrw.org/report/2022/05/25/how-dare-they-peep-my-private-life/childrens-rights-violations-governments>
- Instituto Alana (2024, maio). *Inteligência Artificial para a prioridade absoluta dos direitos e do melhor interesse de crianças e adolescentes no Brasil*. Contribuições do Instituto Alana ao substitutivo preliminar do PL 2.338/2023, divulgado em 25.4.2024. Recuperado de <https://criancaconsumo.org.br/biblioteca/nota-tecnica-do-instituto-alana-sobre-o-pl-2338/>

- Lei n. 8.069/1990. (1990). Estatuto da Criança e do Adolescente (ECA). Recuperado de https://www.planalto.gov.br/ccivil_03/decreto/1990-1994/d99710.htm
- Lei n. 9.394/1996. (1996). Estabelece as diretrizes e bases da educação nacional. Recuperado de https://www.planalto.gov.br/ccivil_03/leis/l9394.htm
- Lei n. 12.965/ 2014. (2014). Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Recuperado de https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm
- Lei n. 13.709/2018. (2018). Lei Geral de Proteção de Dados (LGPD). Recuperado de https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm
- Ministério da Saúde (2022). *Primeira infância*. Recuperado de <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/s/saude-da-crianca/primeira-infancia#:~:text=A%20primeira%20inf%C3%A2ncia%20C3%A9%20o,os%206%20anos%20de%20idade.>
- Organização das Nações Unidas para a Educação, a Ciência e a Cultura - UNESCO (2016). *UNESCO International Bureau of Education*. Recuperado de <https://unesdoc.unesco.org/ark:/48223/pf0000243975.locale=en>
- Organization for Economic Cooperation and Development - OECD (2019). *Recommendation of the Council on Artificial Intelligence*. Recuperado de <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
- Poell, T., Nieborg, D., & Van Dijck, J. (2019). *Plataformização [Platformisation, 2019 – tradução Rafael Grohmann]*. *Revista Fronteiras – estudos midiáticos* (2020) v. 22(1), pp. 2-10
- Projeto de Lei n. 2.338/2023. (2023). *Dispõe sobre o uso da Inteligência Artificial*. Recuperado de <https://legis.senado.leg.br/sdleg-getter/documento?dm=9347622&ts=1694638936361&disposition=inline>
- Sé, L. (2020). *Reconhecimento de emoções? Por que cientistas de IA acham isso errado*. Uol Tilt. Recuperado de <https://www.uol.com.br/tilt/noticias/redacao/2020/01/04/reconhecimento-de-emocoes-por-que-cientistas-de-ia-acham-isso-errado.htm>
- Silva, T. D. O., & Silva, L. T. G. (2016). *Os impactos sociais, cognitivos e afetivos sobre a geração de adolescentes conectados às tecnologias digitais*.
- Tomasello, M., & Call, J. (1997). *Primate cognition*. Oxford University Press, USA.
- Van Dijck, J., Poell, T., & De Waal, M. (2018). *The platform society: public values in a connective world*. New York: Oxford University Press.
- Wolf, M. (2019). *O cérebro no mundo digital: os desafios da leitura na nossa era*. Editora Contexto.
- Zuboff, S. (2021). *A era do capitalismo de vigilância*. Editora Intrínseca.
- Zuboff, S. (2019). *Shoshana Zuboff on Surveillance Capitalism's Threat to Democracy*. [Entrevista concedida a Noah Kelwin]. *New York Magazine*. Recuperado de <https://nymag.com/intelligencer/2019/02/shoshana-zuboff-q-and-a-the-age-of-surveillance-capital.html#:~:text=There's%20data%20and%20information%20collected,control%20and%20power%20that%20these>

Notas finais

1 Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação – Cetic.br (2022). *Resumo Executivo TIC Kids Online Brasil 2022*. Recuperado de <https://cetic.br/pesquisa/kids-online/>

2 Human Rights Watch (2022). “How dare they peep into my private life?” Children’s Rights Violations by Governments that Endorsed Online Learning During the Covid-19 Pandemic. Nova Iorque: HRW.Org. Recuperado de <https://www.hrw.org/report/2022/05/25/how-dare-they-peep-my-private-life/childrens-rights-violations-governments>

3 Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in education: promises and implications for teaching and learning*. Center for Curriculum Redesign. Recuperado de https://www.researchgate.net/publication/332180327_Artificial_Intelligence_in_Education_Promise_and_Implications_for_Teaching_and_Learning

4 Holmes et al. (2019), p.153.

5 Holmes et al. (2019), p.159.

6 Classificação feita por Dr. Ruben Puentedura pelo modelo SAMR (Substitution, Augmentation, Modification, and Redefinition model) dos mecanismos de IA aplicados na educação, demonstrando o maior impacto crescente dos modelos de Redefinição à Substituição. Recuperado de <http://www.hippasus.com/about.html>

7 Estatuto da Criança e do Adolescente, Art. 53. “A criança e o adolescente têm direito à educação, visando ao pleno desenvolvimento de sua pessoa, preparo para o exercício da cidadania e qualificação para o trabalho (...)”. Recuperado de https://www.planalto.gov.br/ccivil_03/leis/l8069.htm

8 Lei n. 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. Brasília: Presidência da República, 1996. Recuperado de http://www.planalto.gov.br/ccivil_03/leis/19394.htm

9 Folha de São Paulo (2023, julho). *SP abre mão de verba para material didático e usará só livro digital a partir do 6º ano*, Recuperado de <https://www1.folha.uol.com.br/educacao/2023/07/sp-fica-fora-de-programa-nacional-de-livro-didatico-e-ira-usar-so-conteudo-proprio-em-escolas.shtml>

10 CNN Brasil (2023, agosto). *Governo de SP recua e diz que vai oferecer livros também impressos para alunos*. Recuperado de <https://www.cnnbrasil.com.br/nacional/governo-de-sp-recua-e-diz-que-vai-oferecer-livros-tambem-impressos-para-alunos/>

11 Barbosa, R., & Blikstein, P. (2019). *Robótica educacional: experiências inovadoras na educação brasileira*. Penso Editora.

12 Os Projetos de Leis que buscam regulamentar o uso da Inteligência Artificial tanto União Europeia, quanto no Brasil, trazem uma preocupação social com a distribuição adequada dos benefícios da IA que é expressa em termos de justiça e equidade. O *Statement on Artificial Intelligence, Robotics and Autonomous Systems* da Comissão Europeia defende que a IA deve “contribuir para a justiça global e a igualdade de acesso aos benefícios” dessa tecnologia. Recuperado de https://research-and-innovation.ec.europa.eu/research-area/industrial-research-and-innovation/key-enabling-technologies/artificial-intelligence-ai_en

13 Os *Princípios Asilomar AI*, coordenados pela FLI e desenvolvidos na *Conferência Beneficial AI 2017* também se preocupam com a aplicação dos valores de benefício compartilhado. Logo, as tecnologias de IA devem beneficiar e capacitar o maior número possível de pessoas a fim de beneficiar amplamente toda a humanidade. Recuperado de <https://futureoflife.org/open-letter/ai-principles/>

- 14 Comitê Gestor da Internet no Brasil (2022). *Pesquisa sobre o acesso, o uso e a apropriação das tecnologias de informação e comunicação (TIC) em escolas públicas e privadas de educação básica brasileira: TIC Educação 2022*. São Paulo: cetic.br/pt/pesquisa/educacao/indicadores/
- 15 Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação - Cetic.br (2022). *Pesquisa sobre o acesso, o uso e a apropriação de tecnologias digitais em escolas públicas e particulares brasileiras que ofereçam Ensino Fundamental e Médio: TIC Educação 2022 – Coletiva de Imprensa*. Recuperado de https://cetic.br/media/analises/tic_educacao_2022_coletiva%20de%20imprensa.pdf
- 16 Marco Civil da Internet (2014). Lei n. 12.965/ 2014. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Recuperado de https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm
- 17 Holmes et al. (2019), p. 188.
- 18 Audi, A. (2023). *Reconhecimento facial no Paraná impõe monitoramento de emoções em escolas*. Apública. Recuperado de <https://apublica.org/2023/10/reconhecimento-facial-no-parana-impoe-monitoramento-de-emocoes-em-escolas/>
- 19 Ministério da Saúde (2022). *Primeira infância*. Recuperado de <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/s/saude-da-crianca/primeira-infancia#:~:text=A%20primeira%20inf%C3%A2ncia%20%C3%A9%20o,os%206%20anos%20de%20idade>
- 20 Fundo das Nações Unidas para a Infância - UNICEF (2006). *Relatório anual do Fundo das Nações Unidas para a infância: crianças de até 6 anos, o direito à sobrevivência e ao desenvolvimento*. Recuperado de https://crianca.mppr.mp.br/arquivos/File/publi/unicef_sowc/sit_inf_brasil_2006_completo.pdf
- 21 Constituição Federal de 1988, art. 227. “É dever da família, da sociedade e do Estado assegurar à criança, ao adolescente e ao jovem, com absoluta prioridade, o direito à vida, à saúde, à alimentação, à educação, ao lazer, à profissionalização, à cultura, à dignidade, ao respeito, à liberdade e à convivência familiar e comunitária, além de colocá-los a salvo de toda forma de negligência, discriminação, exploração, violência, crueldade e opressão.” Recuperado de https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm
- 22 Marco Civil da Internet (2014a).
- 23 Marco Civil da Internet (2014b, art. 26). “O cumprimento do dever constitucional do Estado na prestação da educação, em todos os níveis de ensino, inclui a capacitação, integrada a outras práticas educacionais, para o uso seguro, consciente e responsável da internet como ferramenta para o exercício da cidadania, a promoção da cultura e o desenvolvimento tecnológico.” Recuperado de https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm
- 24 Tomasello, M., & Call, J. (1997). *Primate cognition*. Oxford University Press, USA.
- 25 Silva, T. D. O., & Silva, L. T. G. (2016). Os impactos sociais, cognitivos e afetivos sobre a geração de adolescentes conectados às tecnologias digitais. *Revista Psicopedagogia*. 4(103), pp. 87-97. Recuperado de http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0103-84862017000100009&lng=pt&tlng=pt
- 26 Hoogeveen, M. (1997). Toward a theory of the effectiveness of multimedia systems. *International journal of human-computer interaction*, 9(2), pp.151-168.

27 De Paiva, N. M. N., & Costa, J. (2015). A influência da tecnologia na infância: desenvolvimento ou ameaça. *Psicologia.pt – o portal dos psicólogos*. Teresina, Vol.1 pp. 1-13. Recuperado de <https://www.psicologia.pt/artigos/textos/A0839.pdf>

28 Da Hora, N. (2023). Inteligência Artificial e Educação no Brasil: potencialidades e limitações. *MIT Technology Review*. Recuperado de <https://mittechreview.com.br/inteligencia-artificial-e-educacao-no-brasil-potencialidades-e-limitacoes/>

29 Constituição Federal de 1988. (1988). Recuperado de https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm

30 Constituição Federal de 1988, art. 227. “É dever da família, da sociedade e do Estado assegurar à criança, ao adolescente e ao jovem, com absoluta prioridade, o direito à vida, à saúde, à alimentação, à educação, ao lazer, à profissionalização, à cultura, à dignidade, ao respeito, à liberdade e à convivência familiar e comunitária, além de colocá-los a salvo de toda forma de negligência, discriminação, exploração, violência, crueldade e opressão.” Recuperado de https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm

31 Lei n. 8.069/1990. (1990). Estatuto da Criança e do Adolescente (ECA).

32 Lei n. 13.709/2018. (2018). Lei Geral de Proteção de Dados (LGPD).

33 Lei n. 13.709/2018. (2018). Lei Geral de Proteção de Dados (LGPD). Seção III, Do Tratamento de Dados Pessoais de Crianças e de Adolescentes, Art. 14. “O tratamento de dados pessoais de crianças e de adolescentes deverá ser realizado em seu melhor interesse, nos termos deste artigo e da legislação pertinente.” Recuperado de https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm

34 *Enunciado CD/ANPD nº 1, de 22 de maio de 2023*. (2023). Diário Oficial da União. Recuperado de <https://www.gov.br/anpd/pt-br/asuntos/noticias/anpd-divulga-enunciado-sobre-o-tratamento-de-dados-pessoais-de-criancas-e-adolescentes/Enunciado1ANPD.pdf/view>

35 Lei Geral de Proteção de Dados (LGPD). Lei n. 13.709/2018. (2018). Recuperado de https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm

36 Gawer, A. R., & Srnicek, N. (2021). *Online platforms: Economic and societal effects*. Panel for the Future of Science and Technology (STOA) European Parliament.

37 Hooper, L., Livingstone, S., & Pothong, K. (2022). *Problems with data governance in UK schools: The cases of Google Classroom and ClassDojo*. Digital Futures Commission, 5Rights Foundation, London, UK.

38 O aplicativo Ruangguru (ruangguru.com), desenvolvido na Indonésia, em sua própria descrição, dispõe que “quaisquer que sejam seus desafios de aprendizado, #Não tenha medo, existe Ruangguru que tem soluções para todas as suas necessidades de aprendizagem. A partir de vídeos de aprendizagem, questões práticas para estudar com os melhores professores, podem ser acessados diretamente em um aplicativo”. Disponível para ensino fundamental, médio e superior + níveis UTBK. Recuperado de https://play.google.com/store/apps/details?id=com.ruangguru.livestudents&hl=pt_BR&gl=US

39 A autodeterminação informativa é fundamento da disciplina da proteção de dados pessoais (art. 2º, inciso II da LGPD). Refere-se ao direito que cada indivíduo tem de controlar e proteger seus dados pessoais e que tenha domínio sobre suas próprias informações, *ex*

vi Lei n. 13.709/2018. Lei Geral de Proteção de Dados (LGPD). Recuperado de https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm

40 Assembleia Legislativa do Estado de São Paulo - Alesp (2023, agosto). *Seduc instala aplicativo sem consentimento em celulares pessoais de professores e alunos*. Recuperado de <https://www.al.sp.gov.br/noticia/?10/08/2023/seduc-instala-aplicativo-sem-consentimento-em-celulares-pessoais-de-professores-e-alunos>

41 A Lei Geral de Proteção de Dados (LGPD) determina, no art. 6º, os princípios que devem ser observados para as atividades de tratamento de dados pessoais, entre eles: II - adequação: compatibilidade do tratamento com as finalidades informadas ao titular, de acordo com o contexto do tratamento; VI - transparência: garantia, aos titulares, de informações claras, precisas e facilmente acessíveis sobre a realização do tratamento e os respectivos agentes de tratamento, observados os segredos comercial e industrial. Recuperado de https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm

42 Lei n. 13.709/2018. (2018). Lei Geral de Proteção de Dados (LGPD). Seção III, Do Tratamento de Dados Pessoais de Crianças e de Adolescentes, Art. 14. “O tratamento de dados pessoais de crianças e de adolescentes deverá ser realizado em seu melhor interesse, nos termos deste artigo e da legislação pertinente.” Recuperado de https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm

43 Zuboff, S. (2019, fevereiro). Shoshana Zuboff on Surveillance Capitalism’s Threat to Democracy. [Entrevista concedida a Noah Kelwin]. *New York Magazine*. Recuperado de <https://nymag.com/intelligencer/2019/02/shoshana-zuboff-q-and-a-the-age-of-surveillance-capital.html#:~:text=There's%20data%20and%20information%20collected,control%20and%20power%20that%20these>

44 A Lei Geral de Proteção de Dados (LGPD) determina no art. 14: “o tratamento de dados pessoais de crianças e de adolescentes deverá ser realizado em seu melhor interesse, nos termos deste artigo e da legislação pertinente; § 4º Os controladores não deverão condicionar a participação dos titulares de que trata o § 1º deste artigo em jogos, aplicações de internet ou outras atividades ao fornecimento de informações pessoais além das estritamente necessárias à atividade.” Em observância a essa disposição, é vedado a exigência de informações pessoais além do estritamente necessário. Recuperado de https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm

45 Conforme a LGPD, art. 42, §1º, II, quando mais de um controlador estiver diretamente envolvido no tratamento do qual decorram danos ao titular de dados, estes responderão de forma solidária. Recuperado de https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/2021.05.27GuiaAgentesdeTratamento_Final.pdf

46 O operador dos dados é o agente responsável por realizar o tratamento de dados em nome do controlador e conforme a finalidade por este delimitada. Muito embora o controlador tenha a principal responsabilidade e o operador deva atuar em nome dele, o art. 37 da LGPD determina que ambos partilham obrigações. Nos termos do art. 42 da LGPD, ambos possuem a obrigação de reparação se causarem dano patrimonial, moral, individual ou coletivo a outrem, no âmbito de suas respectivas esferas de atuação. Recuperado de https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/2021.05.27GuiaAgentesdeTratamento_Final.pdf

ARTIGO

Erisictão Já Invadiu o Bosque de Deméter: análise dos riscos sociais do colapso de IAs por retroalimentação

Pedro Khauaja

pedrokhauaja@gmail.com

Advogado; Mestre e Doutorando pelo Programa de Pós-Graduação em Sociologia e Direito da Universidade Federal Fluminense (PPGSD-UFF); editor-assistente da revista CONFLUÊNCIAS; professor substituto no Departamento de Direito Privado da Universidade Federal Fluminense (UFF).

Erisictão Já Invadiu o Bosque de Deméter: análise dos riscos sociais do colapso de IAs por retroalimentação

Palavras-chave

Retroalimentação de IAs

Colapso de sistemas

Risco

Modernidade tardia

Inteligência Artificial

Resumo

Este artigo analisa os riscos sociais que surgem da possibilidade de colapso de sistemas de Inteligência Artificial (IA) por retroalimentação. A retroalimentação ocorre quando IAs generativas são treinadas com dados gerados por outras IAs, o que pode levar ao colapso desses sistemas, conforme demonstrado por Shumailov et al. (2024) e Martinez et al. (2024). A pesquisa toma como ponto de partida o conceito de “risco reflexivo” de Beck (2011), aplicando-o ao contexto tecnológico contemporâneo. A metodologia inclui uma revisão teórica e bibliográfica sobre os riscos associados ao uso crescente de IAs, e uma análise dos dados quantitativos fornecidos pelo Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br), que evidenciam o aumento da presença de IA no Brasil, assim como outros trabalhos que apontam o mesmo

em outros países, muitas vezes sem transparência sobre a origem dos conteúdos gerados. Os resultados indicam que a retroalimentação de IA não só representa um risco técnico para o colapso dos sistemas, mas também um risco social, devido à forma desigual com que os impactos negativos são distribuídos. A análise conclui que o fenômeno de colapso das IAs reflete um padrão característico da modernidade tardia, onde os benefícios tecnológicos e econômicos são concentrados, enquanto os danos potenciais são externalizados, especialmente em contextos de maior vulnerabilidade.

Erysichthon Has Already Invaded Demeter's Grove: an analysis of the social risks of AI collapse due to feedback

Keywords

AI feedback loop
System collapse;
Risk
Late modernity
Artificial Intelligence

Abstract

This paper analyzes the social risks arising from the potential collapse of Artificial Intelligence (AI) systems due to feedback loops. A feedback loop occurs when generative AIs are trained with data generated by other AIs, which can lead to the collapse of these systems, as demonstrated by Shumailov et al. (2024) and Martinez et al. (2024). The research is based on Beck's (2011) concept of "reflexive risk," applying it to the contemporary technological context. The methodology includes a theoretical and bibliographic review of the risks associated with the growing use of AIs and an analysis of quantitative data provided by the Regional Center for Studies on the Development of the Information Society (Cetic.br), which highlights the increasing presence of AI in Brazil, as well as other studies pointing to the same trend in other countries, often without transparency regarding the origin of

generated content. The results indicate that AI feedback not only represents a technical risk for system collapse but also a social risk, due to the unequal distribution of negative impacts. The analysis concludes that the phenomenon of AI collapse reflects a pattern characteristic of late modernity, where technological and economic benefits are concentrated, while potential harms are externalized, particularly in more vulnerable contexts.

1 Introdução

O mito de Erisictão é contado por Ovídio em seu livro “Metamorfoses” como parte do grande acervo de histórias mitológicas gregas. Essa é mais uma que trata do desrespeito aos deuses, da ira divina, e do alto preço da soberba e falta de controle sobre as próprias pulsões. Erisictão era um rei da Tessália conhecido por sua extrema avareza e pelo seu desrespeito às divindades. Em um ato de grande insolência, decidiu cortar uma árvore sagrada de Deméter em um bosque dedicado à deusa da agricultura.

A árvore era considerada sagrada e fazia parte do domínio de Deméter, responsável pela fertilidade da terra e pelo ciclo das colheitas. Em resposta ao ato sacrílego, a deusa decidiu punir Erisictão com uma maldição de fome insaciável. Essa fome fez com que o rei gastasse toda a sua fortuna e posses, e eventualmente começasse a devorar sua própria carne, tamanho nível de seu desespero. A história termina com Erisictão sendo consumido pela própria fome, um destino cruel e simbólico que, naquele contexto, reforçava a mensagem sobre a importância de respeitar os deuses e o seu domínio (Lobato Martins, 2023).

Essa estória, porém, carrega um outro sentido, para além da mensagem e cuidado com os deuses. É mais uma das várias tentativas de, através do mito, avisar dos perigos da gula e da avareza, alertando para o eventual destino de todo descontrole consumista: consumir a si mesmo. Nesse artigo, a estória de Erisictão serve como uma metáfora para o tema central, que é a análise dos riscos sociais implicados nas evidências recentes de que os sistemas de Inteligência Artificial (IA), especialmente as IAs generativas, correm alto risco de colapso quando alimentadas com dados criados por outras IAs.

Essa prática de utilizar dados de uma IA para alimentar o treinamento de outra é o que chamo de retroalimentação. Nesse texto utilizo

o termo retroalimentação como uma tradução possível para “*feedback loop*”, termo que foi utilizado nos primeiros trabalhos sobre o tema, em inglês., com especial atenção para Shumailov et al. (2024) e Martínez et al. (2023). No geral, um “*feedback loop*” é exatamente esse processo em que uma determinada IA recebe dados de outra como se fossem produzidos por humanos. Essa aparência de humanidade se mostrará um elemento essencial para o risco de colapso.

O trabalho de Shumailov et al. (2024) foi o primeiro a apontar essa possibilidade e tem tido muita repercussão, mas outros já vieram confirmar o mesmo efeito em contextos similares, como Martínez et al. (2024). Tomando como pressuposto essa descoberta, qual seja, de que as IAs tendem ao colapso quando retroalimentadas, abordarei o risco social que origina desse fato, quando somado ao amplo espectro de usos das IAs generativas hoje. Farei essa abordagem de análise de risco a partir do arcabouço teórico de Beck (2011) e sua caracterização da modernidade como período de risco reflexivo, distribuído desigualmente e ocultado deliberadamente.

Assim, primeiro trabalharei algumas conceituações iniciais sobre o colapso de IAs por retroalimentação, qualificando o debate ainda novo nas produções brasileiras. Logo em seguida, também trarei dados que corroboram a ideia de que há um aumento perigoso de conteúdo gerado por IA na Internet, e especialmente um conteúdo sem nenhum tipo de aviso ou indicativo de sua origem. Depois, com uso principalmente dos relatórios de pesquisa produzidos pelo Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br)¹, vou trabalhar com a hipótese de um aumento da presença de tecnologias de IA no ecossistema tecnológico brasileiro, e a natureza delicada e sensível de alguns de seus usos.

Esse cenário servirá de base para a análise feita ao final do artigo que, partindo da conceituação teórica de Beck (2011), proporá que o risco de colapso das IAs é característico da

modernidade tardia, funcionando como uma forma desigual de centralizar lucros e compartilhar danos.

2 O Colapso de IAs por retroalimentação

Essa pesquisa se baseia largamente no trabalho de Schumailov et al. (2024) e, portanto, trabalharei com uma conceituação técnica que esteja de acordo com a dos autores. Ainda que no artigo não estejam definidos de forma expressa todos os termos, os autores estão trabalhando com um tipo específico de tecnologia: os *Large Language Models* (LLMs), que são um desdobramento das tecnologias chamadas de IAs generativas.

As Inteligências Artificiais generativas são um subcampo da inteligência artificial focado na criação de modelos capazes de gerar novos dados a partir de um conjunto de dados existente (Ludermir, 2021). Esses sistemas não apenas analisam ou processam informações, mas são projetados para criar conteúdo supostamente “novo”, como texto, imagens, áudio, e até mesmo vídeo, que são indistinguíveis ou dificilmente distinguíveis dos produzidos por seres humanos, como mostra a *survey* de Wu et al. (2023). A geração de conteúdo por essas IAs é baseada em aprendizado de padrões e estruturas presentes nos dados de treinamento, permitindo que elas produzam resultados criativos ou inovadores, sem intervenção humana direta.

Diferentes tipos de IAs generativas têm aplicações vastas e variadas, e não necessariamente funcionam da mesma forma que os LLMs (Ludermir, 2021). Os Modelos de Linguagem de Grande Escala, conhecidos como LLMs, são sistemas de inteligência artificial projetados para compreender, gerar, emular e manipular linguagem natural em grande escala (Naveed et al., 2024). Esses modelos são baseados em

arquiteturas avançadas de redes neurais, como os transformadores, e são treinados em grandes volumes de dados textuais para captar e replicar padrões linguísticos complexos.

O treinamento de LLMs envolve a exposição a vastas quantidades de texto, proveniente de diversas fontes, como livros, artigos, sites e outros documentos. O objetivo é permitir que o modelo aprenda a prever a próxima palavra mais provável em uma sequência de texto, identificar relações contextuais entre palavras e entender o significado geral do texto. Esse processo de treinamento permite que os LLMs aparentem ter uma compreensão profunda da estrutura e do uso da linguagem, enquanto, ao mesmo tempo, emulam um padrão-base humano de respostas (Naveed et al., 2024).

Os LLMs mais avançados, como GPT-3 e GPT-4, são exemplos de modelos que possuem bilhões ou até trilhões de parâmetros, o que lhes confere uma capacidade significativa para capturar nuances e complexidades da linguagem (Colombo e Goulart, 2023). A escala desses modelos permite um desempenho impressionante em várias tarefas linguísticas, mas também levanta questões sobre a ética, a privacidade e o potencial para vieses nos dados de treinamento, assim como debates sobre a legitimidade da sua atividade, com acusações de plágio e outras similares (Wu et al., 2023).

O colapso de sistemas de Inteligência Artificial por retroalimentação se refere a um fenômeno em que sistemas de inteligência artificial, especialmente aqueles baseados em aprendizado de máquina, sofrem um progressivo declínio em sua qualidade de conteúdo gerado por conta do uso repetitivo de seus próprios *outputs*, ou de outras IAs, como *inputs* para futuros processos de treinamento. Esse fenômeno pode resultar em uma degeneração da qualidade das previsões ou decisões produzidas pelo sistema, levando a um ciclo de deterioração autoalimentado, conforme definido por Schumailov et al. (2024).

Os autores apontam que esse tipo de colapso ocorre principalmente em ambientes onde a IA é usada para gerar ou filtrar dados que, posteriormente, são reutilizados no processo de treinamento ou operação contínua do mesmo sistema ou de outros sistemas interligados. Quando esse processo não é cuidadosamente gerido ou monitorado, os erros ou vieses presentes nos *outputs* podem ser amplificados ao longo do tempo, resultando em uma perda de robustez e precisão.

A essência do trabalho de Schumailov et al. (2024), é que quanto maior a presença de substrato gerado por IAs na base de treinamento dos LLMs, maior é a deterioração da qualidade do conteúdo gerado ao final, e maior o risco de um colapso do modelo, que foi definido como sendo um processo degenerativo em que a IA erra a distribuição original dos dados, e passa a reproduzir algo sem sentido. Os autores apontam que, intuitivamente, é provável que esse fenômeno se repita em outras IAs generativas além dos LLMs, e ofereça um risco ainda maior do que o evidenciado.

Matinez et al (2023), que usam de fato o termo “*feedback loop*”, encontraram essencialmente o mesmo resultado com as IAs generativas de imagem, como o DALL-E. Nesse trabalho, cujos parâmetros de metodologia são um pouco menos rígidos que o anterior, os autores essencialmente retroalimentaram uma IA geradora de imagem com imagens de outras IAs a fim de treinamento. Isso gerou imagens progressivamente piores conforme maior presença de imagens de IA no *pool* de treino.

A conclusão é similar: quanto maior a porcentagem de conteúdo gerado por IA presente na hora do treinamento, menor a qualidade, em termos de um parâmetro de avaliação humano, do resultado gerado.

2.1 Aumento da presença de conteúdo gerado por IAs na rede

Schumailov et al. (2024) já terminam seu artigo apontando o perigo que o aumento exponencial de conteúdo gerado por IA cria no contexto do colapso de modelos LLMs. É com base nesse *insight* que desenvolverei agora um panorama geral dessa situação para compreendermos que há, de fato, um risco grande da situação se agravar.

O aumento recente na quantidade de conteúdos gerados por IAs na web é um fenômeno impulsionado por uma combinação de avanços tecnológicos, demanda crescente por automação e a busca por eficiência na produção de conteúdo digital (Wu et al., 2023). Vários fatores têm contribuído para essa proliferação de conteúdo gerado por IA, que vão desde melhorias nos algoritmos de aprendizado de máquina até a expansão do acesso a dados e o desenvolvimento de plataformas mais acessíveis e poderosas para criação automatizada.

Primeiramente, os avanços em aprendizado profundo e técnicas de processamento de linguagem natural (NLP) têm sido fundamentais para o aumento da capacidade das IAs em gerar textos que se assemelham aos produzidos por humanos. Modelos como o GPT-3 e suas iterações subsequentes são bons exemplos de transformadores que conseguem produzir textos longos e coesos a partir de um conjunto limitado de informações ou comandos (Colombo e Goulart, 2023). Esses modelos são treinados em grandes volumes de dados textuais disponíveis na Internet, permitindo-lhes capturar nuances, estilos e contextos variados. A capacidade desses modelos de gerar texto de maneira quase indistinguível do escrito por humanos resultou em sua adoção ampla para diversas finalidades, incluindo blogs, artigos de notícias,

resumos de textos, e até mesmo ficção. (Naveed et al., 2024, Vayadande et al., 2024)

Além desses avanços técnicos, a demanda por automação e escalabilidade na produção de conteúdo também desempenha um papel crucial (Naveed et al., 2024). Empresas de marketing digital, jornalismo e outros setores baseados em conteúdo estão cada vez mais voltadas para soluções que permitam a produção rápida e eficiente de grandes volumes de texto, sem comprometer a qualidade. A utilização de IAs generativas permite às empresas manterem um fluxo constante de conteúdo atualizado, melhorando sua presença online e otimização para mecanismos de busca (SEO), com menos dependência de trabalhadores humanos e menor custo operacional.

Outro fator importante é o aumento do acesso a essas tecnologias. Plataformas como GPT-3 foram disponibilizadas através de APIs, facilitando que desenvolvedores e empresas integrem modelos de linguagem avançados em seus aplicativos e websites. Com uma barreira de entrada mais baixa, um número maior de indivíduos e organizações têm a capacidade de explorar o potencial das IAs para geração de conteúdo. Além disso, ferramentas de IA foram incorporadas a softwares de criação de conteúdo, como editores de texto e geradores de imagem, tornando a geração de conteúdo automatizado mais simplificada.

Nessa pesquisa, para além de questões sobre qualidade desse conteúdo ou a eticidade da sua produção (Colombo e Goulart, 2023), o ponto de maior preocupação é a saturação do ambiente digital com conteúdo gerado por IAs. Como produzem grandes volumes de texto rapidamente, há um risco de inundação de conteúdo repetitivo ou sem valor agregado, mas com toda aparência de conteúdo humano. Além de afetar a confiança dos usuários na veracidade e qualidade dos conteúdos que encontram na web, dificultam muito a diferenciação para construção de bases de treinamento das

próprias IAs—daí o risco de colapso dos modelos (Schumailov et al., 2024, e Martinez et al., 2023).

O aumento do risco de colapso de modelos de IA por retroalimentação pode ser compreendido através de uma análise das dinâmicas de como os modelos de IA são treinados e atualizados, bem como das implicações de uma proliferação massiva de conteúdo automatizado. Primeiro a retroalimentação de dados é um aspecto já inserido no treinamento e aprimoramento dos modelos de IA (Martinez et al., 2023). Esses sistemas frequentemente utilizam dados gerados por si mesmos ou por outros sistemas de IA como entrada para treinamento contínuo, mas de forma acessória. Quando a quantidade de conteúdo gerado por IA aumenta exponencialmente, o volume de dados disponíveis para treinamento pode crescer desproporcionalmente.

Além disso, a proliferação de conteúdo gerado por IA pode levar a uma concentração maior de informações semelhantes ou repetitivas na web (Wu et al., 2023). Modelos de IA, especialmente os baseados em técnicas de aprendizado supervisionado, são treinados em dados históricos que refletem o que já existe na internet. Se uma grande parte do novo conteúdo gerado é similar ou derivada de modelos anteriores, o treinamento subsequente pode ser influenciado por padrões redundantes e pouco diversos. Isso pode resultar na produção de outputs cada vez mais homogêneos e menos inovadores, uma vez que os modelos tendem a se consolidar em padrões repetitivos e menos diversificados.

Outro fator crítico é o impacto da desinformação e dos conteúdos manipulativos (Westerlund, 2019; e Vayadande et al., 2024). Modelos de IA que geram conteúdo para fins de propaganda, desinformação ou outros usos mal-intencionados podem contaminar o pool de dados disponível para treinamento. Quando esses conteúdos prejudiciais são amplamente

disseminados e usados para treinar novos modelos, eles podem contribuir para uma propagação de desinformação e manipulação em larga escala (Mariani, 2023). Isso não só compromete a integridade dos dados utilizados para treinamento, mas também amplifica o risco de que os modelos perpetuem e disseminem desinformação de forma mais intensa.

Adicionalmente, a integração de modelos de IA em sistemas complexos que interagem entre si pode criar redes de retroalimentação mais complexas e difíceis de controlar. Por exemplo, se vários sistemas de recomendação utilizam dados gerados por IA como entrada para suas próprias operações, a interação entre esses sistemas pode resultar em um feedback loop onde o conteúdo gerado se torna cada vez mais tendencioso ou enviesado. Esse fenômeno pode levar a um colapso progressivo na qualidade do conteúdo, onde as decisões e recomendações baseadas em dados automatizados tornam-se cada vez mais insatisfatórias ou prejudiciais.

As IAs geradoras de imagem e vídeo, integradas com outras como LLMs potentes, formam a base do que Westerlund (2019) chama de *deepfake technology*, as ferramentas para emular quase perfeitamente pessoas através de vídeo e áudio e, se devidamente treinadas com uma base de dados suficientemente larga, até mesmo imitar a forma com a pessoa se comunica, cacoetes, maneirismos etc. Vayadande et al. (2024), apontam uma proliferação violenta desses conteúdos, especialmente em contextos de manipulação política e midiática como em eleições.

O problema, em termos de colapso de sistemas de IA, é que esse conteúdo pode passar despercebido e entrar no pool de treinamento, contaminando a base de dados. É o mesmo problema que decorre do que apontam Veselovsky et al. (2023): um aumento muito grande do uso de LLMs pelos chamados *crowd workers*, algo como “trabalhadores de multidão”, que são grandes massas de trabalhadores baratos

contratados para tarefas simples e repetitivas. Uma dessas tarefas, não à toa, é produzir fragmentos textuais para treinamento de IAs.

A dinâmica de trabalho, porém, motiva esses trabalhadores a se utilizarem de IAs para produzir esses textos. No artigo dos autores, são apresentados números baseados numa análise da plataforma de *crowdsourcing* da Amazon, uma das mais utilizadas para terceirização dessas tarefas de forma rápida e com pouco custo, e pouca remuneração também. A essência do *crowdsourcing* é justamente transferir o ônus para um número grande de pessoas, o que dificulta uma verificação concreta do uso de IAs. Os pesquisadores encontraram evidências que algo entre 33% e 46% dos *crowd workers* usaram IA nas suas tarefas textuais (Veselovsky et al., 2023).

Thompson et al. (2024), em um estudo que também tem reverberado bastante, criaram uma metodologia de análise para uma base de dados imensa de frases em dezenas de línguas diferentes, a partir de textos disponíveis publicamente na Internet. Na casa dos bilhões de frases, o estudo focou em investigar instâncias de frases repetidas de tal forma que indicassem serem traduções de uma mesma frase, o que cria um *cluster* formado por frases que são iguais, mas em diferentes línguas.

O que o estudo demonstrou, analisando os parâmetros dessas traduções, é que algo em torno de 57% das frases da base de dados estava em um *cluster* com três ou mais línguas, ou seja, eram frases traduzidas para pelo menos duas outras línguas além da língua de origem. Usando um parâmetro de análise de que quanto maior o número de frases num *cluster*, maior a probabilidade de ser um caso de *Machine Translation*, ou seja, tradução via IAs de texto, o estudo concluiu que, em muitas línguas do estudo, tradução via IA representa a maioria do conteúdo em texto disponível.

Claro que isso não significa que o conteúdo tenha sido originalmente criado por IA, e um texto mal traduzido é muito diferente de um

texto mal escrito. O estudo também aponta alguns vieses, como a impossibilidade de estudos mais individualizados em textos específicos ou que conseguissem identificar diferentes línguas de origens e corrigir distorções. Mas os autores revelam que buscaram algumas frases específicas que apareciam traduzidas em muitas línguas, por vezes dezenas de línguas diferentes, e encontraram coisas como blogs com dicas de comportamento no local de trabalho, relatos de pescaria, e uma série de outros conteúdos que caracterizaram como sendo de baixa complexidade.

Isso reverbera outros estudos, como os de Wu et al. (2023) e Naveed et al. (2024), que apontam uma proliferação muito rápida de textos de baixa complexidade, bastante positivos ou pseudo-informativos, e consumidos de forma muito superficial. Esse fenômeno se encaixa no caráter viral que Parikka (2007) apontava para as redes digitais já no começo do século. Na metáfora do vírus, o autor tenta mostrar três elementos centrais da cultura na Internet: a velocidade de espalhamento e reprodução dos eventos culturais, sua independência do “hospedeiro” (a própria Internet enquanto infraestrutura) e, por fim, sua tendência violenta à disseminação fatal.

É nessa lógica que o conteúdo gerado por IA parece estar se espalhando de forma descontrolada e destrutiva para o hospedeiro, que nesse caso é a própria IA (Mariani, 2023). Agora, com alguma evidência sobre a doença em si, qual seja, o colapso de que falam Schmailov et al., (2024), a metáfora de Parikka (2007) começa a se mostrar de forma mais estruturada, e a virulência da internet de forma mais clara na velocidade e intensidade que caracterizam o aumento desses conteúdos na *web*.

3 Crescente dependência de sistemas de IA generativa

O problema do colapso de sistemas de IA não seria tão grave quando ferramentas como o Chat GPT eram ainda curiosidades usadas aqui e ali. Agora, porém, a curva de dependência de tecnologias de IA generativa parece aumentar tão rápido quanto a do aumento de conteúdo gerado por elas. Esse crescente uso reflete uma tendência global, em que tecnologias emergentes, especialmente as IAs generativas, estão se tornando cada vez mais integradas em diversos setores da sociedade (Cesarino, 2022).

Por um lado, esse fenômeno é impulsionado por avanços tecnológicos rápidos, maior acessibilidade às ferramentas de IA e uma aposta no seu uso para redução de custo com capital humano. No entanto, esse aumento no uso também levanta questões sobre excessiva dependência e implicações éticas para diferentes áreas, como saúde, educação, empresas e negócios, e até mesmo administração pública e uso governamental.

No setor de saúde, por exemplo, a IA está começando a desempenhar um papel significativo no diagnóstico, tratamento e gestão de doenças e pacientes (Lemes e Lemos, 2020). As IAs generativas têm sido estudadas como ferramentas para gerar imagens médicas de alta qualidade, auxiliar na interpretação de exames e até mesmo na descoberta de novos medicamentos, realização de anamnese com pacientes e uma série de outras possibilidades. Ferramentas de IA integradas à bases de dados são usadas para analisar grandes volumes de dados de pacientes e identificar padrões.

No estudo TIC Saúde, realizado pelo NIC (2024c), no módulo de novas tecnologias, foram incluídos indicadores para apontar informações mais detalhadas da aplicação de IAs.

Em 2023 em torno de 3.200 estabelecimentos de saúde usavam ferramentas de IA, a larga maioria na rede privada. Os usos foram: a automatização de fluxos de trabalho (46%), reconhecimento de fala (33%), mineração de texto ou análise de linguagem escrita e falada (32%), reconhecimento e processamento de imagens (21%), e por fim treinamento para predição e análise de dados (16%).

Os estabelecimentos entrevistados apontaram que sua escolha de IA almejou melhorar a segurança digital (45%), apoiar a organização de processos clínicos e administrativos (41%), melhorar a eficiência dos tratamentos (38%), apoiar a gestão de recursos humanos e recrutamento (28%) e auxiliar na dosagem de medicamentos (16%). Esses números revelam que a IA aparece não somente como uma tecnologia importante, por exemplo, para a proteção de dados no setor de saúde, algo esperado, mas também como um potencial elemento central nos atendimentos e consultas.

No campo da educação, as IAs generativas já começaram a ser empregadas na criação de conteúdos educacionais personalizados, exercícios e até mesmo auxiliar na administração de instituições de ensino. Plataformas de ensino online utilizam algoritmos para adaptar o conteúdo às necessidades e ao ritmo de aprendizagem dos alunos, promovendo uma abordagem mais individualizada. Além disso, *chatbots* e assistentes virtuais estão sendo utilizados para responder a perguntas dos alunos e oferecer suporte em tempo real (Tavares, Meira e Amaral, 2020).

O estudo do TIC Educação (2023) revela a consolidação de uma tendência de adoção dessas tecnologias nas escolas, com muitos professores manifestando apoio a uma maior inserção e muitas escolas já começando a adotar essas ferramentas como auxiliares do ensino, e estruturando programas de ensino sobre IA e começando a olhar a possibilidade de tutores automatizados. A crescente dependência dessas

ferramentas pode melhorar a acessibilidade e a eficiência no processo de ensino, mas exige que se considere a qualidade e a equidade do conteúdo gerado, bem como o impacto na interação humana no ambiente educacional.

Já o TIC Empresas (2024a) mostra que empresas brasileiras estão cada vez mais adotando sistemas de IA generativa para otimizar operações, melhorar o atendimento ao cliente e criar novas oportunidades de negócio. Ferramentas de geração automática de texto são utilizadas para criar marketing de conteúdo, gerenciar interações com clientes por meio de chatbots e analisar grandes volumes de dados para tomada de decisões estratégicas. Além disso, a automação de processos e a análise preditiva são apostas das empresas para reduzir custos e aumentar a eficiência.

Esse estudo revelou uma estabilidade no percentual de adoção de tecnologias de IA, que passou de 13%, em 2021, para 14%, em 2023. Esse uso é mais recorrente nas grandes empresas e no setor de informação e comunicação, e a prática mais recorrente é utilização de IA para “automatização de processos de fluxo de trabalho” (73%), com aplicações menos simples, como machine learning (16%) ou geração de linguagem natural (13%), aparecendo de forma mais periférica, o que qualifica uma tendência interessante do ponto de vista da inovação.

No setor governamental, a IA aparece como uma saída para melhorar a eficiência dos serviços públicos e a gestão de recursos, empregada para analisar dados de políticas públicas, tentar prever necessidades sociais e otimizar a alocação de recursos (Toledo e Mendonça, 2023). Além disso, a IA é usada em processos de automação para simplificar burocracias e oferecer serviços mais rápidos e acessíveis à população (Lemes e Lemos, 2020). Na essência, o setor governamental parece também apostar na ferramenta de forma mais ou menos conservadora, ainda que haja uma tendência de aparecer cada vez mais nas interações com cidadãos.

De modo geral, é possível ver uma crescente participação das tecnologias de IA em diversos setores brasileiros, e de origens variadas. Setores como saúde e educação, por exemplo, carregam um peso muito grande em razão de sua natureza mais delicada e sensível. O resultado final, porém, é que há uma maior dependência de uma tecnologia muito recente, e que parece se consolidar de forma muito intensa nos últimos dois ou três anos.

4 O risco social do colapso por retroalimentação

Essa onda de crescente dependência social nas tecnologias de IA pode ser interpretada e conceituada a partir da teoria do risco de Ulrich Beck (2011), que enfatiza como os avanços tecnológicos e científicos, embora tragam benefícios, também introduzem novos riscos que são inerentes à própria modernidade tardia. Para o sociólogo, as sociedades contemporâneas, definidas como sociedades de risco, são caracterizadas por sua capacidade de criar perigos sistêmicos que transcendem fronteiras, têm impactos globais e são frequentemente imprevisíveis e incontroláveis.

Sob essa perspectiva, a difusão das tecnologias de IA pode ser vista como um elemento da sociedade de risco, na medida em que a dependência crescente de sistemas de IA gera uma série de novos riscos, ao mesmo tempo que oferece aparentes soluções promissoras para problemas sociais e econômicos. Esses riscos são tanto tecnológicos quanto sociais, e surgem principalmente porque as tecnologias de IA são altamente complexas, muitas vezes opacas e difíceis de serem completamente compreendidas ou controladas. Essa incerteza está diretamente ligada ao conceito de risco de Beck.

O autor define o conceito de "risco" de maneira a refletir suas preocupações com as sociedades modernas tardias e as mudanças sociais associadas à globalização e à modernidade reflexiva. Para Beck (2011), o conceito de risco deve ser central na análise das dinâmicas contemporâneas e nas formas com que a sociedade lida com as incertezas e perigos associados ao progresso tecnológico e social.

Beck (2011) descreve o risco como um fenômeno social que emerge das condições e consequências do avanço tecnológico e industrial. Mais do que um evento incerto ou perigoso, ele o conceitua como um aspecto intrínseco da modernidade (Beck, 2011), e argumenta que a sociedade moderna tardia está caracterizada por uma crescente preocupação com riscos que são, em grande parte, produtos das próprias inovações e práticas da sociedade industrial.

Para o autor, o risco tem algumas características principais. Uma é seu caráter global, já que os riscos não são mais confinados a áreas locais ou nacionais, mas têm uma dimensão global. Problemas como mudanças climáticas, poluição e crises financeiras são exemplos de riscos que transcendem fronteiras e exigem respostas internacionais. Como Beck (2011) ressalta, os riscos modernos não respeitam fronteiras.

A dependência da IA não é apenas uma questão de uma empresa ou país, mas uma transformação paradigmática global, especialmente com a integração dessas tecnologias em setores cruciais, como saúde, transporte, finanças e governança. A interconectividade digital significa que uma falha em um sistema de IA pode ter consequências de longo alcance. Essa interconectividade transnacional aumenta muito a escala dos riscos, tornando-os mais difíceis de se conter ou controlar.

Outra característica que Beck (2011) vê nos riscos da modernidade tardia é seu aspecto de produção interna. Muitos dos riscos contemporâneos são, em grande parte, uma consequência do próprio avanço tecnológico e das

práticas modernas. A industrialização, a urbanização e a globalização têm gerado novos tipos de riscos que não existiam antes, como os riscos ambientais e tecnológicos.

A IA, como produto da modernidade, ilustra bem esse ponto. Ela é desenvolvida para otimizar processos, automatizar tarefas e gerar valor econômico, mas, ao mesmo tempo, sua implementação cria novos tipos de vulnerabilidades. Por exemplo, algoritmos de IA podem perpetuar vieses discriminatórios ou falhar em decisões críticas em áreas como saúde e segurança, trazendo riscos para os indivíduos e para a sociedade. Essas falhas podem ser imprevisíveis e difíceis de mitigar devido à complexidade dos sistemas e à sua autonomia crescente.

Também podemos pensar na construção de Beck (2011) para a incerteza e a contingência dos riscos contemporâneos. Esses são frequentemente associados a incertezas e à dificuldade de prever ou controlar suas consequências, o que se deve à complexidade dos sistemas tecnológicos e sociais envolvidos, bem como à falta de compreensão completa dos impactos de novas tecnologias e práticas.

Com a IA, esse fenômeno é particularmente relevante, pois os sistemas de IA funcionam em grande parte nos bastidores de muitas atividades cotidianas. A sociedade confia cada vez mais em decisões automatizadas que são tomadas por algoritmos cuja lógica interna é frequentemente opaca (o chamado "problema da caixa-preta"). Isso torna difícil identificar e antecipar quando e onde um risco pode surgir, seja na forma de erro, viés ou uso indevido de dados pessoais.

Outro ponto importante para Beck (2011) é que o enfrentamento dos riscos frequentemente revela desigualdades sociais e injustiças, pois os grupos vulneráveis podem ser desproporcionalmente afetados pelos riscos, enquanto aqueles que têm mais recursos podem se proteger melhor. Beck (2011) aponta que os riscos são muitas vezes distribuídos de maneira

desigual, refletindo e exacerbando as desigualdades existentes. A dependência de sistemas de IA, no geral, cria um grande risco nas pontas do sistema como, por exemplo, em cidadãos usuários dos sistemas de saúde que passam a se integrar com essas tecnologias.

As populações mais vulneráveis, como trabalhadores menos qualificados, vão sofrer as consequências da automação e da substituição de empregos pela IA, enquanto aqueles com maior acesso a capital tecnológico podem se beneficiar desproporcionalmente. Grohmann e Araújo (2021) já apontavam como o "chão de fábrica" das IAs é composto, fundamentalmente, de trabalho precarizado e mal remunerado, que surge a partir de uma relação de exploração baseada em condições econômicas desiguais.

Zajko (2022) aponta que a sociologia, em especial estudos voltados para identificar desigualdades estruturais e estruturantes, oferece uma perspectiva crítica sobre as verdadeiras possibilidades de uma eliminação dos vieses sociais que são reproduzidos nas tecnologias. Sabemos, como mostra Silva (2022), que a vigilância automatizada e os sistemas de IA utilizados por governos e corporações podem impactar mais negativamente grupos marginalizados, aumentando desigualdades sociais e expondo-os a mais riscos, como discriminação algorítmica ou vigilância desproporcional.

Por fim, Beck (2011) levanta uma importante questão sobre mudanças radicais de valores e normas sociais em função da nova dinâmica de riscos. Esses novos problemas desafiam as normas e valores tradicionais e exige uma reavaliação das práticas sociais e políticas. O autor sugere que, à medida que a sociedade toma consciência desses novos riscos, ela se torna reflexiva, ou seja, começa a questionar e reavaliar os próprios fundamentos conceituais da modernidade. É uma mudança de paradigmas sociais que sustentam nosso modelo atual de estrutura coletiva.

5 Considerações Finais

Com o “*framework*” dos riscos da modernidade tardia de Beck (2011) em mente, podemos analisar o problema do colapso das IAs por retroalimentação como um desdobramento dessa nova dinâmica de riscos. Há o surgimento de uma nova tecnologia, com uma série de promessas de avanços, revoluções de custos e eficiência etc., seguido de um espalhamento muito intenso dessa tecnologia de forma financiada e coordenada. Em seguida, essa tecnologia passa a ser integrada em uma série de outros sistemas e serviços sem uma preocupação muito grande com o aumento exponencial da complexidade e dos riscos associados a isso.

Assim, uma vez estabelecido o perigo gerado pelo aumento desproporcional de conteúdo gerado por IA e seu papel nesse colapso por retroalimentação (Schumailov et al., 2024), temos o cenário perfeito de risco conforme descrito por Beck (2011): um perigo globalizado, integrado, que atinge de forma desigual os diferentes atores sociais, e que gera incerteza e dificuldade de contenção, construído pelo próprio avanço da modernidade tardia. Esse risco, ainda segundo Beck, pode ser interpretado como um desdobramento da própria modernidade reflexiva.

Nesse sentido, a atual dinâmica de desenvolvimento de IAs emula o padrão que aparece ao longo do período moderno: há uma enorme capitalização individual do lucro, com uma grande distribuição dos riscos e das consequências. Para além de questões ambientais, trabalhistas e éticas, as tecnologias de IA agora apresentam mais um risco, fruto da proliferação desmedida do seu próprio conteúdo pela internet. A possibilidade de colapso desses sistemas é um perigo concreto para todas as áreas que passarem a depender dos LLMs.

Notas finais

1 Órgão de pesquisa vinculado ao Núcleo de Informação e Coordenação do Ponto BR (NIC.br), que por sua vez é ligado ao Comitê Gestor da Internet do Brasil (CGI.br). Tem a missão de monitorar a adoção das tecnologias de informação e comunicação (TIC) no Brasil. Disponível em: <https://cetic.br/pt/pesquisas/>. Incluir o portal da instituição.

ARTIGO

Gênero e Discriminação Algorítmica: o papel da governança corporativa para mitigar vieses comportamentais em inteligência artificial

Débora Schwartz

debora.schwartz@alumni.usp.br

Mestranda em Direito Comercial
pela Universidade de São Paulo.
Pesquisadora bolsista no Grupo
Direito e Políticas da FDUSP. Integrante
da Comissão Especial de Direito da
Concorrência e Regulação Econômica
(CECORE), da OAB/SP. Advogada.

Gênero e Discriminação Algorítmica: o papel da governança corporativa para mitigar vieses comportamentais em inteligência artificial

Palavras-chave

Discriminação de gênero
Algoritmos
Inteligência artificial
Governança Corporativa

Resumo

O artigo pretende investigar se eventuais malefícios relacionados à discriminação algorítmica de gênero podem ser objeto de disciplina da governança corporativa. A pesquisa apresentará, primeiro, a ideia de que a discriminação algorítmica de gênero se mostra como uma nova roupagem de um problema antigo. Em seguida, argumenta-se que o Direito, e em especial o Direito Empresarial, deve detectar os problemas que essas novas tecnologias acarretam, mediante análise interdisciplinar da interface com a ciência da computação. Por último, levantam-se sugestões práticas de iniciativas, medidas e ferramentas de governança corporativa que podem ser utilizadas para mitigar a discriminação de gênero presente em algoritmos. Ao final, apresenta-se conclusão e breves perspectivas de agendas de pesquisa no campo aqui analisado.

Gender and algorithmic discrimination: the role of corporate governance in mitigating behavioural biases in artificial intelligence

Keywords

Gender discrimination
Algorithms
Artificial Intelligence
Corporate Governance

Abstract

The article aims to investigate whether possible harm related to algorithmic gender discrimination can be the subject of corporate governance discipline. The research will first present the idea that algorithmic gender discrimination shows itself as a new vest for an old problem. It then argues that law, and especially business law, must detect the problems that these new technologies bring, by analysing the interdisciplinary interface with computer science. Finally, practical suggestions are made for initiatives, measures and corporate governance tools that can be used to mitigate the gender discrimination present in algorithms. Finally, there is a conclusion and brief prospects for research agendas in the field analysed.

1 Introdução

No primeiro semestre de 2023, o Instituto InternetLab publicou uma pesquisa que examinava como os aplicativos de música Deezer e Spotify faziam recomendações de artistas aos usuários das plataformas, levando em conta o marcador de gênero na análise¹. Ao se escutar uma música nessas plataformas, o algoritmo automaticamente sugere uma próxima música ao usuário. Com base nesse movimento de recomendações, o resultado da pesquisa indicou que havia diferença significativa na quantidade de artistas mulheres e homens sendo recomendados, no geral, em ambos os aplicativos. Em um caso extremo, chegou-se à marca de uma sequência de 45 músicas recomendadas em que os intérpretes eram, exclusivamente, homens. Por outro lado, quando esse mesmo método foi testado com músicas de intérpretes femininas, o máximo de recomendações seguidas que se chegou de músicas de artistas mulheres foi de 17.

Esse estudo ilustra como as tecnologias de inteligência artificial, movidas por algoritmos, não são tão neutras quanto parecem. Em verdade, sua neutralidade é artificial (Mulholland, 2022). Elas aparentam revelar vieses comportamentais discriminatórios já há tempos identificados na conduta humana. Assim, em alguns casos, a inteligência artificial (IA) poderia reforçar problemas de discriminação de gênero há tanto discutidos, principalmente se considerarmos que suas escolhas virtuais e algorítmicas impactam diretamente na realidade de empresas e consumidores (Van Giffen et al., 2022, p. 93–106).

O conceito de IA é recorrentemente debatido no campo acadêmico e político (Wang, 2019). Segundo estudo recente da União Europeia (2019b, p. 6), pode-se definir IA como:

sistemas de software (e eventualmente também de hardware) concebidos por seres humanos, que, tendo recebido um objetivo complexo, atuam na dimensão física ou digital percebendo o seu ambiente mediante a aquisição de dados, interpretando os dados estruturados ou não estruturados recolhidos, raciocinando sobre o conhecimento ou processando as informações resultantes desses dados e decidindo as melhores ações a adotar para atingir o objetivo estabelecido.

A OCDE, por sua vez, adota uma definição em seu sítio eletrônico que está constantemente se alterando, vez que recebe aportes e sugestões de mudanças em tempo real dada a volatilidade do debate². A proposta mais atualizada da definição de sistemas de IA se aproxima daquela indicada pela UE, mas dá enfoque ao potencial de influência que esses mecanismos possuem, bem como na possibilidade de seus objetivos serem explícitos ou implícitos. Embora não haja um consenso unânime sobre qual deve ser a definição de IA mais adequada, argumenta-se que sua conceitualização é passo importante para se determinar uma moldura regulatória mínima a respeito dessas novas tecnologias (Wang, 2019).

O tema da discriminação algorítmica por sistemas de IA vem tomando força há alguns anos, especialmente por conta de casos em que os setores de recursos humanos de diversas empresas utilizavam IA para contratação de novos funcionários, ou contratavam os serviços de sites de vagas que fizessem essa seleção também por meio de inteligência artificial (Mulholland, 2022). Identificou-se, no entanto, que essas tecnologias acarretavam discriminação de gênero quando da seleção de currículos de possíveis candidatos às vagas disponíveis nas respectivas empresas contratantes.

Entre 2014 e 2015, a empresa Amazon desenvolveu um sistema de inteligência artificial

para ajudar a revisar currículos de candidatos e classificá-los com base em sua adequação para diferentes posições. No entanto, descobriu-se que o sistema tinha uma tendência a discriminar candidatas mulheres. Isso ocorreu porque ele teria sido treinado com currículos enviados à Amazon ao longo de uma década, cujos candidatos eram, em sua maioria, homens. Como resultado, o sistema aprendeu a associar características masculinas às melhores classificações³.

Já em 2018, a plataforma de recrutamento Gupy também apresentou problema similar, em que se questionou a transparência dos critérios de seleção da IA utilizada. Descobriu-se que, apesar de não ser programado para julgar o gênero dos possíveis candidatos ou candidatas, o algoritmo da Gupy não aprovava mulheres em suas seleções⁴.

Em relatório recente denominado *Governing AI for Humanity*, a ONU (2024) reconheceu que sistemas de inteligência artificial podem ser discriminatórios em relação à raça e ao gênero. Dessa forma, classificou tal possibilidade como risco latente a grupos específicos e potencialmente vulneráveis quando do uso de IA.

Tendo em vista esse risco, é corrente o debate sobre a necessidade de regulação das novas tecnologias de inteligência artificial. Em abril de 2021, a Comissão Europeia propôs o primeiro quadro regulatório da União Europeia (UE) para a IA. Ele estabelecia que sistemas de IA que pudessem ser usados em diferentes aplicações fossem analisados e classificados de acordo com o risco que representam aos usuários. Os diferentes níveis de risco significariam maior ou menor regulamentação⁵.

Mais recentemente, em junho de 2023, a UE anunciou que pretende elaborar legislação específica para regular inteligência artificial (também chamada de *Artificial Intelligence Act* ou *AI Act*). A ideia é que o documento estabeleça regras para provedores e usuários, a depender do nível de risco da inteligência artificial. Sistemas de IA que apresentassem risco mínimo teriam

regras e avaliação mais brandas. Já para aqueles sistemas que aparentassem riscos inaceitáveis, como manipulação cognitivo-comportamental de pessoas, scores sociais e identificação biométrica em tempo real, seriam proibidos⁶.

No Brasil, até o presente momento, tem-se notícia da existência de um Projeto de Lei (PL) voltado a esse tema, proposto pelo senador Rodrigo Pacheco (PSD-MG). O PL nº 2.338/2023 estabelece diretrizes para a disponibilização de sistemas de inteligência artificial no Brasil, com especial atenção àqueles impactados pelo seu uso. Além disso, determina critérios para a utilização desses sistemas pelo governo – incluindo sanções para possíveis infrações – e concede ao Poder Executivo a autoridade de designar a entidade responsável por supervisionar e regulamentar o setor⁷.

O debate, portanto, centra-se na perspectiva de atuação do Poder Público face aos eventuais malefícios que as tecnologias envolvendo algoritmos poderiam acarretar, inclusive aqueles relacionados à discriminação de gênero. Mas e as empresas? Que papel devem desempenhar? Devem tomar atitudes para, elas próprias, mitigarem os riscos apresentados por IAs? Como o setor privado se insere nessa questão?

A igualdade de gênero é um direito humano, reconhecido pela ONU⁸, e um direito fundamental, agasalhado pelo art. 5º Constituição de 1988⁹. As empresas, por sua vez, devem possuir papel ativo para com a redução das desigualdades de gênero. Não se trata de preocupação só do Estado ou da sociedade. Desde 2011, os Princípios Orientadores da ONU para Empresas e Direitos Humanos ressaltam essa necessidade de atuação constante e vigilante do setor corporativo (ONU, 2011).

Ademais, com o crescimento do debate envolvendo a pauta ESG (sigla para *Environmental, Social and Governance*), o tema tem se capitalizado e entrado nas estruturas de governança interna das empresas, como será visto adiante. Entende-se, portanto, que as empresas

possuem uma responsabilidade social e devem adquirir posição ativa diante dos problemas que a discriminação algorítmica de gênero pode acarretar. Trata-se de uma questão de governança corporativa.

Por fim, cabe ao ramo do Direito Empresarial endereçar as questões trazidas pelo tema da discriminação algorítmica de gênero. Conforme aponta recente revisão sistemática de literatura conduzida por Sainz et al. (2024, p. 271), o debate aqui exposto tem se concentrado nas áreas do Direito Administrativo e Direitos Humanos. Não há sequer notícia de artigo que trate do tema sob o prisma comercialista. O presente trabalho, portanto, objetiva suprir essa lacuna na literatura e dar início a essa agenda de pesquisa para que mais investigadores se debruçam futuramente sobre o tema.

Para isso, a metodologia adotada foi de levantamento não sistemático da literatura disponível sobre o tema, tendo em vista a escassez de trabalhos já apontada pelo levantamento bibliográfico preliminar no ramo do Direito ora destacado. Após, houve a escolha do referencial teórico para a análise sobre como a discriminação de gênero ocorre ao longo das etapas de aprendizado dos sistemas de IA, sendo elegido o trabalho de Van Giffen et al. (2022), por sua profundidade no tema. Por fim, foram levantados alguns casos de vieses algorítmicos de gênero e ferramentas para mitigá-los com base nas indicações feitas pela própria doutrina selecionada.

Tendo esses pontos como premissa, o presente artigo busca, primeiro, demonstrar que a discriminação algorítmica de gênero se apresenta como uma nova roupagem de um problema antigo. Em seguida, argumenta-se que o Direito Empresarial deve detectar e endereçar os problemas que essas novas tecnologias acarretam, mediante análise interdisciplinar das especificidades trazidas pela interface com a ciência da computação. O último capítulo deste trabalho, por sua vez, busca trazer sugestões

práticas de iniciativas, medidas e ferramentas de governança corporativa para tentar mitigar a discriminação de gênero presente em algoritmos. Ao final, apresenta-se conclusão e breves perspectivas de agendas de pesquisa no campo aqui analisado.

2 Uma velha conhecida em novas roupagens: discriminação algorítmica nas novas tecnologias

Falar em novas tecnologias, por vezes, implica olhar para novos problemas com os quais o Direito deva se preocupar. Embora, à primeira vista, a discriminação algorítmica se apresente como um tema novo no debate jurídico atual, ao se analisar mais detidamente, percebe-se que suas implicações e origens são há muito estudadas pela literatura. Exemplo disso é a falta de representação feminina no campo das ciências exatas, biológicas e humanas, fato que gerou distorções nas pesquisas até então empreendidas por setores que são, majoritariamente, masculinos (Latorre Ruiz & Pérez Sedeño, 2023, p. 66).

Assim, no âmbito algorítmico, essa disparidade social estaria escalada. Segundo revisão de literatura conduzida por Sainz et al. (2024, p. 274), restou pacífico na doutrina que a discriminação por meio de algoritmos reproduz discriminações sociais pré-existentes.

Os crescentes debates sobre discriminação algorítmica serviram, em verdade, para revelar como as desigualdades de gênero, raça etc se mostram e se perpetuam de forma quase “inconsciente” nos vieses comportamentais da sociedade. A particularidade, aqui, evidencia-se no discurso.

Na origem do desenvolvimento de sistemas de IA, difundiu-se a falácia de que tais tecnologias seriam neutras e, portanto, supostamente blindadas dos vieses humanos (Latorre Ruiz & Pérez Sedeño, 2023, p. 65). Vendeu-se a ideia de tecnologia com o intuito de substituição da ação humana pela ação da máquina, justificada pelo fato de a IA ser, em teoria, mais eficiente e livre de subjetividade (Tonucci & Caldeira, 2023). Todavia, isso não se verifica empiricamente (Mulholland, 2022; Lindoso, 2021).

Eroğlu & Karatepe Kaya (2022) destacam, por exemplo, que se cogitou utilizar IAs nos Conselhos de Administração de empresas, para mitigar distorções causadas pela falta de representatividade feminina nesses órgãos corporativos. Resultado: infrutífero. A suposta neutralidade da inteligência artificial não foi comprovada empiricamente e suas decisões, portanto, eram incapazes de substituir a perspectiva feminina – objetivo principal das políticas de cotas.

Esse tema foi mais aprofundado no Brasil sob o recorte da discriminação racial através de algoritmos. Novamente, na revisão de literatura elaborada por Sainz et al (2024, p. 268), a intersecção com o debate racial esteve presente em 69% dos estudos levantados. Há que se destacar também a obra de Silva (2022), a qual procura traçar um panorama sobre como o racismo se mescla e se interpenetra nas redes digitais.

O autor traz o exemplo do *chatbox* Tay, criado em 2016 pela Microsoft (Silva, 2022, p. 67). A inteligência artificial havia sido criada para interagir com o público jovem, usuário da plataforma Twitter, e tinha o objetivo de popularizar os *chatbots*. Após algumas horas de sua implementação, notou-se que a IA passou a reproduzir comentários xenófobos e racistas. Sucede que a base para a coleta e o processamento dos dados que a alimentavam era, precisamente, a interação com humanos através da rede social.

Lobacheva & Kashtanova (2022, p. 99) apontam para um caso nos EUA que possuía

facetas de discriminação ainda mais complexas. Tratava-se de um algoritmo usado em hospitais do país para determinar quais pacientes possivelmente demandariam cuidados médicos adicionais. Em 2019, foi descoberto que esse sistema de IA dava preferência para pacientes brancos em detrimento de pacientes negros. A informação sobre raça não era coletada pela IA, porém era utilizado como variável o valor dispendido pelo paciente em assistência médica. Por uma série de motivos, a parcela negra da população estadunidense alocava menos recursos em saúde do que a parcela branca, o que gerou uma discriminação indireta dos usuários.

Esses exemplos demonstram que os vieses discriminatórios podem influenciar as decisões dos algoritmos, ainda que critérios como raça ou gênero não sejam utilizados para treinamento da IA. Em verdade, os casos narrados demonstram que é justamente na ausência desses dados onde a discriminação tem mais chance de ocorrer indiretamente.

Nesse sentido, entende-se que os aportes feitos sobre a discriminação algorítmica em matéria de raça são aplicáveis e transponíveis para o debate de gênero. Embora o escopo da presente pesquisa não seja a questão racial, este trabalho utiliza esse referencial teórico para embasar a literatura que se debruça sobre discriminação e vieses em sistemas de IA.

Rechaçada a teoria da neutralidade dos sistemas de IA, resta clara a possibilidade de haver discriminação de grupos sociais no processo de aprendizado dessas tecnologias. O conceito de discriminação algorítmica ainda é amplamente debatido na doutrina (Sainz et al, 2024). Todavia, pode ser entendido como a capacidade da IA em replicar ou reforçar “preconceitos existentes na sociedade, ocasionando distinções, preferências ou exclusões capazes de afetar a igualdade de tratamento entre os indivíduos, sobretudo os grupos vulneráveis” (Requião & Costa, 2022, p. 4).

Algumas autoras também argumentam que o próprio conceito de discriminação, por si só, é de difícil determinação (Lobacheva & Kashtanova, 2022). Assim, interpretar a ideia de discriminação algorítmica também se apresenta como desafio. Por essa razão, a literatura tem buscado meios de aprofundar a conceitualização desse termo. Mendes e Mattiuzo (2019), por exemplo, entendem que a discriminação algorítmica pode ocorrer em quatro cenários: (i) por erro estatístico, (ii) por generalização, (iii) por uso de informações sensíveis ou (iv) como limitadora de direitos. Cada uma dessas situações pode ocorrer a depender da etapa de aprendizagem algorítmica que se observa, conforme aprofundado adiante. Pelo exposto, entende-se que a questão da discriminação algorítmica nada mais é do que uma nova roupagem para um problema antigo. No entanto, esse cenário traz novas implicações éticas e políticas, as quais o Direito deve endereçar. Mas, para entender como tratar essas questões, deve-se compreender as particularidades dos sistemas algorítmicos de IA para identificar como a discriminação se mostra nessas novas tecnologias.

3 Em qual momento do desenvolvimento algorítmico ocorre a discriminação?

Para conseguir mitigar o problema da discriminação algorítmica, é necessário compreender, em primeiro lugar, as etapas do chamado *machine learning*. Esse processo consiste, basicamente, na forma pela qual a inteligência artificial recebe, processa e transforma dados em decisões. A União Europeia descreve esse tipo de aprendizagem como técnicas que “permitem que um sistema de IA aprenda a resolver problemas que não podem ser especificados de forma precisa, ou cujo método de resolução não pode ser descrito por regras de raciocínio simbólico” (União Europeia, 2019b, p. 3).

O *machine learning* pode ser dividido em três modalidades: de aprendizagem supervisionada, aprendizagem não supervisionada e aprendizagem por reforço. No primeiro tipo, o algoritmo é treinado a partir de dados fornecidos por um operador humano, de modo que suas decisões serão pautadas com base nos exemplos recebidos. O segundo tipo de aprendizagem decodifica padrões e características comuns aos dados fornecidos e, portanto, não há exemplo correto a ser seguido. Por fim, o terceiro tipo de aprendizagem se desenvolve com base na experiência, pela lógica de tentativa e erro.

A literatura aponta que os problemas de discriminação de gênero são identificados, normalmente, em algoritmos com aprendizagem supervisionada, justamente por haver uma noção de “certo e errado” uma vez que o operador insere os exemplos indicando um viés. Nesse sentido, Van Giffen et al. (2022, p. 96) propõem uma revisão da literatura para compreender como a discriminação pode ocorrer em cada uma das etapas do *machine learning*.

Utiliza-se o método CRISP-DM (*Cross Industry Standard Process for Data Mining*) para se identificar as etapas, são elas: (i) entendimento do negócio; (ii) entendimento dos dados; (iii) preparação dos dados; (iv) modelagem dos dados; (v) avaliação do modelo; e (vi) *deployment* (fase de implantação da IA). Conforme a Figura 1, o momento em que mais há chance de ocorrer discriminação é a etapa de preparação dos dados, pois é nela em que os vieses comportamentais dos próprios seres humanos mais podem incidir quando do desenvolvimento do algoritmo.

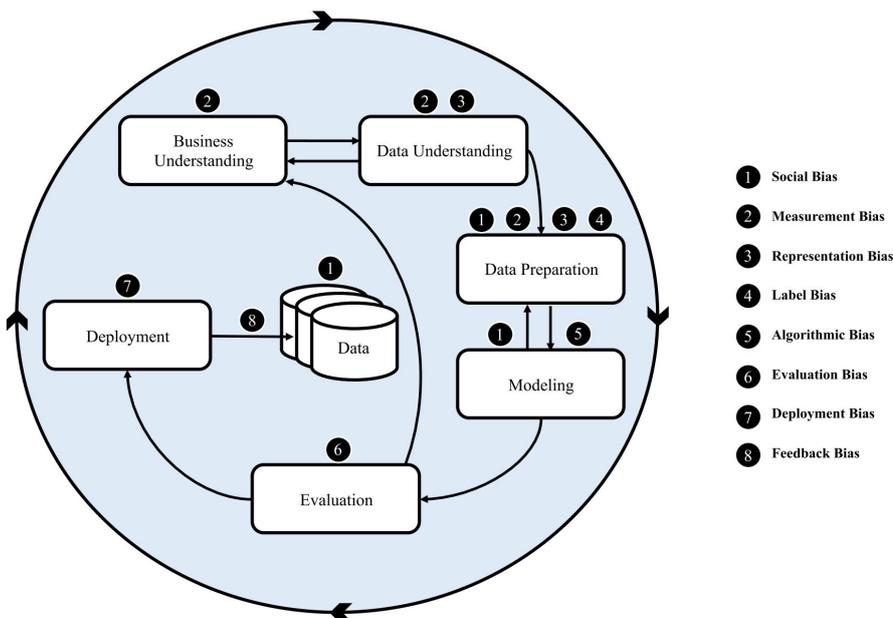


Figura 1: Localização de vieses inconscientes nas etapas do machine learning.

Fonte: Recuperado de "Overcoming the pitfalls and perils of algorithms: A classification of machine learning biases and mitigation methods", de Van Giffen et al., 2022, p. 95.

A partir de então, os autores indicam os tipos de vieses inconscientes passíveis de serem encontrados em cada uma dessas etapas. Descrevem oito tipos de vieses: (i) sociais; (ii) de medição; (iii) de representação; (iv) de classificação; (v) algorítmicos; (vi) de avaliação; (vii) de *deployment*; e (viii) de *feedback*.

Os vieses sociais são reflexo dos próprios vieses inconscientes da amostra/base de dados em que foram coletados os dados (podem ser de uma população, de uma empresa etc). Os de medição consistem na omissão ou na descon sideração de alguma informação da amostra, o que acarretaria discriminá-la (por exemplo, apagar a informação do gênero ou de raça em um algoritmo que preveja a chance de reincidência criminal). Os de representação indicam que a base de dados utilizada para treinar o algoritmo não é adequada ou representativa o suficiente (como ao se pesquisar por imagens de "CEO" no Google e encontrar somente figuras de homens brancos). Já os de classificação indicam erro interpretativo, pois há categorização de padrões que não são socialmente unânimes (é o caso de se pesquisar por imagens de "casamento" no Google e deparar-se apenas com o modelo de festa ocidental, excluindo-se os de tradição indiana, por exemplo).

Quanto aos vieses algorítmicos, trata-se de considerações técnicas inapropriadas do próprio sistema, durante a fase de modelagem da IA. Vieses de avaliação ocorrem quando uma amostra não representativa é usada para testar e reavaliar o algoritmo (o que acarreta resultado similar ao já descrito nos vieses de representação). Os de *deployment* ocorrem, especificamente, no momento de implantação da tecnologia, quando uma IA é utilizada em um contexto diferente daquele que ela fora criada para atuar. Por fim, os vieses de *feedback* incidem ao final do *machine learning*, quando uma outra IA é treinada usando dados fornecidos por um algoritmo eivado de vieses inconsciente. Por isso a literatura chama

esse fenômeno de *feedback loop* (O’Neil, 2018), já que, em determinado ponto, não é mais possível descobrir em que momento da cadeia algorítmica foi gerada a discriminação (Van Giffen et al., 2022, p. 98).

Essa ampla gama de vieses inconscientes passíveis de acarretar discriminação fez com que a União Europeia, em 2019, desenvolvesse um Guia de Orientações Éticas para uma IA de confiança (no título em inglês, *Ethics Guidelines for Trustworthy AI*), que estabelece sete requisitos para que se tenha um grau razoável de confiabilidade em um algoritmo: (i) ação e supervisão humanas; (ii) solidez técnica e segurança; (iii) privacidade e governação dos dados; (iv) transparência; (v) diversidade, não discriminação e equidade; (vi) bem-estar ambiental e societal; (vii) responsabilização (União Europeia, 2019a, p. 6). Especificamente para o presente artigo, destacam-se os requisitos (iv) e (v).

Sobre diversidade, não discriminação e equidade, o Guia pontua que a utilização de sistemas de IA deve vir acompanhada (i) da Prevenção de enviesamentos injustos, (ii) da noção de acessibilidade e conceção universal e (iii) da participação das partes interessadas. São diretrizes genéricas que resumem as preocupações de Van Giffen et al. (2022) descritas acima. O problema da discriminação algorítmica é, portanto, completamente reconhecido pela União Europeia.

A questão da transparência dos algoritmos – requisito (iv) – se dá pelo fato de que, com o avanço dessa tecnologia, os algoritmos têm ficado cada vez mais complexos e sofisticados. Sua documentação, portanto, é dificultada, bem como o seu acesso por terceiros. Tal situação fez com que as decisões tomadas por algoritmos ficassem opacas, não apenas para a sociedade e as autoridades, mas também para os próprios desenvolvedores de algoritmos (Latorre Ruiz & Pérez Sedeño, 2023, p. 71).

Ademais, é do interesse das empresas que esses algoritmos permaneçam na forma de

segredos de negócio, pois consistem em vantagem competitiva perante seus concorrentes. Assim, ainda que existentes, tais documentos não são comumente divulgados, sob a justificativa de se evitar a ocorrência de “efeito carona” (no jargão em inglês, *free-riding*). Pasquale (2015) explica que esse âmbito de opacidade pode ser chamado de sigilo legal, isto é, quando o conteúdo do algoritmo não pode ser revelado por determinações normativas.

O Guia da UE, por sua vez, propõe três pilares para se endereçar o problema da transparência: (i) rastreabilidade; (ii) explicabilidade e (iii) comunicação. No primeiro, há grande foco na necessidade de documentação dos algoritmos. A explicabilidade técnica, por sua vez, exige que as decisões tomadas por um sistema de IA possam ser compreendidas e rastreadas por seres humanos. Por fim, o pilar da comunicação estipula que não se deve entender os sistemas de IA como se humanos fossem, de forma que as pessoas têm direito de serem informadas de que estão interagindo com uma IA.

O Guia, contudo, não apresentou iniciativas práticas para se implementar os requisitos de uma IA de confiança (Hickman & Petrin, 2021). A tentativa da doutrina, desde então, foi buscar estabelecer mecanismos e instrumentos que pudessem endereçar os problemas discriminatórios ocasionados pelo uso de inteligência artificial. Acredita-se que a governança corporativa, por sua vez, possa ter um papel relevante na mitigação dessas questões.

4 Mecanismos de governança corporativa para mitigar a discriminação algorítmica de gênero

A partir do cenário até aqui narrado, doutrina, empresas, organizações internacionais e governos vêm tentando buscar alternativas para amenizar ou solucionar a discriminação identificada em tecnologias de inteligência artificial. Um exemplo recente é o kit de ferramentas da IBM chamado *AI Fairness 380*, que procura facilitar adequação dos algoritmos de uso industrial aos princípios éticos destacados nesta pesquisa, bem como para auxiliar pesquisadores a compartilharem e avaliarem algoritmos entre si (Bellamy et al., 2021). Assim como ela, outras ferramentas foram desenvolvidas por grandes empresas de tecnologias para tentar amenizar os problemas aqui discutidos, como se verá adiante.

Tais medidas nada mais são do que práticas de governança corporativa elaboradas pelo setor privado para endereçar o problema dos vieses em IA, também chamada nesse contexto de governança algorítmica (Mendes & Mattiuzzo, 2019). A ideia principal é que, ao projetar sistemas algorítmicos, devem ser seguidos os princípios de conscientização, explicabilidade, precisão, auditabilidade, justiça (*fairness*) e fiscalização e reparação de eventuais danos causados pela IA, sob a perspectiva do chamado *accountability* (Mendes & Mattiuzzo, 2019, p. 56).

Preconiza-se que as empresas devem ser responsabilizadas pelo fato de que as pessoas serão afetadas pelo processo decisório dos sistemas de IA. Como forma de mitigar eventuais efeitos negativos, recorre-se também à noção de transparência, justamente como forma de aclarar as *black boxes* mencionadas por Pasquale (2015), que consistiriam no fenômeno da opacidade algorítmica.

Todavia, a mera listagem de tais princípios não é suficiente para garantir seu efetivo *enforcement*. É necessário pensar medidas mais concretas e aplicáveis à prática do setor privado.

Nessa linha, acerca das possíveis discriminações algorítmicas descritas no capítulo anterior, Van Giffen et al. (2022) propõem um modelo com diversas sugestões para se mitigar os efeitos que vieses comportamentais podem causar no âmbito do *machine learning*, aqui entendidas como práticas de governança corporativa. Destacam-se três delas: (i) documentação e transparência do algoritmo; (ii) supervisão humana obrigatória e (iii) times diversos e multidisciplinares.

Cumprido desde logo destacar que tais sugestões em muito se aproximam dos princípios destacados pelo Guia da União Europeia para se atingir uma “IA confiável”. O elemento adicional aqui consiste no aspecto prático dessas possíveis soluções. Frisa-se, também, que essas alternativas para se lidar com a discriminação algorítmica devem incidir principalmente na etapa de preparação de dados do *machine learning* – etapa (iii) descrita no capítulo anterior –, onde mais se encontra o risco de incidência de vieses.

Quanto à solução de documentação e transparência do algoritmo, novas ferramentas vêm sendo desenvolvidas e testadas pelos cientistas da computação, cabendo destaque para duas: *a Model Cards*, da Google, e *AI System Cards*, da Meta. A tecnologia de *Model Cards* busca simplificar e automatizar a geração de documentos de *machine learning* que oferecem contexto e transparência para o desenvolvimento e o desempenho de um modelo de algoritmo¹⁰. Ela seria integrada ao processo de *machine learning*, de forma a possibilitar que os criadores de um sistema de IA compartilhem as métricas e os metadados do modelo com demais pesquisadores, desenvolvedores, informantes etc.

A *AI System Cards*, por sua vez, tem uma lógica semelhante. A diferença residiria no fato de que

esse modelo foca, principalmente, em informações de nível de sistema. Um System Card forneceria uma visão geral de vários modelos de aprendizado de máquina que compõem um sistema de aprendizado de máquina, bem como detalhes sobre esses componentes e um guia passo-a-passo com um exemplo de entrada¹¹.

Tais ferramentas podem ser úteis não apenas na fase de preparação dos dados, mas também na de entendimento desses dados e de implementação da IA – respectivamente etapas (ii) e (vii) mencionadas por Van Giffen et al. (2022). Uma vez ampliada a transparência no *machine learning*, é possível haver maior treinamento das equipes que o alimentam e que dele se utilizam para a criação de novos sistemas de IA. Assim, o risco de haver transmissão de vieses comportamentais entre diferentes IAs fica reduzido.

Contudo, ressalta-se que as empresas podem encontrar dificuldades na implementação dessa prática de transparência, tendo em vista a vantagem competitiva que reveste a opacidade algorítmica. Uma vez agasalhado pelo segredo de negócio, o algoritmo se transforma em um ativo a ser protegido do chamado “efeito corona”. Daí a necessidade de medidas regulatórias, capitaneadas pelo Poder Público, para que a disciplina do segredo empresarial não seja um empecilho para se atingir um grau adequado de confiabilidade e transparência algorítmicas.

Quanto à segunda sugestão, sobre a possibilidade de haver supervisão humana obrigatória dos sistemas de IA, tem sido frequente o debate sobre a criação de equipes de *Responsible AI* ou de *Explainable AI*, mais conhecidas pela sigla XAI (Arrieta et al., 2020). A fim de se garantir que haja redução nas discriminações que seus respectivos algoritmos poderiam causar, propõe-se a existência de equipes responsáveis por aplicar os princípios da transparência e da explicabilidade nos sistemas de IA utilizados no bojo de suas respectivas empresas.

Destaca-se que a supervisão humana também pode ser de extrema valia na etapa (v) do

machine learning, de avaliação do modelo, sugerida por Van Giffen et al. (2022). Isso porque, ao passar pelo crivo humano, há maiores chances de que eventual incorporação de vieses pela IA seja identificada e corrigida. Todavia, a supervisão acrítica não basta, razão pela qual se destaca a importância da criação de equipes diversas e inclusivas para realizar essa avaliação da modelagem algorítmica.

Chega-se, dessa forma, à sugestão de estabelecimento de times multidisciplinares e diversos para análise e supervisão da IA, mencionados por Van Giffen et al. (2022) e outros expoentes da literatura (Tonucci & Caldeira, 2023, p. 336). No bojo dessas equipes, o cientista da computação estaria imbuído de coletar evidências suficientes para se testar o modelo e garantir a correta documentação do algoritmo. Ao operador do Direito, caberia examinar se os resultados e o comportamento do algoritmo implicam discriminação ou ilegalidade.

Ademais, o problema da sub-representação feminina no bojo das ciências exatas, especialmente na ciência da computação e na ciência de dados, é fenômeno latente. Como apontado por Tonucci & Caldeira (2023, p. 330), estudos realizados em 2020 indicaram que 15% das cientistas de dados do mundo eram mulheres e, das profissionais que ocupavam funções de análise de dados, eram apenas 26%.

A respeito dos vieses inconscientes envolvidos na contratação de profissionais, Pisanelli (2022) desenvolveu interessante trabalho acerca da possibilidade de reduzir as disparidades de gênero usando, precisamente, inteligência artificial.

Como largamente demonstrado no início do artigo, a grande parte das notícias sobre o tema apontam para casos de aumento da discriminação de gênero, e não de redução. Todavia, o estudo empírico ora mencionado aponta para outras modelagens possíveis do *machine learning*.

Em suma, a autora explica que os problemas principais da discriminação de IAs como da

Gupy, da Amazon etc residiam no fato de que essas empresas faziam o *machine learning* com base em seus próprios funcionários, adotando uma estratégia de IA preditiva (*predictive AI*). Todavia, caso fosse adotada uma estratégia de triagem através da inteligência artificial (*screening AI*), isto é, que o sistema tivesse como base os próprios currículos e escolhesse os melhores a partir tão somente da amostra dos candidatos, a discriminação de gênero era mitigada.

O trabalho de Pisanelli (2022), portanto, indica que a adoção de modelos de IA baseados em triagem pode ser mais benéfica quando da contratação de profissionais, inclusive aumentando a probabilidade de haver contratadas mulheres.

Todavia, ainda que se transponha a barreira da contratação, compor equipes diversas e multidisciplinares não é uma tarefa simples, tampouco óbvia. Isso porque a própria integração dessas equipes é um desafio enfrentado pelas empresas quando da implementação dessa medida de governança.

Uma política comumente vista nesse âmbito é a criação de cotas para grupos minorizados¹² em processos seletivos de contratação ou em altos cargos de gestão, de modo a permitir que eles acessem espaços de poder. Esse é, definitivamente, um primeiro passo importante. Todavia, essa medida deve vir acompanhada de mecanismo de ascensão e permanência das mulheres em tais ambientes majoritariamente masculinos.

Abreu (2024), ao estudar as políticas de cotas de gênero nas companhias brasileiras, identifica que ações afirmativas de reservas de vagas a mulheres em altos cargos de gestão não são suficientes para endereçar o problema da sub-representação feminina nesses ambientes¹³. Primeiro, porque a grande maioria das empresas possui, no máximo, 30% de cargos de diretoria ocupados por mulheres, o que, segundo a autora, seria insuficiente para se falar em efetiva representação. Segundo, porque se identificou grande número de diretoras que atuam em

vários conselhos de administração – ou seja, acumulam cargos como membros independentes, fenômeno também conhecido como *trophy directors* ou *golden skirts*. São mulheres que possuem dupla ou tripla jornada para garantir a representação feminina nos cargos decisórios das empresas. Tendo em vista o elevado comprometimento exigido nesses cargos, além de poder acarretar sobrecarga, a autora aponta que esse poderia ser um indício de baixa participação destas conselheiras nas companhias em que atuam (Abreu, 2024). Não basta, pois, criar condições de acesso.

Assim, o que se debate hoje é que, para favorecer o surgimento de ambientes diversos, é preciso, também, criar políticas de inclusão e permanência de grupos minorizados. O tema já fora muito debatido, por exemplo, no âmbito do ingresso desses grupos às universidades públicas. As primeiras políticas de cotas no ambiente universitário foram pioneiras ao possibilitar o acesso de grupos minorizados aos espaços acadêmicos, antes destinados a uma pequeníssima parcela da sociedade. Com o tempo, percebeu-se que tais políticas não eram suficientes. Deveriam vir acompanhadas, também, de medidas que possibilitassem a permanência e a inclusão desses grupos na universidade. Assim surgiram as diversas iniciativas que vemos hoje nas instituições de ensino superior do país¹⁴.

A mesma lógica pode ser aplicada para se debater a questão da sub-representação feminina no ambiente de trabalho. Se queremos combater o problema da discriminação algorítmica com a inserção de mulheres nos mercados de tecnologia, especialmente na computação, não basta garantir o simples acesso a esses espaços (Requião & Costa, 2022). É necessário traçar estratégias de ascensão e de permanência das mulheres nesse ambiente.

A ONU Mulheres, a WE Impact e a Consultoria GEMA elaboraram uma série de iniciativas práticas possíveis de serem utilizadas

por startups que buscam diversificar seu quadro de funcionários. O relatório “Princípios de Empoderamento das Mulheres para *Startups*” (GEMA, 2021) sugere, por exemplo, (i) realização de mentorias entre as mulheres da empresa; (ii) treinamento de novos colaboradores e novas colaboradoras voltado a questões de gênero; (iii) criação de espaços inclusivos *networking* para pessoas que estão em posições de cuidado, como mães e aquelas que são responsáveis por idosos; dentre outras políticas e programas de governança que fomentem o desenvolvimento profissional do quadro feminino em empresas de tecnologia.

Ademais, pode-se dizer que uma empresa comprometida em mitigar a discriminação algorítmica (seja de gênero, raça, classe social etc) é também uma empresa que se pretende alinhada aos princípios ESG, tão em voga nas atuais discussões de Direito Empresarial (Eroğlu & Karatepe Kaya, 2022, p. 548).

No Brasil, este debate está, em certa medida, inserido no referido PL nº 2338/2023, em trâmite perante o Senado. Os arts. 19 a 26 de seu texto original estabelecem medidas gerais de governança dos sistemas de inteligência artificial. Por sua vez, o art. 20, IV, estabelece que:

Art. 20. Além das medidas indicadas no art. 19, os agentes de inteligência artificial que forneçam ou operem sistemas de alto risco adotarão as seguintes medidas de governança e processos internos:

[...]

IV – medidas de gestão de dados para mitigar e prevenir vieses discriminatórios, incluindo:

a) avaliação dos dados com medidas apropriadas de controle de vieses cognitivos humanos que possam afetar

a coleta e organização dos dados e para evitar a geração de vieses por problemas na classificação, falhas ou falta de informação em relação a grupos afetados, falta de cobertura ou distorções em representatividade, conforme a aplicação pretendida, bem como medidas corretivas para evitar a incorporação de vieses sociais estruturais que possam ser perpetuados e ampliados pela tecnologia; e

b) composição de equipe inclusiva responsável pela concepção e desenvolvimento do sistema, orientada pela busca da diversidade.

Como se vê, o trecho acima incorpora algumas das sugestões esboçadas na literatura e largamente debatidas neste artigo. Todavia, embora o PL se esforce para prever a necessidade de governança corporativa quanto aos vieses algoritmos, não avança muito quanto à forma que essas medidas podem tomar.

A alínea ‘a’ destaca a necessidade de avaliação dos dados objeto do *machine learning*. Todavia, não detalha como poderia ser feita essa avaliação, se ela necessariamente deve passar pelo crivo humano e nem como isso se daria dentro das empresas e demais entidades do mercado. Já a alínea ‘b’ menciona a implementação de equipes diversas, mas não ressalta a questão da multidisciplinariedade para se enfrentar a complexidade do problema, por exemplo.

Dessa forma, entende-se que o debate sobre vieses discriminatórios de sistemas de IA ainda caminha a passos lentos, principalmente quando se trata de medidas a serem adotadas pelo setor privado para mitigar tal situação.

De fato, essa ausência de robustez regulatória também é um obstáculo enfrentado pelo mercado quando da implementação dessas práticas de governança. Uma vez que não há clareza sequer na definição de IA, tampouco nas formas

de tratá-la, há um vácuo normativo acerca de quais medidas mais concretas de governança corporativa devem ser incentivadas e/ou seguidas pelas empresas.

Tendo em vista essa lacuna, a grande maioria das propostas apresentadas neste capítulo perpassam o aprofundamento de medidas voltadas à governança corporativa. A urgência em se regular tecnologias de inteligência artificial não é um debate que cabe apenas ao âmbito do Direito Público, mas também ao Direito Privado. A discriminação algorítmica de gênero é um problema interseccional, multisetorial e, portanto, deve ser objeto de estudo de todos os campos do saber jurídico.

5 Conclusão

Apesar da falácia de neutralidade das tecnologias de inteligência artificial, a crença de que as IAs são eficientes e livres de subjetividade foi desafiada e desmistificada. Não há, pois, que se falar em tecnologia neutra ou livre de vieses comportamentais. Existem diversas variáveis invisíveis quando falamos da tomada de decisões por algoritmos. Uma delas é, precisamente, a questão de gênero.

Isso traz implicações éticas e políticas que o Direito deve enfrentar, exigindo uma compreensão profunda e interdisciplinar das particularidades dos sistemas algorítmicos de IA para lidar com a discriminação nesse contexto.

A literatura aponta que, em todas as etapas do chamado *machine learning*, pode haver discriminação algorítmica, com especial destaque para a fase de preparação dos dados. Isso porque é nela que os vieses comportamentais dos próprios seres humanos mais podem incidir quando do desenvolvimento do algoritmo.

Tais preocupações fizeram com que União Europeia, por exemplo, propusesse um Guia de diretrizes éticas para se atingir um maior

nível de confiabilidade na inteligência artificial, com especial enfoque para as noções de transparência e não discriminação. Embora referido Guia não tenha apresentado iniciativas práticas, a doutrina buscou mecanismos para abordar os problemas discriminatórios causados pela IA. O Direito Empresarial, por sua vez, teria o papel de propor abordagens regulatórias capazes de mitigar esse problema no âmbito do setor privado

Nesse contexto, várias instituições, empresas, organizações internacionais e governos têm buscado alternativas para atenuar ou resolver a discriminação identificada em tecnologias de inteligência artificial. Por exemplo, o desenvolvimento de ferramentas de transparência por grandes empresas de tecnologia – como IBM, Google e Meta –, na tentativa de abordar os problemas ora discutidos.

Como sugestões práticas, propõe-se um modelo que oferece alternativas para mitigar os efeitos dos vieses comportamentais no *machine learning*, em especial a documentação e transparência do algoritmo, a supervisão humana obrigatória e a criação de equipes multidisciplinares.

Especialmente quanto à última sugestão, o debate se intensifica. A composição dessas equipes não é simples e exige a criação de políticas de inclusão e permanência de grupos minorizados. Isso pode ser observado no debate sobre a sub-representação feminina no campo das ciências exatas e no ambiente de trabalho, onde estratégias de ascensão e inclusão são essenciais para combater a discriminação.

Tais estratégias de mitigação do problema, no entanto, são pouco reguladas ou disseminadas na atualidade. Para que haja um combate mais robusto dos vieses algorítmicos de gênero, faz-se necessário não apenas um papel ativo e integrado do setor privado, mas também dos próprios ramos do saber jurídico.

Referências bibliográficas

- Abreu, T. (2024). A diversidade de gênero na alta administração: tokenismo, conexões pessoais e diretoras de enfeite. *Revista de Direito Mercantil industrial, econômico e financeiro – RDM*, 180/181, 161–212.
- Arrieta, A., B. et al. (2020). Explainable Artificial Intelligence (XAI): *Concepts, taxonomies, opportunities and challenges toward responsible AI*. *Information Fusion*, 58, 82–115.
- Bellamy, R., K., E. et al. (2021). AI Fairness 360: An Extensible Toolkit for Detecting and Mitigating Algorithmic Bias. *IBM Journal of Research and Development*.
- Eroğlu, M., & Karatepe Kaya, M. (2022). Impact of Artificial Intelligence on Corporate Board Diversity Policies and Regulations. *European Business Organization Law Review*, 23(3), 541–572. <https://doi.org/10.1007/s40804-022-00251-5>
- Gema, ONU Mulheres & We Impact (2021). *Princípios de Empoderamento das Mulheres para Startups*. Recuperado de: http://www.onumulheres.org.br/wp-content/uploads/2021/09/WEPs_Startups_Toolkit_POR_BRA_2021_WinWin.pdf.
- Hickman, E., & Petrin, M. (2021). Trustworthy AI and Corporate Governance: The EU's Ethics Guidelines for Trustworthy Artificial Intelligence from a Company Law Perspective. *European Business Organization Law Review*, 22(4), 593–625. <https://doi.org/10.1007/s40804-021-00224-0>
- Latorre Ruiz, E., & Pérez Sedeño, E. (2023). Gender Bias in Artificial Intelligence. In J. Vallverdú (Org.), *Gender in AI and Robotics* (Vol. 235, p. 61–75). Springer International Publishing. https://doi.org/10.1007/978-3-031-21606-0_4
- Lobacheva, A., & Kashtanova, E. (2022). Social Discrimination in the Epoch of Artificial Intelligence. *Wisdom*, 2(1), 97–105. <https://doi.org/10.24234/wisdom.v2i1.767>
- Lindoso, M., C., B. (2021). *Discriminação de gênero no Tratamento Automatizado de Dados Pessoais: como a automatização incorpora vieses de gênero e perpetua a discriminação de mulheres*. Rio de Janeiro: Processo.
- Mulholland, C. (2022). Inteligência artificial e discriminação de gênero. In Schreiber, A.; Martins, G., M., & Carpena, H. (eds.) *Direitos Fundamentais e Sociedade Tecnológica* (pp. 169-182). São Paulo: Foco.
- O'NEIL, C. (2018). *Weapons of Math Destruction: How Big Data increases Inequality and Threatens Democracy*. Nova York: Crown.
- ONU. (2024). *Governing AI for Humanity: Final Report*. Recuperado de: https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf.
- ONU. (2011). *Guiding Principles on Business and Human Rights*. Disponível em: https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinessshr_en.pdf.
- Pasquale, F. (2015). *The Black Box Society: The secret Algorithms that control Money and Information*. Cambridge, Massachusetts, London, England: Harvard University Press.
- Pisanelli, E. (2022). *A new turning point for women: artificial intelligence as a tool for reducing gender discrimination in hiring*. Presentation at the European University Institute. Recuperado de: https://congress-files.s3.amazonaws.com/2022-08/Pisanelli_paper.pdf.
- Requião, M., & Costa, D., C. (2022) Discriminação algorítmica: ações afirmativas como estratégia de combate. *Civilistica*, 11(3), 1–24.

- Sainz, N., Gabardo, E., & Ongaratto, N. (2024). Discriminação algorítmica no Brasil: uma análise da pesquisa jurídica e suas perspectivas para compreensão do fenômeno. *Revista Direito Público*, 21(110), 258-289. doi: <https://doi.org/10.11117/rdp.v21i110.7295>
- Santos, R. (2020). *Maioria minorizada*. Rio de Janeiro: Telha.
- Silva, T. (2022). *Racismo algorítmico: Inteligência artificial e discriminação nas redes digitais*. São Paulo: Edições Sesc.
- Tonucci, D., C., G., & Caldeira, M., M. (2023). Feminismo de dados: uma nova perspectiva para a ciência e os vieses "inconscientes" de gênero. In Barbosa, B., Tresca, L., & Lauschner, T. (eds.) *3a Coletânea de Artigos—TIC, Governança da Internet, Gênero, Raça e Diversidade—Tendências e Desafios*. Núcleo de Informação e Coordenação do Ponto BR.
- União Europeia. (2019a). *Ethics guidelines for trustworthy AI*. Recuperado de: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.
- União Europeia. (2019b). *Uma definição de IA: principais capacidades e disciplinas científicas*. Recuperado de: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.
- Van Giffen, B., Herhausen, D., & Fahse, T. (2022). Overcoming the pitfalls and perils of algorithms: A classification of machine learning biases and mitigation methods. *Journal of Business Research*, 144, 93–106. <https://doi.org/10.1016/j.jbusres.2022.01.076>
- Wang, P. (2019). On Defining Artificial Intelligence. *Journal of Artificial General Intelligence*, 10(2), 1-37. doi: <http://dx.doi.org/10.2478/jagi-2019-0002>.

Notas finais

1 Projeto Algo_Ritmos. Disponível em: <https://algoritmos.internetlab.org.br>.

2 Updates to the OECD's definition of an AI system explained. Disponível em: <https://oecd.ai/en/wonk/ai-system-definition-update>.

3 Disponível em: <https://epocanegocios.globo.com/Empresa/noticia/2018/10/amazon-desiste-de-ferramenta-secreta-de-recrutamento-que-mostrou-vies-contramulheres.html>.

4 Disponível em: <https://www.intercept.com.br/2022/11/24/como-plataformas-de-inteligencia-artificial-podem-discriminar-mulheres-idosos-e-faculdades-populares-em-processos-seletivos/>.

5 Disponível em: <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>.

6 Disponível em: <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>.

6 Disponível em: <https://www12.senado.leg.br/noticias/materias/2023/05/12/senado-analisa-projeto-que-regulamenta-a-inteligencia-artificial>.

7 Disponível em: <https://www12.senado.leg.br/noticias/materias/2023/05/12/senado-analisa-projeto-que-regulamenta-a-inteligencia-artificial>.

8 A igualdade de gênero passou a fazer parte do direito internacional dos direitos humanos pela Declaração Universal dos Direitos Humanos, que foi adotada pela Assembleia Geral em 10 de dezembro de 1948. Disponível em: <http://www.onumulheres.org.br/noticias/igualdade-de-genero-e-assembleia-geral-da-onu-fatos-e-historia-a-saber/>.

9 “Art. 5º Todos são iguais perante a lei, sem distinção de qualquer natureza, garantindo-se aos brasileiros e aos estrangeiros residentes no País a inviolabilidade do direito à vida, à liberdade, à igualdade, à segurança e à propriedade, nos termos seguintes: I - homens e mulheres são iguais em direitos e obrigações, nos termos desta Constituição [...]”.

10 Disponível em: <https://modelcards.withgoogle.com/about>.

11 Disponível em: <https://modelcards.withgoogle.com/about>.

12 Opta-se pela expressão “grupo minorizado” por entender que ela reflete melhor a realidade e o processo que as chamadas “minorias sociais” enfrentam cotidianamente. Conforme Santos (2020), o termo passou a ser utilizado a partir dos anos 90 e deriva da ideia de que determinados segmentos sociais, independentemente da quantidade de indivíduos, têm pouca representação social, econômica e política. Estão, portanto, à margem dos interesses sociais.

13 O regulamento de emissores da B3, por exemplo, contemplou medidas relacionadas a temas Ambientais, Sociais e de Governança Corporativa (conhecidos como ESG), voltadas para a temática da representação feminina nos conselhos de administração ou nas diretorias estatutárias de companhias abertas. Disponível em: https://www.b3.com.br/data/files/3B/31/0A/CF/394798101DBF7498AC094EA8/Regulamento%20de%20Emissores%20_20.07.2023_.pdf#page30.

14 Exemplo disso é o Programa de Apoio à Permanência e Formação Estudantil (PAPFE), na USP, vide: <https://jornal.usp.br/radio-usp/programa-estudantil-na-usp-pretende-assegurar-maior-inclusao-e-permanencia-de-alunos/>.

ARTIGO

Opacidade algorítmica, segregação epistêmica e alterações no conceito de trabalho

Catarina Spagnol

catarinaspagna@yahoo.com.br

Website: <https://catarinaspagnol.wordpress.com/>

Formada em Comunicação Social pela Universidade Metodista de Piracicaba (Brasil) e em Filosofia pela UNED (Espanha), onde também concluiu o Mestrado em Filosofia Teórica e Prática, com especialização em Lógica, História e Filosofia da Ciência. Seus interesses de pesquisa incluem filosofia da linguagem, relação linguagem-pensamento, teorias do significado e mudanças conceituais em contextos sociais e políticos, com foco em contextos digitais de argumentação coletiva.

Opacidade algorítmica, segregação epistêmica e alterações no conceito de trabalho

Palavras-chave

Tecnologia

Opacidade algorítmica

Inteligência Artificial

Segregação epistêmica

Mudanças Conceituais

Resumo

Como a opacidade algorítmica está moldando a prática social e transformando nossa concepção sobre o trabalho? Neste artigo reflito sobre o impacto dos algoritmos de aprendizado que, ao operarem de maneira opaca, têm efeitos significativos sobre o ambiente laboral. Analiso como a falta de transparência nas decisões automatizadas pode comprometer os direitos dos trabalhadores e alterar as normas e práticas laborais. Evidencio que a opacidade dos algoritmos frequentemente resulta em decisões arbitrárias, marginalizando trabalhadores e promovendo injustiças no local de trabalho. Apresento uma perspectiva que parte da ideia que as tecnologias disruptivas e opacas modificam a rede conceitual do trabalho, gerando insegurança e desigualdade. Para concluir, afirmo que é crucial exigir maior transparência nos algoritmos e responsabilizar as empresas para evitar a erosão dos direitos trabalhistas. Enfatizo a necessidade de um engajamento crítico para garantir que a evolução tecnológica não perpetue injustiças e incertezas no ambiente laboral.

Algorithmic opacity, epistemic segregation and changes in the concept of work

Keywords

Technology

Algorithm Opacity

Artificial Intelligence

Epistemic Segregation

Conceptual Change

Abstract

How is algorithmic opacity shaping social practices and transforming our conception of work? In this article, I reflect on the impact of learning algorithms that, by operating opaquely, have significant effects on the workplace. I analyze how the lack of transparency in automated decisions can compromise workers' rights and alter labor norms and practices. I highlight that algorithmic opacity often results in arbitrary decisions, marginalizing workers and promoting injustices in the workplace. I present a perspective that disruptive and opaque technologies modify the conceptual framework of work, generating insecurity and inequality. In conclusion, I assert that it is crucial to demand greater transparency in algorithms and hold companies accountable to prevent the erosion of labor rights. I emphasize the need for critical engagement to ensure that technological evolution does not perpetuate injustices and uncertainties in the workplace.

1. Introdução

A tentativa de prever o futuro da humanidade diante do avanço tecnológico é uma pretensão arriscada e, por isso, não faz parte do objetivo do presente trabalho. Em vez disso, o objetivo é propor uma análise sobre os possíveis efeitos da mediação algorítmica na prática social do trabalho. Convém mencionar que apesar de fazer uma análise sobre o uso de tecnologias nas práticas sociais, o presente trabalho evita visões alarmistas e deterministas que sugerem cenários catastróficos provocados pela inteligência artificial. Para isso, adota-se uma definição mais simples de artificialidade como algo que foi construído por um grupo de organismos com um propósito específico (Romero, 2024). Assim, ainda que se reconhece a não neutralidade da tecnologia, especialmente das tecnologias disruptivas que afetam o trabalho, defende-se que o modo de utilização dessa artificialidade reside nos seres humanos, que devem estar conscientes dos riscos e desafios éticos a ela relacionados (Romero, 2024).

Com a ampliação dos sistemas sociais por meio da tecnologia, surgem novas dinâmicas e interações. Os sistemas sociais, incluindo os tecnológicos, são compostos por elementos interdependentes que transformam as dinâmicas e estruturas das práticas sociais (Romero, 2024). Sob essa ótica, tecnologias como a inteligência artificial (IA) impactam o trabalho ao introduzirem novos elementos —como novas formas de opacidade— nas práticas laborais. Nesse contexto, o problema central investigado neste artigo é o impacto da opacidade algorítmica —e novas formas de segregação como consequência dessa opacidade— sobre as práticas e valores do trabalho. Quando algoritmos opacos tornam-se mediadores de decisões no ambiente laboral, observa-se uma segregação epistêmica que compromete a autonomia e a transparência dessas práticas, perpetuando

desigualdades e incertezas. A questão que se coloca é se essa opacidade —ao tornar o ambiente de trabalho mais arbitrário e desigual— pode também alterar a concepção e o valor social do trabalho, deixando-o mais suscetível a práticas exploratórias e à erosão fragmentada dos direitos trabalhistas.

Uma perspectiva crítica sobre os efeitos da opacidade algorítmica no ambiente de trabalho sugere que essa característica dos algoritmos pode não apenas transformar práticas laborais, mas também modificar a noção mesma do próprio conceito de trabalho —isto é, como representamos o trabalho no mundo e como o compreendemos socialmente. A opacidade algorítmica, ao dificultar o entendimento dos processos decisórios mediados por IA, reduz a capacidade dos trabalhadores de compreender, contestar ou participar efetivamente dessas decisões. Em vez de explorar cenários especulativos de uma possível dominação da IA, o presente trabalho enfatiza que os efeitos concretos e já visíveis da inteligência artificial exigem atenção, pois, em vez de eliminar empregos de forma abrupta, a IA vem contribuindo para a centralização de poder das empresas, intensificando desigualdades pré-existentes e limitando a autonomia dos trabalhadores (Kapoorsayash Kapoor e Narayanarvind Narayanan, 2023). Partindo da inspiração da noção de “disrupção conceitual” (Lohr, 2023), este trabalho propõe que a opacidade algorítmica desafia as normas e valores do trabalho, introduzindo nessa esfera da vida novos elementos que demandam uma reavaliação crítica e urgente dessa prática social. Para mitigar os efeitos prejudiciais da opacidade, torna-se essencial implementar regulamentações que promovam a transparência e protejam os direitos dos trabalhadores.

Para desenvolver esta análise, inicialmente será apresentada uma definição de trabalho e dos valores envolvidos nesta prática social. Essa definição ajudará a relacionar as práticas laborais normativas com os problemas gerados pela

opacidade algorítmica. Em seguida, será introduzido o conceito de algoritmo e esclarecida a questão da opacidade associada a sistemas de aprendizado de máquina, que gera uma preocupação epistêmica devido à dificuldade de entender como os algoritmos chegam aos resultados classificatórios usados nos processos decisórios no ambiente de trabalho.

Na sequência, aborda-se diretamente a opacidade algorítmica na mediação do trabalho. Em primeiro lugar, discute-se o aspecto sistêmico da opacidade e o potencial desta, quando aplicada em larga escala, de gerar resultados indesejáveis no contexto laboral, interferindo na realidade sociopolítica. Em segundo lugar, será explorado como as preocupações epistêmicas abrangem também valores não epistêmicos relacionados ao impacto da opacidade, exemplificados pela segregação epistêmica que afeta trabalhadores precarizados nas plataformas digitais. O modelo de negócios dessas empresas demonstra como a tecnologia disruptiva pode gerar mal-estar laboral, corroendo os direitos e valores associados ao trabalho.

Com inspiração na perspectiva de Guido Löhr (2023), argumenta-se que as tecnologias socialmente disruptivas têm o potencial de modificar nossas concepções sobre conceitos estabelecidos. Se os processos decisórios no trabalho forem mediados por uma opacidade algorítmica sistêmica que resulta em consequências indesejáveis, como a segregação epistêmica, e se a introdução de perturbações na realidade altera o conceito representado pela prática, então a deliberação mediada pela opacidade pode contribuir para transformar tanto a prática social do trabalho quanto o conceito que ela representa.

Finalmente, este artigo argumenta que a opacidade algorítmica mediando o trabalho é capaz de alterar a rede conceitual do trabalho, uma vez que nossos conceitos estão vinculados ao acesso tecnológico ao mundo. Mudanças nas normas e práticas sociais introduzidas pelo

desenvolvimento tecnológico impactam a rede conceitual, afetando o próprio conceito de trabalho. A análise demonstrará como a transformação dessa rede começa com a introdução ou eliminação de novos conceitos e novas relações inferenciais entre os conceitos existentes. E assim começamos.

2. A prática social do trabalho

Para esboçar uma definição para a noção de trabalho, recorre-se à análise de Raymond Geuss (2021) no livro *A Philosopher Looks at Work*. Raymond Geuss (2021) propõe três componentes centrais e comuns para designar uma atividade como trabalho: (i) exigir esforço mensurável e extenuante —dependendo de condições físicas, psicológicas e estados de treinamento; (ii) suprir uma necessidade no sentido de sobrevivência individual e coletiva e (iii) participar de uma objetividade —no sentido de que pode ser entendido como um processo separado e com possibilidade de avaliação e execução independentes daqueles que o realizam. Além desses três componentes, Geuss (2021) também identifica mais três características adicionais para definir uma atividade como trabalho: (iv) ter distinção e independência em relação a outras atividades; (v) divergir de diversão, prazer ou brincadeira; (vi) contar com remuneração.

Por objetividade, Geuss (2021) refere-se a: a) possibilidade do trabalho seguir procedimentos objetivos e manipular coisas por meio de processos mecânicos identificáveis em produção contínua; b) a possibilidade dos procedimentos realizados no trabalho poderem ser seguidos por outros em caso de substituição do trabalhador; e c) a capacidade do trabalhador agir de acordo com uma série de formas que podem ser identificadas com um resultado desejado. Para melhor compreensão, convém especificar

que "objetividade" aqui se refere ao caráter sistemático e replicável das tarefas realizadas no trabalho, de maneira a permitir sua análise e avaliação independente. Sob a perspectiva de Danaher e Nyholm (2021), essa objetividade pode ser compreendida como uma estrutura de tarefas que envolve a contribuição de um grande número de pessoas (os trabalhadores). Eles definem o espaço que estrutura essas tarefas como o espaço de trabalho.

Para o propósito desta análise, serão utilizados integralmente os itens (i), (ii) e (iii) propostos por Geuss (2021). Essa escolha é basicamente pragmática, sendo os itens mencionados, segundo Geuss (2021), os mais comuns para designar uma atividade como trabalho. Considera-se que os demais seriam apenas complementos que podem variar segundo o tipo de atividade e que podem gerar desacordos. Por exemplo, certos trabalhos como a função de "influencer" não é exercida com independência de outros âmbitos da vida, pelo contrário, explora-se outros âmbitos, como a intimidade, para o ganho financeiro. Parte-se também da ideia de Marx (2010) de que o trabalho emerge na sociedade a partir dos seus modos de organização, como um processo entre o homem e a natureza. Adicionalmente, se considerará a visão de Danaher e Nyholm (2021) de que, do ponto de vista econômico, o trabalho envolve o desenvolvimento de competências específicas, referidas como "*skills*". Essas competências são habilidades técnicas e não técnicas adquiridas e aplicadas ao longo do tempo. Acredita-se que essas ideias podem ser combinadas para formular uma definição conceitual de "trabalho" útil para a análise proposta. Assim, propõe-se que algo é trabalho (T) se e somente se: (i) T é um processo que nasce de uma dada organização social X, (ii) T está envolvido em uma objetividade P, (iii) T é uma atividade Q (que requer esforço ou gasto de energia que pode ser mensurado, depende de condições físicas, psicológicas e estados de treinamento, se diferencia do

lazer e consiste no desenvolvimento de competências), (iv) T supre uma necessidade Y (seja biológica, econômica, por exemplo, a remuneração, ou psicológica).

Em outras palavras, T é trabalho se e somente se é um processo que nasce em uma dada organização social X dentro de uma objetividade P, subdividido em um tipo de atividade Q para suprir uma necessidade Y. Essa definição permite uma visão abrangente do trabalho, integrando aspectos teóricos e práticos.

Além de definir o conceito, afirmar que uma prática social é considerada trabalho implica, segundo Horgan (2021), que essa prática segue um marco de direitos definidos politicamente. Esse marco faz do trabalho uma prática social regulamentada por um conjunto específico de leis e regulamentos que definem direitos e responsabilidades, como a Lei X e a Norma Y, que estabelecem diretrizes claras para a proteção dos trabalhadores e a responsabilização dos empregadores (Horgan, 2021, p. 1113). Isso significa que a definição de trabalho não é apenas teórica, mas também envolve a aplicação de normas e direitos específicos. Danaher e Nyholm (2021) destacam que o trabalho ocupa um espaço central na vida das pessoas.

Essas normas e direitos, reconhecidos pela Organização Internacional do Trabalho —OIT (2004), fazem do trabalho uma parte central da vida e podem ser sintetizados no conceito de "trabalho decente". Segundo a Organização Internacional do Trabalho, o trabalho decente deve: (i) "dignificar", (ii) "permitir o desenvolvimento das capacidades individuais", (iii) "distinguir-se de outros", (iv) "cumprir com a lei", (v) "remunerar justamente", (vi) "não discriminar" e (vii) "cumprir com normas de proteção social, diálogo e tripartismo". A definição de trabalho decente inclui valores fundamentais associados ao trabalho, "incluindo dignidade, realização, autonomia, mérito, distinção, responsabilidade, justiça, igualdade e comunalidade". Na perspectiva adotada aqui, esses

valores incluem, por exemplo, dignidade, realização, autonomia, mérito, distinção, responsabilidade, justiça, igualdade e comunalidade. Antes de abordar o tema da opacidade algorítmica na mediação do trabalho, será realizada uma breve análise isolada do problema da opacidade, examinando as suas causas e problemas.

2. O problema epistêmico da opacidade

Esta seção começa com a introdução da noção de algoritmo. Para isso, adota-se a definição de Hill citada por Tsamados et al. (2021) em *The Ethics of Algorithms: Key Problems and Solutions*. Segundo essa definição, um algoritmo é uma construção matemática composta por uma estrutura de controle finita, abstrata, efetiva e imperativamente definida, que cumpre um propósito específico sob certas condições (Tsamados et al., 2021, p. 3). Para alcançar esse propósito, o algoritmo deve estar integrado em um sistema, que, por sua vez, é implementado em programas e aplicações (Mittelstadt et al., 2016, p. 2). Dentro de um sistema, o algoritmo — projetado pelo raciocínio humano — segue uma racionalidade que determina sua função de analisar um conjunto de dados de um domínio específico e, a partir disso, incorpora uma estrutura matemática abstrata por meio de processos operacionais (Mittelstadt et al., 2016, p. 2). No entanto, é importante observar que a complexidade e a opacidade dos algoritmos podem variar amplamente, o que nem sempre é refletido nas definições mais gerais. Justin Joque (2022) explica que Donald Knuth, cientista da computação, define "a arte da programação de computadores" como um sistema baseado em cinco requisitos: (i) a finitude, pois um algoritmo tem um tempo limitado para ser executado e concluído; (ii) a definição, pois cada passo do algoritmo é definido

"rigorosamente e sem ambiguidade"; (iii) uma entrada, que constitui um processo de entrada que resulta em (iv) uma saída; e (v) "a efetividade do processo, marcada pela simplicidade de cada passo para uma compreensão fácil e possível execução" (Joque, 2022, p. 67).

Os algoritmos podem ser de diferentes tipos, mas compartilham um ponto em comum, que é o "intento de trazer algo ao mundo material", seja esse algo uma ideia, um cálculo ou uma experiência anterior (Toscano, 2022, p. 50). Apesar da definição clara de algoritmos, a implementação e a complexidade dos sistemas de aprendizado de máquina (ML — *machine learning*) podem levar a diferentes níveis de opacidade. O foco, neste caso, está nos sistemas que utilizam algoritmos de aprendizado de máquina (ML). Esses sistemas, segundo Loi et al., "passam por um processo de treinamento baseado em dados digitais" (Loi et al., 2021, p. 13). Nos sistemas de ML, os algoritmos recebem dados de entrada (*inputs*) e, a partir deles, criam uma cadeia de dados que são avaliados como saídas (*outputs*) (Joque, 2022, p. 45). A opacidade desses sistemas pode variar de acordo com o tipo de modelo de ML e a complexidade do treinamento. É relevante notar que esses algoritmos são utilizados em programas e aplicações que automatizam processos deliberativos, como sistemas de classificação de crédito, uso clínico e análise de reincidência em sistemas de justiça e policiais (Loi et al., 2021, p. 13; Mittelstadt et al., 2016, p. 3). Neste trabalho, considera-se um processo deliberativo qualquer processo que envolve tomada de decisão, como decidir sobre ações em relação a situações específicas, por exemplo, a demissão de um funcionário. Quando algoritmos são usados em processos deliberativos, a dificuldade em entender como produzem suas decisões pode comprometer a transparência e a justiça.

A função classificatória dos sistemas de ML suscita algumas preocupações. Segundo Tsamados et al. (2021), essas preocupações se

dividem em duas categorias: (i) epistêmicas e (ii) normativas. As preocupações epistêmicas referem-se à evidência inconclusiva, inescrutável e incorreta. Por outro lado, as preocupações normativas estão relacionadas a resultados injustos e impactos transformadores. Além disso, Tsamados et al. (2021) também discutem a questão da rastreabilidade, que pode ser tanto epistêmica quanto normativa. No presente trabalho, o foco está na evidência inescrutável, associada à opacidade e à falta de transparência nos processos algorítmicos.

A evidência inescrutável, uma das preocupações éticas levantadas pela tecnologia algorítmica nos sistemas de aprendizado de máquina, resulta em opacidade. A opacidade ocorre porque os algoritmos de ML operam de maneira que o processo interno de tomada de decisão não é plenamente compreendido nem explicado. Para compreender melhor a opacidade, é necessário retornar ao tema dos algoritmos de aprendizado. Nos sistemas de aprendizado, os algoritmos passam por um processo de capacitação e treinamento que pode ser descrito tecnicamente como o desafio de “melhorar alguma medida de desempenho P ao executar uma tarefa T, por meio de algum tipo de experiência de treinamento E” (Loi et al., 2021, p. 254). No entanto, a capacidade de aprimorar o desempenho pode ser limitada pela dificuldade em interpretar os processos internos do algoritmo. Na prática, isso ocorre porque o processo de programação algorítmica envolve etapas como a entrada e a saída de dados. Durante o treinamento, os dados são inseridos (entrada) e, posteriormente, integrados em uma infraestrutura que gera novas previsões (saída), automatizando decisões (Loi et al., 2021, p. 255). Assim, o sistema opera como um “motor cognitivo” (Loi et al., 2021, p. 255) que que processa os dados de entrada e classificando-os para a tomada de decisão (Burrell, 2016, p. 1). Esse funcionamento confere aos sistemas um certo grau de autonomia (Mittelstadt et al.,

2016, p. 3). No entanto, esse grau de autonomia aliado à dificuldade de interpretação pode resultar em decisões que não são totalmente compreendidas, gerando um nível de opacidade preocupante.

Loi et al (2021) afirmam que os processos deliberativos em algoritmos de aprendizado operam dentro de um limite epistêmico que impede a compreensão completa de seus processos internos, tornando-os inescrutáveis. De acordo com Loi et al. (2021, p. 256), “uma explicação sobre X é eficaz quando o que é explicado sobre X é claro e acessível para as pessoas que desejam entender X”. Porém, a inescrutabilidade dos sistemas de ML pode comprometer essa clareza e, conseqüentemente, a eficácia comunicativa. Esse fator deve-se principalmente à falta de transparência no “comportamento do sistema algorítmico” (Loi et al., 2021, p. 256). Esse problema não seria crítico caso as decisões baseadas em ML não impactassem direta ou indiretamente a vida das pessoas. Contudo, elas têm um impacto significativo.

O conceito de transparência algorítmica é definido pela sua eficácia comunicativa. Se a comunicação fornece as “informações essenciais necessárias” para explicar o design do sistema de forma satisfatória, então ela é transparente (Loi et al., 2021, p. 257). Isso é crucial para determinar se as decisões algorítmicas são “moralmente arbitrarias” (Loi et al., 2021, p. 260). Para entender o que constitui uma decisão moralmente arbitrária, é importante compreender o conceito de moral. Segundo González e Teixidó (2023) em “*Derechos, deberes, justicia y simetría moral desde una filosofía científica*”, a moral pode ser vista como um conjunto de normas sociais, construídas individual ou coletivamente, com o propósito de motivar ações ou inibições em situações envolvendo valores humanos em conflito. Nesse contexto, a moral orienta o comportamento humano quando diferentes valores estão em jogo. Os algoritmos, ao reproduzirem sistemas de regras que orientam decisões,

podem incorporar elementos de teor moral. A arbitrariedade dessas decisões, portanto, pode surgir (i) da reprodução de sistemas morais desiguais e (ii) da falta de transparência sobre as regras seguidas pelos algoritmos. González e Teixidó (2023) argumentam que a equidade moral busca o equilíbrio simétrico entre normas estabelecidas socialmente. Assim, se aplicamos essa ideia para nosso tema, podemos afirmar que a arbitrariedade algorítmica viola princípios morais ao gerar ou perpetuar desequilíbrios já existentes. Os algoritmos de *machine learning* aprendem com os dados que lhes são fornecidos durante o treinamento. Se esses dados refletirem preconceitos e desigualdades da sociedade, as decisões algorítmicas tenderão a replicar essas distorções. Por exemplo, um algoritmo treinado com dados históricos de contratações pode associar cargos de liderança predominantemente ao gênero masculino, discriminando candidatas mulheres.

A transparência é, portanto, fundamental para permitir que as partes interessadas compreendam e avaliem as decisões algorítmicas de forma adequada. Se a comunicação falha em fornecer as informações necessárias, a transparência não é alcançada, resultando em opacidade. A opacidade dos sistemas algorítmicos é uma preocupação epistêmica central, pois a dificuldade em entender como os algoritmos chegam a determinadas conclusões pode levar a decisões moralmente arbitrárias (Burrell, 2016, p. 1).

Alguns autores, como Ferreira (2021), argumentam que a opacidade não é um problema sério, pois a natureza inescrutável dos sistemas algorítmicos não representa maior risco de arbitrariedade do que as decisões humanas. Ferreira defende que, apesar dos desafios epistêmicos, os sistemas de ML podem ser justos e eficazes em processos deliberativos. Contudo, mesmo reconhecendo que os algoritmos podem produzir resultados justos em certos casos, a opacidade continua sendo uma preocupação ética relevante. A falta de compreensão

dos processos internos pode gerar sérias implicações éticas, especialmente em contextos que exigem justiça e responsabilidade. Portanto, é crucial abordar a opacidade de forma crítica e abrangente, assegurando que os sistemas de aprendizado de máquina sejam empregados de maneira ética e responsável, particularmente em ambientes onde as decisões afetam diretamente a vida das pessoas.

3. A prática social do trabalho mediada pela tecnologia algorítmica

Até aqui, as noções de trabalho e opacidade foram discutidas separadamente. Agora, é necessário integrá-las para avançar em direção à proposta do presente trabalho. Como mencionado, algo pode ser considerado trabalho se e somente se é um processo que surge em uma organização social X e se envolve em uma objetividade P , dividida em um tipo de atividade Q para suprir uma necessidade Y . Discutir a mediação algorítmica no trabalho é examinar como os sistemas algorítmicos estão presentes na objetividade do ambiente laboral, influenciando as regras e protocolos que determinam a atividade em questão.

Se um sistema de aprendizado algorítmico classifica dados em um domínio específico, os processos deliberativos relacionados a esse domínio são impactados pela natureza dessa classificação. No contexto do trabalho, quando os processos inferidos por sistemas classificatórios são opacos, isso afeta diretamente o domínio laboral, potencialmente gerando arbitrariedade, uma vez que as decisões seguem um processo não transparente. Decisões arbitrárias podem levar a consequências indesejáveis, como a segregação epistêmica, que, neste estudo, é entendida como a divisão de indivíduos

em grupos com acesso desigual a informações e conhecimentos. A segregação epistêmica representa um desafio para a democracia ao limitar a diversidade de perspectivas, dificultar a construção de conhecimento compartilhado e comprometer a qualidade do debate público e da tomada de decisões coletivas. Antes de explorar a segregação epistêmica em mais profundidade, é importante esclarecer os tipos de trabalho mediado por sistemas algorítmicos.

Exemplos de funções laborais mediadas por algoritmos incluem processos de seleção de trabalhadores (Páez e Ramírez-Bustamante, 2022), avaliações de desempenho e produtividade em plataformas de entrega de alimentos como *Glovo*, *Deliveroo* e *UberEats*, e até sistemas para aplicação de medidas disciplinares e demissões (Asquerino Lamparero, 2022). Essas classificações têm impactos concretos: influenciam a contratação, a remuneração e a desativação de trabalhadores (Suarez, 2023).

Retornando à questão da segregação, entendemos que ela é uma consequência da opacidade algorítmica. Portanto, antes de abordá-la diretamente, é preciso esclarecer o conceito de opacidade algorítmica no contexto sistêmico.

3.1. A opacidade algorítmica sistêmica

Ferreira (2021) argumenta que a opacidade faz parte da natureza dos algoritmos e, por si só, não é problemática. Burrell (2016), por sua vez, identifica três formas de opacidade: (i) opacidade por direito de propriedade, associada à proteção corporativa e ocultamento institucional (Burrell, 2016, p. 3); (ii) opacidade por "analfabetismo técnico", que exige conhecimento especializado para decodificar algoritmos; e (iii) opacidade resultante do desajuste entre o aprendizado de máquina e a interpretação humana (Burrell, 2016, p. 2). Essas formas ajudam

a compreender por que Ferreira (2021) defende que sempre haverá algum grau de opacidade. Contudo, essa perspectiva não é suficiente.

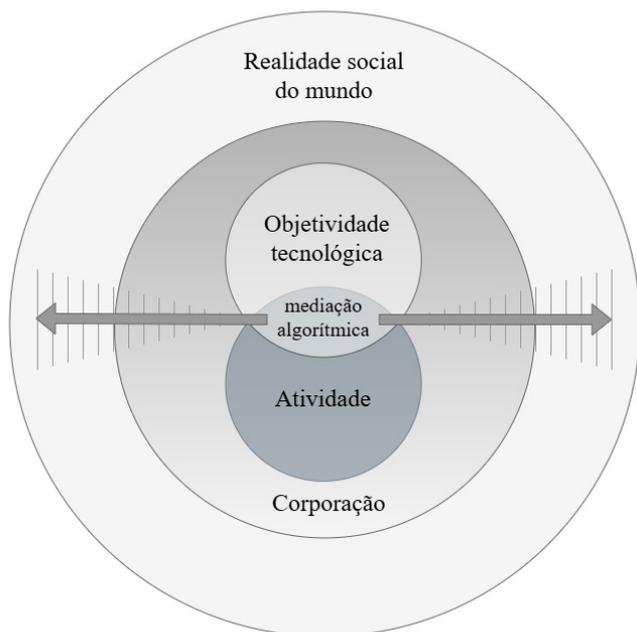
Creel e Hellman (2022) argumentam que o problema da opacidade é sistêmico, não isolado. Quando a opacidade se torna sistemática em grande escala, ocorre uma aplicação uniforme de classificações que pode resultar em segregação social e econômica, o que as autoras chamam de opacidade epistêmica sistêmica. Esse problema começa com o monopólio de fornecedores privados que projetam a maioria dos sistemas algorítmicos, padronizando preconceitos e critérios de design (Creel e Hellman, 2022, p. 1).

A análise de Creel e Hellman (2022) permite avaliar a dimensão do uso desses sistemas no contexto do trabalho. Por exemplo, a concentração de controle e a falta de transparência nos sistemas algorítmicos utilizados por plataformas digitais de entrega de alimentos ilustram como a opacidade pode levar ao que Stuart et al. (2023) chamam de um "mal-estar laboral" (p. 3). Com poucas grandes empresas dominando o setor, a falta de transparência nos critérios de avaliação pode resultar em práticas uniformemente prejudiciais. Segundo o documento "Infografía—Panorámica de las plataformas digitales en la UE" (2023), havia 28,3 milhões de trabalhadores de plataformas digitais na União Europeia em 2022 e a expectativa é que esse número atinja 43 milhões até 2025. Se nenhuma mudança for implementada, mais pessoas estarão sujeitas a ambientes digitais marcados pela opacidade e segregação.

Portanto, a perspectiva de Creel e Hellman (2022) sugere que, quando aplicada uniformemente em larga escala, a opacidade sistêmica afeta não apenas o ambiente de trabalho, mas toda a realidade sociopolítica. Esse impacto se dá de dentro — da objetividade tecnológica das empresas — para fora, afetando o mundo sociopolítico, como ilustrado na imagem abaixo:

Figura 1: Visão global e sistêmica da mediação algorítmica no trabalho

Fonte:Elaborada pela autora (2024)



Voltando a definição de trabalho apresentada na parte 1 do presente trabalho, o que se pode mencionar é que ao tornar-se parte integrante das estruturas de trabalho, os algoritmos inserem-se no processo social em que o trabalho é produzido, transformando a organização do próprio ambiente laboral (condição (i) do conceito). A interferência dessa tecnologia gera um novo tipo de objetividade (ii), centrada nas decisões automatizadas que escapam à compreensão direta dos trabalhadores e, com isso, introduzem arbitrariedades e segregações epistêmicas que prejudicam a transparência. Além disso, o esforço e o gasto de energia dos trabalhadores (iii) são medidos e monitorados por sistemas algorítmicos que determinam as métricas de produtividade e desempenho sem transparência, limitando a autonomia dos trabalhadores para ajustar suas competências em

função dos critérios objetivos. Por fim, a opacidade dos algoritmos pode também impactar a satisfação das necessidades dos trabalhadores (iv) ao enfraquecer as normas que tradicionalmente garantem remuneração justa e segurança no trabalho. Dessa forma, os elementos que constituem o conceito de trabalho são alterados, pois o caráter opaco dessas tecnologias redefine as expectativas sobre o que constitui um “trabalho decente” e digno de reconhecimento social, comprometendo valores centrais como autonomia, justiça e igualdade.

Em resumo, a prática social do trabalho mediada pela tecnologia algorítmica é impactada por processos deliberativos caracterizados pela opacidade, levando à arbitrariedade e à segregação epistêmica. Como discutido, essa segregação é evidente em processos de seleção e avaliação, muitas vezes projetados por poucos fornecedores que seguem padrões que exacerbam injustiças e limitam a compreensão dos trabalhadores sobre as decisões que os afetam. A opacidade favorece as empresas, prejudica os trabalhadores e limita a capacidade de contestação.

3.2. Opacidade epistêmica e segregação epistêmica

De acordo com Sullivan (2022), as preocupações epistêmicas relacionadas à opacidade também envolvem valores não epistêmicos, como a preocupação por justiça, responsabilidade e equidade. A complexidade dos modelos algorítmicos de aprendizado automático dificulta a explicação de como chegam às suas conclusões. Isso é particularmente relevante em contextos em que esses modelos influenciam decisões significativas, como as que afetam trabalhadores em processos de seleção e gestão de pessoal. Considerando as preocupações epistêmicas

como a dificuldade de entender os processos de classificação dos sistemas, a falta de explicações claras compromete a transparência das informações que afetam a vida das pessoas (Sullivan, 2022, p. 6). Essa ausência de clareza afeta a capacidade de explicação, prejudicando sistematicamente os interesses dos trabalhadores.

Para ilustrar casos concretos, considere as plataformas digitais. Uma carta do sindicato dos trabalhadores dessas plataformas na França revela que a opacidade impede a compreensão de como as tarifas horárias são definidas, dificultando a avaliação e reivindicação de valores justos (Suarez, 2023). Além disso, a falta de clareza nos critérios de desativação repentina das plataformas impede os trabalhadores de questionar tais decisões. O relatório *Algorithmic Management: Opportunities for Collective Action* (UNI Global Union, 2023) aponta que os sistemas algorítmicos monitoram rigorosamente o desempenho, impondo metas frequentemente irrealistas e afetando a saúde física e mental dos trabalhadores. Dubal (2023) observa que, na plataforma *Uber*, a opacidade nos cálculos de pagamento promove discriminação salarial, resultando em desigualdades injustificadas. Essa falta de transparência intensifica a segregação epistêmica, prejudicando a capacidade dos trabalhadores de compreenderem as causas das diferenças salariais. Portanto, a opacidade sistêmica afeta negativamente a dimensão laboral, refletindo-se em vários países e contextos. Esse problema não se restringe a casos específicos, mas é um fenômeno global com amplas implicações.

Além disso, Páez e Ramírez-Bustamante (2022) destacam que a opacidade algorítmica afeta processos de contratação e avaliação de desempenho. Algoritmos de seleção podem perpetuar preconceitos ao basear-se em dados que refletem desigualdades sociais. Consequentemente, certos grupos são excluídos de oportunidades de emprego, enquanto outros são favorecidos por critérios opacos e

injustos. Avaliações de desempenho também sofrem com a falta de transparência, dificultando a contestação de decisões e o entendimento de como melhorar dentro dos parâmetros das plataformas. Essa opacidade contribui para a segregação epistêmica, privando os trabalhadores das informações necessárias para compreender seus direitos e desempenho.

A combinação dessas formas de opacidade — que vai desde a propriedade e o analfabetismo técnico até o desajuste epistêmico sistêmico — revela um quadro complexo onde a transparência e a equidade no ambiente de trabalho são comprometidas. Portanto, abordar criticamente esse fenômeno e implementar soluções que promovam a transparência são medidas essenciais. Tais soluções podem incluir regulamentações mais rigorosas sobre algoritmos em processos laborais e práticas que assegurem decisões mais justas e compreensíveis.

Em resumo, a análise da opacidade algorítmica e seus impactos sistêmicos revela a necessidade urgente de ações para garantir que as tecnologias utilizadas no trabalho sejam projetadas e aplicadas de maneira transparente e equitativa. Somente assim é possível mitigar os efeitos prejudiciais da segregação epistêmica e assegurar que o trabalho mediado por algoritmos respeite os direitos e a dignidade dos trabalhadores, promovendo um ambiente de trabalho mais justo e inclusivo.

3.3. Alterações no conceito de trabalho

Antes de discutir os impactos da opacidade algorítmica nas práticas laborais, é importante reconhecer, como argumenta Stephanie Hare (2022, p. 16), que a tecnologia não é neutra. Hare entende a tecnologia como parte do que nos torna humanos e destaca que ela media a relação entre cidadãos e governos,

consumidores e empresas, entre outros. Além disso, Hare (2022, p. 42 e 43) argumenta que a tecnologia não é apenas uma ferramenta, mas um sistema complexo que envolve aspectos como racionalidade, poder, controle, linguagem e conhecimento. Nesse sentido, Hare considera que o termo "caixa preta", usado para descrever a falta de transparência dos sistemas algorítmicos também pode ser apropriado para representar o desconhecimento sobre o funcionamento do sistema, o que dificulta não só a compreensão de seus processos internos, mas também a previsão de como o uso contínuo do sistema pode afetar a realidade ao longo do tempo (Hare, 2022, p. 59).

Embora seja fascinante refletir sobre as possíveis mudanças que a tecnologia pode provocar na realidade, o objetivo aqui é analisar especificamente a opacidade algorítmica na mediação do trabalho. Em vez de explorar as múltiplas implicações dessa mediação, a análise se concentrará em refletir em como os mecanismos de representação podem ser afetados. Interessa pensar como as pequenas mudanças na normatividade das práticas laborais podem influenciar em como representamos a prática que conhecemos como trabalho. Considerando que a linguagem se organiza em um sistema de representações, o objetivo não é apenas discutir o trabalho como prática social, mas também como um conceito que envolve interações objetivas e estruturadas, passíveis de análise.

Inspirada na perspectiva de Guido Löhr (2023) sobre como as tecnologias socialmente disruptivas podem mudar nossas concepções de determinados conceitos, argumenta-se que, se os processos deliberativos no trabalho são mediados por uma opacidade algorítmica que gera efeitos indesejáveis, como a segregação epistêmica sistêmica, e se a introdução de perturbações na realidade pode mudar o conceito que essa prática representa, então a deliberação mediada pela opacidade pode contribuir para transformar tanto a prática social do trabalho

quanto o conceito associado a ela. Em outras palavras, quando produz segregação epistêmica, a opacidade sistêmica pode transformar não apenas a prática social conhecida como trabalho, mas também a nossa concepção do trabalho e dos valores que ele representa. De maneira semelhante, Horgan (2021) propõe que o deslocamento do conceito de trabalho em múltiplas direções — como trabalho emocional, trabalho autônomo, etc. — adiciona novos elementos (ou novos conceitos) ao conceito principal, alterando assim como o assimilamos e representamos.

Hare (2022) afirma que a tecnologia faz parte do que nos torna humanos, e Löhr (2023) destaca que, de acordo com os filósofos da tecnologia, nosso acesso ao mundo ocorre necessariamente através da mediação tecnológica. Löhr sugere que nossos conceitos estão envolvidos nesse acesso, implicando que discutir o desenvolvimento tecnológico é também discutir nosso aparato conceitual. A tecnologia socialmente disruptiva (TSD) é definida como aquela que provoca modificações significativas nas práticas sociais, gerando grandes transformações (Rueda, 2021, p. 6). Por essa definição, as plataformas digitais podem ser consideradas tecnologias socialmente disruptivas. Quando aplicadas ao trabalho, essas tecnologias causam mudanças e perturbações na percepção sociopolítica que o trabalho representa. O exemplo das plataformas ilustra como casos específicos de opacidade podem resultar em segregação epistêmica que afeta uma classe social específica, criando perturbações no ambiente laboral.

Löhr (2023) destaca que qualquer alteração em nossas percepções sobre um conceito provoca mudanças na forma como pensamos sobre ele. Partindo dessa ideia, argumenta-se que a opacidade dos sistemas algorítmicos pode transformar nossa percepção sobre o trabalho, conduzindo a alterações futuras em nosso entendimento do conceito. Uma perturbação pode ser definida como qualquer alteração no

estado atual. Em outras palavras, rupturas gradativas na prática social do trabalho têm potencial para modificar nossa percepção dessa dimensão social, resultando em mudanças conceituais. Não se pode afirmar se essas mudanças serão positivas ou negativas. Löhr (2023) usa o exemplo do ventilador mecânico para ilustrar como a tecnologia alterou nossa compreensão sobre a morte — antes definida por critérios cardiopulmonares, agora por parâmetros neurológicos. As transformações nas condições objetivas do trabalho, como as mudanças tecnológicas e sociais, influenciam diretamente a estrutura e as práticas que o definem. Dessa forma, a percepção alterada do trabalho, resultante da interferência de tecnologias complexas e opacas, pode conduzir a modificações gradativas em sua conceituação, assim como a introdução de novos critérios de avaliação.

Löhr (2023, p. 1) explica que o desenvolvimento tecnológico pode alterar a compreensão dos propósitos e normas de um conceito, gerando mudanças sociais e morais significativas. Quando ocorre uma mudança em torno de um conceito, há uma ruptura que afeta nosso entendimento sobre ele, levando à "incerteza classificatória" (Löhr, 2023, p. 1). Isso implica uma incerteza nas condições usadas para classificar o conceito. Por exemplo, a discriminação salarial enfrentada por trabalhadores de plataformas ilustra como práticas discriminatórias podem violar normas de não discriminação e segurança no trabalho, introduzindo novos elementos como instabilidade e desigualdade. Essas mudanças podem, gradualmente, minar os direitos trabalhistas e contribuir para alterações na definição. À luz dessas considerações, podemos afirmar que as perturbações na realidade laboral, tais como mudanças tecnológicas, sociais ou políticas, podem contribuir para alterar o conceito de trabalho ao modificar as condições objetivas em que ele ocorre, afetando a estrutura de suas práticas, as relações entre trabalhadores

e empregadores, e as formas de organização e distribuição das tarefas.

Nesse ponto do presente trabalho, é pertinente apresentar um exemplo mais empírico de como ocorre uma mudança no próprio conceito de trabalho.¹ No capítulo intitulado "Why and How Precarious Workers Support Neo-illiberalism" do livro "Beyond Neoliberalism and Neo-illiberalism", Rosana Pinheiro-Machado (2024) afirma que trabalhadores de plataforma de baixa renda no Brasil, apesar da situação de precariedade, identificam-se com o discurso de empreendedorismo promovido pelas plataformas digitais. De acordo com os estudos de Pinheiro-Machado (2024), realizados ao longo de 23 anos, há uma incorporação dos valores relacionados às práticas empreendedoras entre esses trabalhadores. Entre esses valores, podemos mencionar a valorização da autonomia e da responsabilidade individual na busca pela prosperidade. Segundo a autora, essa alteração na perspectiva faz com que a responsabilidade pela desigualdade não recaia sobre o Estado, mas sim sobre o esforço do indivíduo. Nesse sentido, há uma transferência de responsabilidade na percepção do conceito de trabalho. Pinheiro-Machado (2024) explica que, de acordo com seus estudos, a maioria desses trabalhadores considera que um trabalho convencional, ou seja, aquele em que se trabalha para outra pessoa, é visto como sinônimo de acomodação na vida. Em outras palavras, o trabalho convencional não é mais valorizado e foi substituído pela valorização do empreendedorismo. Segundo Pinheiro-Machado (2024), houve uma transformação na percepção da ascensão social via trabalho, onde, se antes se valorizava a estabilidade e segurança de um emprego, hoje valoriza-se a promessa de sucesso por meio do trabalho duro (entendido como a normalização da precariedade) e da prática do empreendedorismo. A autora conclui que a precariedade econômica favorece a adesão dos trabalhadores aos

tipos de trabalho fomentados pelas plataformas. Podemos afirmar que esse exemplo ilustra bem a ideia de mudança conceitual.

Nesse contexto, podemos observar que a introdução de novos valores nas normas da prática laboral, como a autonomia e a responsabilidade pessoal sobre a desigualdade, contribui para a erosão dos valores anteriores. Isso pode ser facilmente relacionado às perspectivas de segregação e opacidade mencionadas anteriormente. Nesse caso, a opacidade contribui para criar um ambiente de trabalho em que os processos decisórios são vistos como arbitrários, o que leva à desconfiança e ao enfraquecimento das normas laborais convencionais. Por sua vez, o enfraquecimento das normas laborais convencionais afeta a forma como o trabalho é representado e compreendido socialmente, enfatizando a individualidade e o empreendedorismo em detrimento de valores como segurança e transparência. Vale também mencionar que, no que diz respeito à opacidade, o limite ao acesso dos trabalhadores às informações sobre os critérios e métodos usados para decisões que afetam suas vidas reforça a desigualdade e tende a fomentar um ambiente de desigualdade, transferindo a responsabilidade para o indivíduo em vez do sistema e dos mediadores. Isso pode contribuir para a percepção equivocada de que a condição de precariedade é uma falha pessoal, e não uma alteração nas normas sociais. Por fim, o que podemos observar é que segregação e opacidade contribuem para a normalização de práticas precárias e desiguais, influenciando a forma como o trabalho é entendido. A introdução de novos elementos, como a autonomia e a responsabilidade individual pelo próprio sucesso financeiro, tal como impulsionada pelas plataformas digitais na prática social do trabalho, conforme destacada por Pinheiro-Machado (2024), é um exemplo de como a opacidade e a segregação contribuem para transformar a percepção coletiva do trabalho,

fazendo com que elementos como estabilidade e proteção sejam substituídos por flexibilidade e autossuficiência.

Segundo Löhr (2023, p. 2), a perspectiva de mudança num conceito X qualquer surge a partir da introdução de novos elementos na rede de conceitos relacionada a X. A mudança em uma rede conceitual começa com a introdução ou eliminação de conceitos nela ou por causa de novas relações inferenciais entre conceitos já existentes. Löhr observa que a mudança pode ocorrer mesmo quando alguns "conceitos individuais" da rede permanecem estáveis. Conforme o argumento de Löhr, pequenas mudanças na rede conceitual gradualmente causam perturbações na percepção da normatividade, resultando em uma erosão sistêmica e gradual. Segundo Löhr, nossas concepções e formas de aplicar um conceito variam entre pessoas e refletem muitas das nossas crenças, mudando conforme aprendemos, vivenciamos e conhecemos o mundo (Löhr, 2023, p. 2). Portanto, mudanças nas normas e práticas sociais alteram o conteúdo da rede conceitual, alterando assim o conceito mesmo (Löhr, 2023, p. 3). É possível ilustrar esse processo da seguinte maneira:

Figura 2: A opacidade algorítmica sistêmica introduz perturbações nas normas sociais, produzindo alterações na rede de conceitos do trabalho

Fonte: Elaborado pela autora (2024)

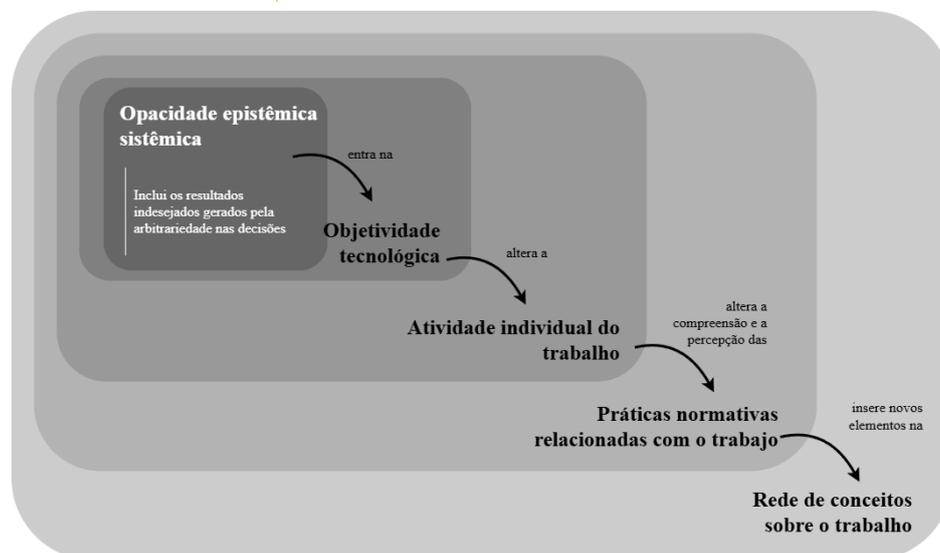
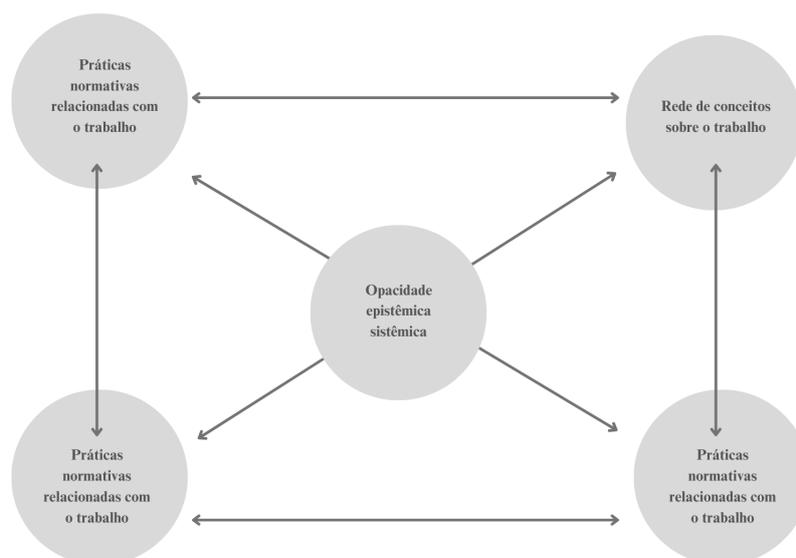


Figura 3: De maneira diferente da figura 2, essa imagem ilustra uma perspectiva de interconectividade, onde cada processo pode interferir no outro, contribuindo assim para a introdução de perturbações

Fonte: Elaborado pela autora (2024)



No entanto, se pensamos no contexto de não-neutralidade tecnológica, é razoável afirmar que esse processo pode, de fato, ser retroalimentado.² Pensar desde uma perspectiva de retroalimentação sugere que uma vez alterada, a rede de conceitos sobre o trabalho pode influenciar a forma como novos algoritmos são desenvolvidos e implementados. Essa perspectiva de continuidade permite repensar o processo descrito na figura 2 de uma maneira que permite contemplar também a visão de um sistema interconectado e não só uma sequência

linear. Essa perspectiva encontra-se ilustrada na figura 3 a continuação. É importante ressaltar que se pensada desde uma perspectiva de interconectividade, a relação entre a rede conceitual sobre o trabalho, as práticas normativas, a prática individual, a opacidade tecnológica relacionada com o trabalho e a opacidade algorítmica ressaltaria que, assim como os algoritmos influenciam as normas e percepções sobre o trabalho, as mudanças nessas normas e conceitos também podem redefinir a forma como os algoritmos são projetados e utilizados.

Nas duas perspectivas (linear ou de interconectividade), essa maneira de considerar como um desenvolvimento tecnológico pode alterar nossa concepção coletiva sobre um determinado conceito é interessante. No entanto, na filosofia da linguagem, não há consenso sobre como possuímos conceitos, especialmente os mais abstratos, como liberdade, justiça, poder, trabalho, etc. Autores como Cappelen (2016), Soysal (2020) e Löhr (2019) sugerem que a posse de um conceito está conectada ao ambiente social. Se isso for verdade, então deve haver um grau de disposição para ajustar crenças em conformidade com mudanças no contexto do ambiente (Soysal, 2020, p. 7).

É interessante a ideia de que Löhr defende a relação entre mudanças tecnológicas socialmente significativas e a mudança na rede conceitual (Löhr, 2023, p. 4), convidando-nos a pensar que os conceitos não possuem apenas uma condição de aplicação, mas também podem ter consequências de sua aplicação. Portanto, se aceitarmos que a tecnologia não é neutra, devemos estar atentos às consequências sociais e às mudanças nas consequências da aplicação, ou seja, à relação inferencial entre os conceitos (Löhr, 2023, p. 4) relacionados ao trabalho.

Considere o caso das plataformas digitais e a segregação imposta aos trabalhadores que sofrem as consequências de práticas de precarização laboral. O fato de que um grupo de pessoas precisa se submeter ao modelo de negócios imposto por essas empresas cria um dilema moral para a sociedade: permitiremos que algumas pessoas trabalhem à margem da justiça trabalhista sem seus direitos laborais garantidos? Segundo Danaher e Saetra (a publicar), um dos efeitos negativos da tecnologia é a imposição de uma pressão sobre as decisões, limitando-as. No caso concreto mencionado, essa flexibilização dos direitos representa uma mudança em nossos valores que gera novas permissões.

Nesse sentido, é possível questionar: O que acontecerá se interrompemos nossas práticas e

normas sociais que garantem direitos no trabalho, como o direito à não discriminação ou à segurança? O que acontecerá se adicionarmos um novo contexto tecnológico à realidade sociopolítica que possibilita aceitar o não cumprimento dos direitos no marco legal do trabalho? E será que isso já não está ocorrendo de forma sistêmica na dinâmica de trabalho promovida, por exemplo, pelo modelo de negócios utilizado pelas plataformas digitais? Essas questões não são fáceis de responder. Para complementar, não se pode deixar de comentar que Cristina Bicchieri (2005) faz uma consideração muito relevante sobre mudanças normativas. Ela supõe que mudanças normativas ocorrem não apenas por alteração das regras, mas também através da adaptação dos comportamentos às novas normas e às expectativas sociais em torno delas. Assim, a dinâmica de trabalho mediada por algoritmos pode não apenas transformar práticas, mas também redefinir nossas expectativas e normas sociais relacionadas ao trabalho.

4. Conclusão

A modo de conclusão, vale a pena reforçar que este trabalho tentou analisar como a opacidade algorítmica —ao distanciar trabalhadores dos processos decisórios e obscurecer as bases de decisões automatizadas— pode contribuir para transformar a própria noção de trabalho. A opacidade erige uma barreira entre trabalhadores e empresas e com isso, promove uma segregação epistêmica que ameaça a democracia no ambiente laboral ao promover injustiças. Como examinado ao longo do texto, o problema não é meramente técnico, mas estrutural, uma vez que a opacidade algorítmica altera as normas de trabalho e a forma como os trabalhadores exercem sua autonomia e entendem o papel do trabalho na esfera social.

Segundo a forma como este problema foi exposto aqui, entende-se que o enfrentamento desses desafios passa pela implementação de regulamentações de transparência algorítmica. Essa implementação pode espelhar-se em exemplos práticos das políticas. Nos Estados Unidos, a *New York City Local Law 144* (DCWP, s.d.) da cidade de Nova Iorque tem o objetivo de combater a discriminação no emprego através da proibição do uso de ferramentas automatizadas de decisão de emprego (*Automated Employment Decision Tools* —AEDTs) a menos que a ferramenta tenha sido submetida a uma auditoria de viés, que deve ser tornada pública, e que certos avisos sejam fornecidos aos candidatos ou funcionários. O objetivo dessa lei busca garantir que as ferramentas automatizadas usadas em processos de recrutamento não perpetuem ou ampliem vieses existentes e, portanto, exigem transparência por parte das empresas, permitindo que candidatos e funcionários avaliem a justiça da ferramenta. Na União Europeia, o *General Data Protection Regulation* (GDPR), especialmente em seu Artigo 22 (s.d.), já assegura o direito à explicação de decisões automatizadas. Este artigo visa proteger os indivíduos de decisões automatizadas discriminatórias. As auditorias podem ajudar a identificar e mitigar o viés nos algoritmos, garantindo que as decisões de contratação sejam justas e equitativas. Outro exemplo que podemos citar é o novo *AI Act* da União Europeia. No site “*EU AI Act: first regulation on artificial intelligence | Topics | European Parliament*” (*EU AI Act*, 2023) é possível encontrar a descrição da Lei de IA da União Europeia, a primeira lei abrangente sobre Inteligência Artificial (IA) do mundo. O objetivo principal da lei é garantir que os sistemas de IA utilizados na UE sejam seguros, transparentes, rastreáveis, não discriminatórios e ambientalmente amigáveis. Para atingir esse objetivo, a lei classifica os sistemas de IA em quatro níveis de risco: inaceitável, alto, limitado e mínimo, cada um com diferentes níveis

de regulamentação. Sistemas de IA considerados uma ameaça às pessoas são proibidos, enquanto sistemas de alto risco, como aqueles usados em dispositivos médicos ou aplicação da lei, exigem registro em um banco de dados da UE. A lei também exige transparência, como a rotulagem clara de conteúdo gerado por IA, e visa apoiar a inovação, fornecendo ambientes de teste para *startups* e pequenas e médias empresas. O Parlamento Europeu e o Conselho já aprovaram a lei, que entrará em vigor em diferentes etapas, com a proibição de sistemas de IA de risco inaceitável sendo aplicada primeiro e os sistemas de alto risco tendo o maior prazo para conformidade.

No entanto, enquanto este trabalho aborda a opacidade algorítmica e seus desafios estruturais, é importante também reconhecer as perspectivas otimistas sobre o impacto da IA no mercado de trabalho. Podemos ver um exemplo desse otimismo em textos como “*AI Could Actually Help Rebuild The Middle Class*” (Autor, 2024). Nele, argumenta-se que a IA, quando utilizada para aumentar as capacidades dos trabalhadores e democratizar o acesso a tarefas complexas, pode fortalecer a classe média. Essa visão —que não deixa de ser interessante e promissora— contrasta com as preocupações predominantes sobre a substituição de trabalhadores e sugere que a IA pode oferecer suporte ao trabalho humano ao invés de substituí-lo, particularmente ao possibilitar que profissionais menos qualificados executem tarefas mais complexas. Para que esse potencial se realize, contudo, é essencial garantir que o desenvolvimento de IA permaneça ancorado em políticas públicas robustas, como as defendidas pelo *AI Act* da União Europeia, e em uma estrutura ética que assegure que a IA sirva para empoderar e valorizar o trabalho humano, promovendo a justiça social.

Esse debate está alinhado à perspectiva que critica a especulação dos danos reais causados pela IA —tal como argumentam Kapoorayash

Kapoor e Narayananarvind Narayanan (2023). Eles identificam três categorias de danos reais: a dependência excessiva em LLMs para tomada de decisões críticas, que amplifica a desinformação; o impacto sutil da IA no trabalho, onde o poder é transferido dos trabalhadores para as empresas; e a exploração de trabalhadores que desempenham funções como filtragem de conteúdo tóxico, sem garantias de compensação justa ou direitos adequados. Em vez de uma moratória no desenvolvimento da IA, os autores defendem uma regulamentação prática e imediata que proteja os direitos dos trabalhadores e ofereça compensação justa aos criadores cujas produções servem de base para os modelos de IA. Compartilhando dessa perspectiva crítica, o presente trabalho reforça que o caminho para uma evolução tecnológica justa requer não apenas maior transparência, mas também um compromisso com o fortalecimento dos direitos trabalhistas e a criação de um ambiente de trabalho equitativo. Somente por meio de regulamentações claras e ações que promovam uma IA justa e ética será possível garantir que a tecnologia contribua para a dignidade e autonomia dos trabalhadores, em vez de perpetuar desigualdades e concentrar poder nas mãos de poucos.

Referências bibliográficas

- Art. 22 GDPR – Automated individual decision-making, including profiling. (s. f.). *General Data Protection Regulation (GDPR)*. Recuperado 7 de noviembre de 2024, de <https://gdpr-info.eu/art-22-gdpr/>
- Autor, D. H. (2024). *AI Could Actually Help Rebuild The Middle Class*. <https://www.noemamag.com/how-ai-could-help-rebuild-the-middle-class>
- Asquerino Lamparero, M. J. (2022, abril 30). Algoritmos y Discriminación [Site]. *Trabajo, Persona, Derecho Y Mercado Nuevos retos del mercado laboral: Pobreza en el trabajo, economía colaborativa y envejecimiento*.
- Bicchieri, C. (2005). *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press.
- Burrell, J. (2016). How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms. *Big Data and Society*, 3(1), 205395171562251.
- Cappelen, H. (2018). *Fixing Language an Essay on Conceptual Engineering*. Oxford University Press.
- Creel, K., & Hellman, D. (2022). The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision-Making Systems. *Canadian Journal of Philosophy*, 52(1), 26-43.
- Danaher, J., & Nyholm, S. (2021). Automation, Work and the Achievement Gap. *AI and Ethics*, 1(3), 227-237.
- Danaher, J., & Sætra, H. S. (forthcoming). Mechanisms of Techno-Moral Change: A Taxonomy and Overview. *Ethical Theory and Moral Practice*, 1-22.
- DCWP - *Automated Employment Decision Tools (AEDT)*. (s. d.). Recuperado 7 de noviembre de 2024, de <https://www.nyc.gov/site/dca/about/automated-employment-decision-tools.page>
- Dubal, Veena, On Algorithmic Wage Discrimination (January 19, 2023). UC San Francisco Research Paper No. Forthcoming.
- EU AI Act: *First regulation on artificial intelligence*. (2023, junio 8). Topics | European Parliament. <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>
- Ferreira, M. (2021). Inscrutable Processes: Algorithms, Agency, and Divisions of Deliberative Labour. *Journal of Applied Philosophy*, 38(4), 646-661.
- Geuss, R. (2021). *A Philosopher Looks at Work* (1ª). Cambridge University Press.
- González, J. R. O., & Teixidó Durán, Ó. F. (2023). Derechos, deberes, justicia y simetría moral desde una filosofía científica. *UNIVERSITAS. Revista de Filosofía, Derecho y Política*, 42, Article 42. <https://doi.org/10.20318/universitas.2023.7850>
- Hare, S. (2022). *Techonlogy is not neutral A short Guide to Technology Ethics*. London Published Partnership.
- Horgan, A. (2021). Creeping and Ameliorative Accounts of “Work”. *Theory & Event*, 24(4), 1110-1129.
- Infografía—Panorámica de las plataformas digitales en la UE. (2023, marzo 23). [Site]. *Consejo Europeo*.
- Joque, J. (s. f.). *Revolutionary Mathematics Artificial Intelligence, Statistics and the Logic of Capitalism* (2022.a ed.). Verso.
- Kapoorsayash Kapoor, S., & Narayananarvind Narayanan, Y. (2023, marzo 29). A misleading open letter about sci-fi AI dangers ignores the real risks [Site]. *Snake Oil*.

- Loi, M., Ferrario, A., & Viganò, E. (2021). Transparency as Design Publicity: Explaining and Justifying Inscrutable Algorithms. *Ethics and Information Technology*, 23(3), 253-263.
- Löhr, G. (2019). Embodied cognition and abstract concepts: Do concept empiricists leave anything out? *Philosophical Psychology*, 32(2), 161-185.
- Löhr, G. (2023). Do socially disruptive technologies really change our concepts or just our conceptions? *Technology in Society*, 72.
- Marx, K. (2010). *El Capital Libro Primero El proceso de Producción del Capital* (P. Scaron, Trad.). Siglo XXI editores.
- Mittelstadt, B., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data and Society*, 3(2).
- OIT. (2004, agosto 9). ¿Qué es el Trabajo Decente? [Site]. *Organización Internacional del Trabajo*.
- Páez, A., & Ramírez-Bustamante, N. (2022). Análisis jurídico de la discriminación algorítmica en los procesos de selección laboral. En N. Angel & R. Urueña (Eds.), *Derecho, poder, y datos: Aproximaciones críticas al derecho y las nuevas tecnologías*. Ediciones Uniandes.
- Pinheiro-Machado, R. (2024). Why and How Precarious Workers Support Neo-illiberalism. En M. Gabriel, A. Katsman, T. Liess, & W. Milberg (Eds.), *Beyond Neoliberalism and Neo-Illiberalism: Economic Policies and Performance for Sustainable Democracy*. The New—Transcript and Columbia University Press.
- Romero, G. E. (Director). (2024, julio 9). *Inteligencia artificial* [Video recording]. <https://www.youtube.com/watch?v=-DNqEOUbwFQ>
- Rueda, J. (2021). Tecnologías socialmente disruptivas (Introducción). *Dilemata*, 31(13), 5-9.
- Soysal, Z. (2021). From Metasemantics to Analyticity. *Philosophy and Phenomenological Research*, 103(1), 57-76.
- Stuart, M., Trappmann, V., Bessa, I., Joyce, S., Neumann, N., & Umney, C. (2023). Labor Unrest and the Future of Work: Global Struggles against Food Delivery Platforms. *Labor Studies Journal*.
- Suarez, T. (2023, junio 7). Le Dialogue Social pour les Travailleurs des Plateformes en France est de la Poudre aux Yeux. *Euractiv.fr*.
- Sullivan, E. (2022). How Values Shape the Machine Learning Opacity Problem. En I. Lawler, K. Khalifa, & E. Shech (Eds.), *Scientific Understanding and Representation* (pp. 306-322). Routledge.
- Toscano, J. (2022). But seriously: What do Algorithms Want? Implying Collective Intentionalities in Algorithmic Relays. A Distributed Cognition Approach. *Zagadnienia Filozoficzne W Nauce*, 73, 47-76.
- Tsamados, A., Aggarwal, N., Cows, J., Morley, J., Roberts, H., Taddeo, M., & Floridi, L. (2021). The ethics of algorithms: Key problems and solutions. *En AI and Society* (pp. 1-16).

Notas finais

1 A inclusão desse parágrafo na presente seção do artigo foi motivada por comentários de um revisor anônimo, a quem agradeço.

2 A inclusão desse parágrafo na presente seção do artigo foi motivada por comentários de um revisor anônimo, a quem agradeço.

ARTIGO

Inteligência artificial e desigualdade social: o impacto do colonialismo digital nas políticas públicas

Letícia Lé Oliveira

leticialeolv@gmail.com

Bacharel pela Faculdade de Direito
da Universidade de São Paulo e Pós-
Graduada em Direito Digital pela
Fundação Getúlio Vargas.

Inteligência artificial e desigualdade social: o impacto do colonialismo digital nas políticas públicas

Palavras-chave

Inteligência Artificial
Colonialismo Digital
Direitos Fundamentais
Políticas Públicas

Resumo

Este artigo tem como objetivo analisar a intersecção entre os direitos fundamentais e a inteligência artificial, com um foco particular no colonialismo digital e na proteção constitucional. Busca-se compreender como as práticas de vigilância e controle de dados por grandes empresas e eventualmente poderes estatais desafiam os princípios constitucionais e os direitos fundamentais, bem como explorar as respostas jurídicas e políticas necessárias para mitigar esses desafios.

Artificial intelligence and social inequality: the impact of digital colonialism on public Policy

Keywords

Artificial Intelligence
Digital Colonialism
Fundamental Rights
Public Policies

Abstract

This article aims to analyze the intersection between fundamental rights and artificial intelligence, with a particular focus on digital colonialism and constitutional protection. It seeks to understand how surveillance practices and data control by large corporations and eventually government powers challenge constitutional principles and fundamental rights, as well as to explore the legal and political responses needed to mitigate these challenges.

1 Introdução

O avanço acelerado da tecnologia, com especial destaque para o campo da inteligência artificial (IA)¹, tem gerado uma série de desafios e oportunidades que impactam diretamente a proteção dos direitos fundamentais. De acordo com Russel e Norwig (2021), entende-se que Inteligência Artificial é

“o campo de estudo e desenvolvimento de sistemas computacionais capazes de realizar tarefas que normalmente requerem inteligência humana, como reconhecimento de fala, tomada de decisão, tradução de idiomas e percepção visual. A IA pode ser dividida em subcampos específicos, incluindo aprendizado de máquina, processamento de linguagem natural, visão computacional, e robótica, entre outros”. (Russel & Norwig, 2021)

A crescente interação entre o mundo digital e o mundo real tem transformado as formas de socialização - econômica e política - demandando uma reavaliação dos conceitos tradicionais de privacidade, liberdade e controle. Nesse contexto, emerge a compreensão da categoria que é conhecida como colonialismo digital, definido por Cassino et al. (2022, p. 17) como uma nova forma de dependência que combina práticas predatórias do colonialismo histórico com métodos abstratos de quantificação computacional.

A proteção dos direitos fundamentais, tais como a privacidade e a dignidade da pessoa humana, enfrenta desafios significativos diante dessa nova realidade digital. A Constituição Federal de 1988, que consagra a dignidade da pessoa humana como um de seus fundamentos, deve também ser reinterpretada e aplicada de modo a proteger os cidadãos contra os abusos

potenciais do colonialismo digital. Neste sentido, a promulgação da Lei Geral de Proteção de Dados (LGPD), em 2018, representa um marco importante, estabelecendo diretrizes claras para a coleta, tratamento e armazenamento de dados pessoais². Todavia, a efetividade dessa legislação depende de sua implementação rigorosa e da adaptação contínua às rápidas mudanças tecnológicas.

A discussão aqui proposta será estruturada em quatro seções principais. Na primeira, serão abordados os conceitos de direitos fundamentais na era digital, destacando a evolução da proteção de dados e privacidade como direitos essenciais. Na segunda seção, será explorado o conceito de colonialismo digital e suas implicações, incluindo a análise do capitalismo de vigilância. A terceira seção discutirá os desafios jurídicos e tecnológicos impostos pelo colonialismo digital à proteção dos direitos fundamentais. Finalmente, a quarta seção apresentará propostas de políticas públicas e intervenções legislativas para proteger os direitos na era digital.

Diante da crescente importância da inteligência artificial e da digitalização na sociedade contemporânea, a proteção dos direitos fundamentais torna-se uma questão urgente e complexa.

A metodologia deste artigo adota uma abordagem mista que combina análise teórica e empírica para uma compreensão aprofundada do impacto da inteligência artificial (IA) e do colonialismo digital sobre direitos fundamentais. A análise teórica será fundamentada na literatura acadêmica existente sobre proteção de dados, discriminação algorítmica e vigilância digital, integrando conceitos e debates presentes nas áreas de direito e políticas públicas. Serão incorporados dados empíricos sobre violações de privacidade e casos de discriminação algorítmica, extraídos de relatórios, estudos de ONGs e instituições de pesquisa, bem como um estudo comparativo entre diferentes países que implementaram regulamentações de proteção de dados.

Desse modo, este estudo pretende contribuir para o debate acadêmico e jurídico sobre como o Direito pode responder de forma eficaz a esses novos desafios, promovendo um desenvolvimento tecnológico que respeite e proteja os direitos humanos.

1. A inclusão da proteção de dados e privacidade como direitos fundamentais.

A proteção de dados pessoais emergiu como uma questão central no século XXI, impulsionada pelo crescimento exponencial da coleta, armazenamento e processamento de dados pelas empresas e governos. Este fenômeno, frequentemente descrito como a "era dos dados", trouxe à tona preocupações significativas sobre privacidade, segurança e autonomia individual. Mayer-Schönberger e Cukier (2013) destacam que a capacidade de coletar e analisar grandes volumes de dados transformou diversos aspectos da sociedade, desde negócios até políticas públicas.

O reconhecimento da proteção de dados como um direito fundamental tem implicações profundas para a estruturação das interações sociais e econômicas. A legislação de proteção de dados, como a GDPR e a LGPD, busca criar um equilíbrio entre a necessidade de inovação tecnológica e a proteção dos direitos individuais. Ao garantir que os dados pessoais sejam tratados de maneira ética e transparente, essas leis não apenas protegem a privacidade, mas também promovem um ambiente de confiança que é essencial para o desenvolvimento econômico e social.

A privacidade, tradicionalmente entendida como o direito de estar só, passou por uma transformação significativa com o advento da era digital. A definição clássica de privacidade,

conhecida como "o direito de ser deixado em paz", articulada por Warren and Brandeis (1890) e baseada na ideia de que a privacidade é um direito essencial à dignidade e ao bem-estar individual, sendo fundamental para proteger a integridade moral e psicológica das pessoas, não é mais suficiente para abordar as complexidades introduzidas pela tecnologia moderna. Na era digital, a privacidade envolve não apenas a proteção contra intrusões físicas, mas também a proteção contra a coleta e uso indevido de informações pessoais.

Zuboff (2019) argumenta que a privacidade deve ser contextualizada em um mundo onde os dados pessoais são constantemente coletados, analisados e monetizados por empresas de tecnologia. Ela introduz o conceito de "capitalismo de vigilância", no qual o comportamento humano é transformado em dados que são utilizados para prever e influenciar futuras ações. Zuboff (2019) destaca ainda que essa prática cria formas de poder e controle, onde as empresas detêm vastas quantidades de informações sobre os indivíduos, muitas vezes sem o seu conhecimento ou consentimento. Este novo paradigma exige uma reavaliação das normas jurídicas de privacidade para proteger os indivíduos contra a exploração e manipulação de seus dados pessoais.

A proteção de dados e a privacidade não são apenas elementos essenciais para a dignidade e a liberdade dos indivíduos, mas também fundamentais para a autonomia pessoal. Segundo Rodotà et al. (2008), a proteção de dados é uma condição indispensável para a realização da dignidade humana na sociedade contemporânea. Este reconhecimento da proteção de dados como um direito fundamental tem implicações profundas e necessárias para a estruturação das interações sociais e econômicas. Ao garantir que os dados pessoais sejam tratados de maneira ética e transparente, essas leis não apenas protegem a privacidade, mas também promovem um ambiente de confiança

que é essencial para o desenvolvimento econômico e social.

Neste sentido, a dignidade humana, como princípio fundamental, está intrinsecamente ligada à proteção da privacidade e dos dados pessoais. A vigilância e a coleta indiscriminada de dados podem levar à manipulação e ao controle das escolhas individuais, comprometendo a liberdade e a autonomia. Portanto, proteger os dados pessoais é proteger a capacidade dos indivíduos de tomar decisões livres e informadas sobre suas próprias vidas.

A implementação de regulamentações como a GDPR e a LGPD apresenta tanto desafios quanto oportunidades no campo dos direitos fundamentais. Por um lado, essas leis impõem novas responsabilidades e custos às organizações que coletam e processam dados pessoais. As empresas devem investir em sistemas de compliance, treinamento de funcionários e infraestrutura tecnológica para assegurar a conformidade com as normas de proteção de dados. Além disso, a aplicação extraterritorial dessas leis, como é o caso da GDPR, cria complexidades adicionais para as organizações que operam em múltiplas jurisdições.

Os desafios também incluem a necessidade de adaptar rapidamente as regulamentações às novas tecnologias. A Inteligência Artificial (IA), a Internet das Coisas (IoT) e o big data trazem novos riscos e oportunidades para a proteção de dados. A IA, por exemplo, tem sido usada para analisar grandes volumes de dados e fazer previsões ou decisões automatizadas, o que levanta questões significativas sobre transparência, responsabilidade e vieses. Segundo Crawford (2021), os sistemas de IA operam frequentemente com opacidade, dificultando a compreensão de seus critérios de decisão e introduzindo vieses que refletem as limitações dos dados usados para treinamento, o que pode resultar em discriminação e injustiça algorítmica (Crawford, 2021, p. 215). Complementarmente, Obermeyer et al. (2019)

destacam que algoritmos de IA, especialmente em setores críticos como saúde e finanças, têm reproduzido preconceitos presentes nos dados históricos, exacerbando desigualdades e afetando desproporcionalmente grupos vulneráveis (Obermeyer et. al, 2019, p. 447-453).

Já a Internet das Coisas (IoT), com dispositivos conectados que coletam dados em tempo real, aumenta exponencialmente a quantidade de dados disponíveis e o risco de violações de privacidade. De acordo com Zuboff (2019), a coleta contínua de dados por dispositivos IoT alimenta o que ela chama de “capitalismo de vigilância,” onde a privacidade dos indivíduos é comprometida para sustentar modelos de negócio baseados na exploração de dados pessoais (Zuboff, 2019, p. 340). Esses dispositivos, como observa Greengard (2020), geram novos desafios de segurança e privacidade, pois muitas vezes carecem de proteções robustas contra invasões, colocando os dados dos usuários em constante vulnerabilidade (Greengard, 2020, p. 56).

Por outro lado, essas regulamentações oferecem uma oportunidade para fortalecer a confiança dos consumidores e promover práticas comerciais mais éticas e transparentes. A proteção de dados eficaz pode ser um diferencial competitivo para as empresas, atraindo clientes que valorizam a privacidade e a segurança de suas informações pessoais. Além disso e principalmente: a proteção de dados contribui para a construção de um ambiente digital mais seguro e confiável, incentivando a inovação e o desenvolvimento econômico sustentável.

Um dos principais desafios na proteção de dados é garantir a conformidade em um ambiente globalizado, onde os dados frequentemente atravessam fronteiras nacionais. A cooperação internacional é essencial para abordar questões transfronteiriças de privacidade e segurança de dados. A criação de frameworks internacionais e a harmonização de normas de proteção de dados podem facilitar a cooperação e a coordenação entre diferentes jurisdições,

promovendo uma abordagem mais coerente e eficaz para a proteção de dados.

A literatura destaca que a proteção de dados e a privacidade não são apenas questões de conformidade regulatória, mas também de ética e responsabilidade social. Autores como Nissenbaum (2009), defendem que a privacidade deve ser vista como um bem social que contribui para a integridade e coesão das comunidades. Nissenbaum (2009) argumenta que a privacidade contextualizada, onde os dados são tratados de acordo com as normas sociais e expectativas do contexto específico, é fundamental para garantir que a proteção de dados seja eficaz e respeite os valores sociais.

Em uma análise comparativa, é possível observar que a implementação de legislações de proteção de dados na América Latina é uma área que tem visto avanços significativos nas últimas décadas, refletindo a crescente conscientização sobre a importância da privacidade e segurança dos dados. No entanto, cada país da região tem adotado abordagens distintas, influenciadas por contextos políticos, econômicos e sociais únicos, além de enfrentar o desafio adicional do colonialismo digital.

Na Colômbia, a proteção de dados pessoais é garantida pela Constituição Política do país, que reconhece o direito de cada cidadão de conhecer, atualizar, retificar e cancelar suas informações pessoais. A abordagem colombiana enfatiza a transparência e o consentimento do usuário, refletindo um compromisso robusto com a proteção de dados (Velooso Meireles, 2023). No entanto, a implementação prática dessas diretrizes ainda encontra obstáculos, especialmente em áreas com menor acesso à tecnologia e informação.

O Chile promulgou a Lei de Proteção de Dados Pessoais (LFPD) em 1999, com revisões importantes em 2018. A LFPD proíbe o tratamento inadequado de dados sensíveis e impõe sanções para violações. Embora o Chile tenha uma estrutura legal para a proteção de dados,

a lei ainda carece de um órgão regulador independente, o que limita sua eficácia e a confiança do público (Veronese et al., 2023).

O México implementou a Lei Federal de Proteção de Dados Pessoais (LFPDPPP) em 2010, atualizada em 2018, que estabelece princípios como minimização de dados e responsabilidade. A LFPDPPP exige consentimento explícito do usuário antes da coleta e processamento de dados (Baptista Luz Advogados et al., 2022). No entanto, a aplicação prática da legislação ainda enfrenta obstáculos, especialmente em áreas rurais e entre comunidades indígenas, conforme relatado pela Asociación por los Derechos Civiles (ADC, 2021). Essas desigualdades geográficas e econômicas limitam o alcance da lei e permitem que empresas internacionais se aproveitem de lacunas regulatórias para coleta de dados, reforçando uma dependência que reflete o colonialismo digital.

No contexto do colonialismo digital, a América Latina enfrenta a dominação de grandes corporações e países desenvolvidos sobre suas infraestruturas digitais e dados. Essa dinâmica de poder pode resultar na exploração dos dados dos cidadãos latino-americanos, restringindo a autonomia digital da região. Empresas de tecnologia globalmente dominantes frequentemente controlam vastos volumes de dados, o que levanta preocupações sobre soberania digital e proteção adequada dos dados dos cidadãos (Velooso Meireles, 2023).

Deste modo, a complexidade crescente das interações digitais e a expansão da coleta de dados em escala global evidenciam que a privacidade e a proteção de dados transcendem o domínio do controle individual sobre informações pessoais e se inserem profundamente em questões éticas, sociais e políticas. Em um ambiente digital marcado pela interdependência global e por novas formas de poder corporativo, a proteção de dados assume o papel de um contrapeso indispensável à expansão do “capitalismo de vigilância” e à exploração

econômica dos dados pessoais. Nesse contexto, mais do que salvaguardar a dignidade e a liberdade dos indivíduos, as regulamentações como a GDPR e a LGPD representam uma tentativa de estabelecer novos paradigmas de justiça e responsabilidade na era digital.

2. Impacto da digitalização e da IA nos Direitos Fundamentais

A digitalização e a IA revolucionaram a forma como os dados pessoais são coletados, armazenados e processados. A digitalização, aqui, trata-se da "transformação de representações contínuas e analógicas (como som, imagem, texto e formas físicas) em dados digitais que podem ser armazenados, processados e transmitidos em redes computacionais" (Brennen & Kreiss, 2014). Esse processo de digitalização inclui não só a criação de representações digitais de objetos e informações físicas, mas também a reestruturação de práticas sociais, econômicas e culturais para operar dentro do contexto digital.

Estas tecnologias e transformações tecnológicas modificaram radicalmente também a forma como a informação é produzida, distribuída e consumida. Plataformas digitais e redes sociais utilizam aplicações de IA customizando os algoritmos de modo a personalizar o conteúdo que os usuários veem, com base em seus comportamentos e preferências passadas. Embora isso possa aumentar o engajamento e a satisfação dos usuários, também pode criar "bolhas de filtro" e câmaras de eco, limitando a exposição a diferentes pontos de vista e polarizando o discurso público. As "bolhas de filtro", aqui citadas, são, como explica Pariser (2011, p. 9) "seu próprio universo pessoal e único de informações em que você vive online. E o que está na sua bolha de filtro depende de quem você é

e do que você faz. Mas você não decide o que entra – e, mais importante, você não vê o que é filtrado." Ou seja, surgem da personalização feita por algoritmos de recomendação, como os usados em redes sociais, mecanismos de busca e plataformas de conteúdo. Já as câmaras de eco, conforme conceituadas por Sunstein (2018, p. 10), trata-se de "uma situação em que muitas pessoas ouvem ecos de suas próprias vozes, porque estão escutando pessoas com ideias afins ou com pontos de vista semelhantes (...)".

Além disso, o uso de IA para moderar conteúdo nas plataformas digitais levanta preocupações sobre censura e liberdade de expressão. Crawford (2021) explica que os sistemas de IA usados na moderação são, por vezes, "opacos e tendenciosos," porque operam sob algoritmos que refletem preconceitos embutidos, resultando em uma censura inadvertida de conteúdos de minorias ou de temas sensíveis. Ela menciona: "A censura algorítmica frequentemente opera sob uma lógica binária, onde nuances de expressão e contexto cultural são sacrificados em prol da padronização e da obediência a diretrizes muitas vezes inflexíveis" (Crawford, 2021, p. 215). Ademais, Douek (2020), professora e pesquisadora de liberdade de expressão, aponta que a IA aplicada à moderação automatizada frequentemente toma decisões de remoção de conteúdo com base em categorias binárias e inconsistentes com a complexidade da linguagem e do contexto humano. Ela argumenta que "a tecnologia de IA carece da capacidade de contextualizar o discurso, o que frequentemente resulta em uma aplicação desproporcional de normas de censura" (Douek, 2020, p. 231).

Assim, conteúdos que seriam permitidos no contexto humano podem ser censurados erroneamente pela IA. Algoritmos de moderação de conteúdo podem remover injustamente postagens e perfis, muitas vezes sem explicação ou recurso adequado. Isso pode resultar na supressão de vozes dissidentes e na limitação

do debate democrático, comprometendo o direito à liberdade de expressão protegido pela Constituição e por tratados internacionais de direitos humanos.

No trabalho de Wu (2017) se explora como as plataformas digitais competem pela atenção dos usuários, muitas vezes manipulando informações e criando ambientes propícios à desinformação e à manipulação. Para proteger a liberdade de expressão na era digital, é essencial equilibrar a moderação de conteúdo com a transparência e a responsabilidade das plataformas, além de garantir que os usuários tenham acesso a um processo justo de apelação contra decisões automatizadas.

A digitalização e a IA também oferecem oportunidades para melhorar o acesso à justiça, mas apresentam desafios que precisam ser cuidadosamente gerenciados. Ferramentas de IA podem ser usadas para automatizar tarefas jurídicas, como a revisão de documentos e a previsão de resultados judiciais, aumentando a eficiência e reduzindo os custos legais. Isso pode tornar os serviços jurídicos mais acessíveis para indivíduos e pequenos negócios que de outra forma não poderiam arcar com os custos.

No entanto, a automação do processo judicial também levanta questões sobre transparência, responsabilidade e justiça. Algoritmos de IA utilizados em decisões judiciais e administrativas devem ser transparentes e passíveis de auditoria para garantir que suas decisões sejam justas e imparciais. Além disso, é crucial garantir que os indivíduos afetados por essas decisões automatizadas tenham o direito de contestá-las e buscar reparação.

Yeung (2017) destaca a necessidade de um quadro regulatório robusto para governar o uso de IA no sistema judicial. Ela argumenta que a regulação algorítmica deve ser orientada por princípios de justiça procedimental, transparência e responsabilidade, a fim de proteger os direitos fundamentais e garantir a equidade nas decisões automatizadas.

Ademais, a IA tem o potencial de perpetuar e amplificar discriminações existentes na sociedade. A discriminação algorítmica é uma consequência direta do colonialismo digital, afetando desproporcionalmente os grupos vulneráveis. Algoritmos treinados em conjuntos de dados históricos tendem a replicar os vieses e preconceitos presentes nesses dados. Quando esses algoritmos são usados para tomar decisões em áreas críticas, como emprego, crédito, habitação e justiça criminal, eles podem replicar e até exacerbar esses preconceitos.

O'Neil (2016) discute como algoritmos de IA podem ser "armas de destruição matemática", criando ciclos de discriminação e desigualdade. Ela destaca casos em que algoritmos de pontuação de crédito, predição de crimes e contratação de funcionários resultaram em discriminação contra minorias e grupos marginalizados. "Os modelos matemáticos podem perpetuar a discriminação existente, pois são treinados com dados históricos que refletem essas desigualdades" (O'Neil, 2016, p. 45).

Essas preocupações são ecoadas por Eubanks (2018), onde ela examina como os sistemas automatizados utilizados para fornecer serviços públicos podem reforçar a exclusão social. Eubanks argumenta que "os algoritmos que determinam quem recebe assistência social, quem é elegível para moradia pública e quem é monitorado pelo sistema de justiça criminal muitas vezes refletem e ampliam preconceitos e desigualdades sociais existentes" (Eubanks, 2018, p. 29).

À medida que a digitalização e a IA reconfiguram o espaço dos direitos fundamentais, emergem, portanto, questões profundas sobre a interseção entre controle automatizado e dignidade humana. Essas tecnologias, ao mesmo tempo em que facilitam o acesso à informação e a justiça, também introduzem um novo paradigma: a de uma sociedade que, inadvertidamente, delega suas decisões e julgamentos a sistemas opacos e impessoais. Se a digitalização, na

busca por eficiência, não contemplar o necessário respeito ao pluralismo e à autonomia individual, ela corre o risco de substituir os valores democráticos pela tirania algorítmica. Emerge, assim, a necessidade de um amplo debate sobre o tema, abarcando as mais diversas áreas dos setores público e privado.

3. Colonialismo Digital e Capitalismo de Vigilância

Couldry e Mejias (2019, p.1) definem o colonialismo de dados como uma nova forma de colonialismo, onde as práticas extrativistas do período colonial histórico são replicadas através da captura de dados em massa. Eles afirmam, ainda, que "o colonialismo de dados combina as práticas extrativas predatórias do colonialismo histórico com os métodos abstratos de quantificação da computação, transformando a vida humana em dados que podem ser apropriados para o capitalismo".

O colonialismo digital é um conceito, portanto, utilizado para descrever as dinâmicas de poder e controle exercidas por empresas de tecnologia e Estados sobre os dados, informações e infraestruturas digitais de outros países ou regiões, especialmente aquelas em desenvolvimento. Este conceito é empregado para analisar como as práticas de coleta, processamento e monetização de dados podem perpetuar desigualdades econômicas, sociais e políticas, replicando padrões históricos de exploração e dominação colonial.

Em sua análise, Faustino e Lippold (2023) destacam que o colonialismo digital não é apenas uma metáfora, mas uma dominação concreta que envolve a "subordinação econômica, política, social e racial de determinados territórios" por meio de tecnologias digitais. Eles afirmam que esse modelo de dominação reproduz as lógicas coloniais, agora reconfiguradas pela

dependência tecnológica, onde as infraestruturas e o acesso a dados são controlados majoritariamente pelos países do Norte Global, como Estados Unidos.

Uma das características centrais do colonialismo digital é a centralização do poder nas mãos de poucas corporações tecnológicas. Essas empresas possuem não apenas os recursos financeiros, mas também a infraestrutura e os dados necessários para influenciar comportamentos e decisões em uma escala global. O controle sobre os dados permite que essas corporações direcionem a inovação tecnológica, moldem políticas públicas e influenciem economias inteiras.

Assim, no contexto do colonialismo digital, a assimetria de poder entre as corporações tecnológicas e os usuários individuais se torna evidente. Os dados pessoais, coletados muitas vezes sem consentimento informado, são transformados em commodities valiosas. Os usuários, por outro lado, frequentemente desconhecem a extensão da coleta de dados e a forma como seus dados são utilizados. Essa falta de transparência e de consentimento informado perpetua uma dinâmica de exploração, onde os benefícios econômicos são concentrados nas mãos de poucos, enquanto os riscos e as vulnerabilidades são distribuídos entre muitos.

O colonialismo digital também cria dinâmicas de dependência tecnológica e econômica. Países em desenvolvimento, sem a infraestrutura ou os recursos para competir com gigantes tecnológicas, tornam-se dependentes das tecnologias e serviços oferecidos por essas corporações. Essa dependência limita a capacidade desses países de desenvolverem suas próprias soluções tecnológicas e econômicas, perpetuando um ciclo de subordinação e exploração. Empresas e governos de países em desenvolvimento são frequentemente obrigados a confiar nesses serviços para suas operações diárias, criando vulnerabilidades significativas. A centralização dos dados em infraestruturas controladas por corporações

estrangeiras também levanta questões de soberania e segurança nacional.

Outra forma de enxergar o colonialismo digital é a partir da compreensão de que ele é uma forma contemporânea de extrativismo, onde os dados pessoais são explorados como recursos naturais. Couldry e Mejias (2020) afirmam que "a exploração de dados é uma forma contemporânea de extrativismo, onde os dados pessoais são coletados e transformados em mercadorias exploráveis, perpetuando desigualdades globais" (Couldry & Mejias, 2019, p.2). A coleta massiva de dados é realizada sem a devida compensação para os indivíduos ou comunidades de onde esses dados são extraídos. Os benefícios econômicos dessa exploração são concentrados nas mãos das corporações tecnológicas, enquanto os custos e os riscos são externalizados para os indivíduos e as sociedades.

Essa dinâmica de extrativismo de dados reflete práticas históricas de colonialismo, onde recursos naturais eram extraídos de regiões colonizadas para o benefício econômico dos colonizadores. No colonialismo digital, os dados pessoais substituem os recursos naturais, mas as dinâmicas de poder e exploração permanecem semelhantes. Os dados são coletados e processados de maneiras que maximizam o lucro para as corporações, muitas vezes à custa da privacidade e da autonomia dos indivíduos.

A prática de coleta e monetização de dados pessoais por empresas tecnológicas globais em países africanos, como Gana e Nigéria, exemplifica o fenômeno conhecido como colonialismo digital. Conforme destacado pela Privacy International (n.d), essas empresas frequentemente utilizam dados dos cidadãos africanos para fins de publicidade direcionada e personalização de serviços, muitas vezes sem o consentimento claro dos usuários.

Essa prática levanta sérios questionamentos sobre a soberania digital, uma vez que os dados produzidos dentro desses países são explorados economicamente por corporações estrangeiras,

sem que o valor gerado retorne para a economia local. Cidadãos ganeses e nigerianos, por exemplo, geram grandes volumes de dados ao usar aplicativos e serviços móveis, mas, em muitos casos, esses dados são monetizados por empresas estrangeiras sem transparência. Como consequência, países africanos perdem o potencial de desenvolver suas próprias economias digitais, ficando tecnologicamente dependentes de soluções e infraestruturas externas (Stevenson, 2024).

A Privacy International (2020) destaca que a falta de regulamentações robustas de proteção de dados em muitos países africanos permite que grandes empresas de tecnologia operem praticamente sem supervisão. Diferente da Europa, que conta com a GDPR para regulamentar e limitar o uso de dados pessoais, muitos países africanos ainda não dispõem de marcos legais que protejam os dados de seus cidadãos. Essa lacuna regulatória facilita práticas que perpetuam relações econômicas e tecnológicas desiguais, típicas do colonialismo, onde os recursos – no caso, dados pessoais – são extraídos e explorados por empresas do Norte Global.

O conceito de capitalismo de vigilância, por sua vez, dialoga diretamente com a lógica do colonialismo digital. Em seu trabalho, Zuboff (2019) define o "capitalismo de vigilância" como a prática de coleta massiva de dados pessoais por grandes corporações para prever e influenciar comportamentos futuros, transformando a vigilância em um modelo de negócios lucrativo. Zuboff (2019, p. 15) explica que "o capitalismo de vigilância monetiza a experiência humana transformando-a em dados comportamentais que podem ser vendidos e comprados, criando novas formas de poder e controle".

O capitalismo de vigilância representa uma ameaça significativa à privacidade e à autonomia individual. A coleta constante e a análise de dados pessoais permitem que as empresas construam perfis detalhados dos indivíduos, frequentemente sem o seu consentimento

explícito. Esses perfis são utilizados para prever e influenciar comportamentos, comprometendo a capacidade dos indivíduos de tomar decisões livres e informadas. Zuboff (2019, p. 15) destaca que "a vigilância constante cria um ambiente onde os indivíduos são manipulados sem seu conhecimento, comprometendo a liberdade e a dignidade humanas".

A capacidade das empresas de influenciar o comportamento individual também tem implicações significativas para a democracia. A personalização de conteúdo nas plataformas digitais limita a exposição a diferentes pontos de vista e polarizando o discurso público, conforme exposto anteriormente. Além disso, a manipulação da informação pode ser utilizada para influenciar processos eleitorais e minar a confiança nas instituições democráticas.

As práticas de colonialismo digital e capitalismo de vigilância exemplificam, portanto, em escala internacional, como as dinâmicas de poder e exploração históricas podem ser replicadas e amplificadas na era digital. Estas dinâmicas impõem desafios práticos na proteção de dados e na governança da internet.

3.1. Desafios do Colonialismo Digital para a Proteção de Dados

A era digital trouxe consigo uma série de avanços tecnológicos que revolucionaram a maneira como a sociedade opera, interage e consome informações. No entanto, essa transformação também gerou desafios significativos para a proteção dos direitos fundamentais, especialmente no contexto do colonialismo digital. O controle e a exploração de dados pessoais por grandes corporações tecnológicas criam barreiras jurídicas e tecnológicas que dificultam a defesa desses direitos.

A primeira barreira jurídica significativa é a inadequação das legislações existentes. As leis de proteção de dados, como a GDPR (General Data Protection Regulation) na Europa e a LGPD (Lei Geral de Proteção de Dados) no Brasil, foram desenvolvidas para fornecer um marco regulatório robusto para a proteção dos dados pessoais. No entanto, a rápida evolução das tecnologias digitais muitas vezes supera a capacidade dessas leis de oferecer proteção adequada. Como destaca Zuboff (2019, p. 156) em seu trabalho, a velocidade com que as tecnologias avançam frequentemente excede a capacidade das legislações de se adaptar, criando lacunas que podem ser exploradas pelas corporações tecnológicas.

A GDPR, por exemplo, estabelece princípios fundamentais de transparência e consentimento, mas a implementação prática desses princípios enfrenta desafios significativos. A complexidade dos mecanismos de consentimento e a opacidade das práticas de coleta de dados por grandes corporações dificultam a aplicação desses princípios de maneira eficaz. Na visão de Veale e Binns (2017, p. 10), a exigência de consentimento informado e explícito é muitas vezes comprometida pela falta de clareza e pela complexidade dos termos de serviço apresentados aos usuários.

Além disso, a falta de harmonização global das leis de proteção de dados cria um ambiente desigual onde corporações tecnológicas podem explorar jurisdições com regulamentações mais fracas. Enquanto a GDPR e a LGPD estabelecem padrões elevados de proteção, muitos países ainda não possuem legislações robustas ou mecanismos de aplicação eficazes. Esta disparidade permite que as corporações movam suas operações para regiões com menos regulamentação, exacerbando a exploração de dados. Nissenbaum (2011) argumenta em seu trabalho que a fragmentação das leis de proteção de dados cria uma 'terra de ninguém' regulatória onde os direitos dos indivíduos são vulneráveis à exploração.

Outro desafio jurídico é a dificuldade de responsabilizar grandes corporações tecnológicas. Empresas como Google, Facebook e Amazon possuem recursos jurídicos consideráveis e operam em várias jurisdições, complicando a aplicação das leis nacionais. A natureza transnacional da internet permite que os dados frequentemente cruzem fronteiras nacionais, escapando à jurisdição de uma única entidade reguladora. Como observam De Hert e Papakonstantinou (2016), a globalização dos fluxos de dados requer uma abordagem internacional coordenada para a regulamentação, o que muitas vezes é difícil de alcançar devido às diferenças culturais e políticas entre os países.

A resistência das corporações tecnológicas à transparência e à responsabilização também representa uma barreira significativa. Muitos algoritmos e práticas de coleta de dados são protegidos como segredos comerciais, o que impede a fiscalização e a responsabilização por violações de direitos. Citando Pasquale (2015, p. 146), "a 'caixa preta' dos algoritmos corporativos oculta os processos de tomada de decisão, dificultando a identificação e a correção de práticas discriminatórias ou invasivas".

As barreiras tecnológicas incluem a complexidade dos sistemas de coleta e processamento de dados, que são frequentemente opacos e incompreensíveis para a maioria dos usuários. A falta de transparência nos algoritmos utilizados para processar dados pessoais é uma preocupação significativa. Esses algoritmos, muitas vezes protegidos por segredos comerciais, são complexos demais para serem auditados ou compreendidos por pessoas comuns. Esta opacidade impede a responsabilização e dificulta a proteção dos direitos dos indivíduos.

A centralização das infraestruturas de dados nas mãos de poucas corporações tecnológicas representa outra barreira tecnológica crítica. Essas corporações possuem vastos recursos para desenvolver tecnologias avançadas de coleta e análise de dados, criando uma disparidade de

poder tecnológico entre elas e os usuários individuais ou governos de países em desenvolvimento. Esta centralização de poder tecnológico não apenas limita a capacidade desses países de desenvolver suas próprias soluções tecnológicas e econômicas, mas também os torna dependentes das tecnologias oferecidas pelas grandes corporações. Isso perpetua uma dinâmica de dependência e subordinação, onde os países em desenvolvimento não têm autonomia para controlar suas próprias infraestruturas digitais.

A infraestrutura tecnológica centralizada permite que grandes corporações tecnológicas exerçam um controle significativo sobre os fluxos de dados globais. Essa centralização não só facilita a coleta e análise massiva de dados pessoais, mas também cria um ponto único de vulnerabilidade, onde violações de dados ou ataques cibernéticos podem ter impactos devastadores. Schneier (2015) argumenta, dentre outros temas, que a centralização dos dados em grandes plataformas tecnológicas cria alvos para atores maliciosos, aumentando os riscos de segurança para os indivíduos e as sociedades.

A complexidade técnica dos sistemas de IA e aprendizado de máquina utilizados pelas corporações tecnológicas também representa uma barreira significativa para a proteção dos direitos. Esses sistemas frequentemente operam como "caixas pretas", onde os processos de tomada de decisão são opacos e difíceis de interpretar. Isso dificulta a identificação e a correção de práticas discriminatórias ou invasivas. Burrell (2016) observa em seu trabalho que a opacidade dos sistemas algorítmicos cria um 'déficit de responsabilidade', onde as decisões automatizadas que afetam profundamente a vida dos indivíduos são tomadas sem a devida transparência ou supervisão.

A falta de alfabetização digital entre os usuários também contribui para a vulnerabilidade aos abusos de dados. Muitos indivíduos não possuem o conhecimento técnico necessário para compreender as implicações das práticas

de coleta de dados ou para proteger adequadamente sua privacidade. Esta falta de compreensão impede que os usuários tomem decisões informadas sobre o uso de seus dados pessoais. Citando Selwyn (2004, p. 352), "a alfabetização digital é essencial para a participação plena na sociedade da informação, mas a lacuna de conhecimento técnico entre os usuários e as corporações tecnológicas perpetua uma dinâmica de exploração e vulnerabilidade".

Desse modo, a governança da proteção de dados na era digital enfrenta desafios significativos devido à natureza global e interconectada da internet. A criação de um quadro regulatório eficaz requer a cooperação internacional e a harmonização das leis de proteção de dados. No entanto, diferenças culturais, políticas e econômicas entre os países dificultam essa harmonização. Como observa Bennett (2011) em sua obra, a governança global da privacidade requer um equilíbrio delicado entre a soberania nacional e a necessidade de normas internacionais consistentes.

4. Políticas Públicas para Proteger Direitos na Era Digital

Diante dos desafios e barreiras impostos a partir do colonialismo digital e do capitalismo de vigilância, políticas públicas robustas são essenciais para garantir que os direitos dos cidadãos sejam protegidos contra abusos e explorações no ambiente digital. Essas políticas devem abordar a privacidade, a segurança de dados, a igualdade de acesso e a transparência.

Uma das principais áreas de foco das políticas públicas na era digital é a privacidade e a proteção de dados. Governos em todo o mundo têm implementado legislações para regular a coleta, o armazenamento e o uso de dados

pessoais por empresas e entidades governamentais. A GDPR na União Europeia e a LGPD no Brasil são exemplos de tais legislações que estabelecem diretrizes claras para o tratamento de dados pessoais e garantem os direitos dos titulares dos dados. Essas leis exigem que as organizações obtenham consentimento explícito dos indivíduos antes de coletar seus dados, informando-os sobre como esses dados serão usados e armazenados. Além disso, proporcionam aos indivíduos o direito de acessar, corrigir e excluir suas informações pessoais.

Outra dimensão crítica das políticas públicas na era digital é a segurança de dados. As violações de dados e os ataques cibernéticos representam ameaças significativas à segurança nacional, à economia e aos direitos individuais. Os governos devem investir em tecnologias avançadas de segurança e promover a cooperação entre o setor público e o privado para combater ameaças cibernéticas. A criação de centros nacionais de resposta a incidentes cibernéticos (CSIRTs) e a implementação de programas de conscientização sobre segurança cibernética são exemplos de iniciativas políticas que podem fortalecer a resiliência contra ataques cibernéticos.

As políticas públicas também devem abordar com centralidade a questão da desigualdade de acesso à tecnologia e à internet. Políticas de inclusão digital são essenciais para garantir que todos os cidadãos tenham acesso equitativo à internet e às tecnologias digitais. Essas políticas podem incluir, por exemplo, a expansão da infraestrutura de banda larga para áreas rurais e remotas, a oferta de subsídios para dispositivos tecnológicos para famílias de baixa renda e a promoção da alfabetização digital. A inclusão digital também envolve o desenvolvimento de conteúdos e serviços acessíveis a pessoas com deficiência, garantindo que todos os cidadãos possam participar plenamente da sociedade digital. Warschauer (2003) observa que a inclusão digital é fundamental para a equidade social, o

que permite que todos os indivíduos tenham acesso às oportunidades educacionais, econômicas e sociais proporcionadas pela era digital.

Por fim, a transparência e a responsabilidade são pilares fundamentais das políticas públicas na era digital. A implementação de políticas de transparência pode incluir a obrigatoriedade de relatórios de impacto sobre a privacidade, a realização de auditorias regulares de segurança de dados e a criação de mecanismos para que os cidadãos denunciem abusos ou violações de dados.

4.2. A Importância da Colaboração Internacional

A natureza global da internet e das tecnologias digitais torna a colaboração internacional essencial para a proteção dos direitos na era digital. A harmonização das leis de proteção de dados, a cooperação na aplicação da lei e a troca de informações são fundamentais para enfrentar os desafios transnacionais do colonialismo digital.

Nesse sentido, a cooperação entre as autoridades de proteção de dados e as agências de aplicação da lei de diferentes países também é essencial para combater violações de dados transnacionais. Isso pode incluir a realização de investigações conjuntas, a partilha de informações e a assistência mútua na aplicação da lei. A criação de redes internacionais de cooperação, como a Global Privacy Enforcement Network (GPEN), pode facilitar essa colaboração, bem como a troca de informações e melhores práticas entre países pode ajudar a melhorar a eficácia das políticas de proteção de dados. Isso pode incluir a organização de conferências internacionais, a publicação de relatórios conjuntos e a criação de plataformas de conhecimento compartilhado. Essas iniciativas

podem ajudar os países a aprenderem uns com os outros e a adotar as melhores práticas em suas próprias jurisdições.

Por outro lado, tratados internacionais e os acordos de cooperação são instrumentos igualmente importantes para garantir a proteção dos dados em um contexto global. Esses acordos podem estabelecer normas mínimas de proteção de dados, promover a cooperação na aplicação da lei e garantir que os direitos dos indivíduos sejam respeitados em todas as jurisdições. No sentido do que observa o trabalho de Kuner (2013), cabe ressaltar que os tratados internacionais são essenciais para criar um quadro regulatório global que proteja os direitos dos indivíduos e promova a confiança no ambiente digital.

A colaboração internacional também pode incluir o desenvolvimento de capacidades e a assistência técnica para países em desenvolvimento que estão implementando e aprimorando suas leis de proteção de dados. Isso pode incluir a oferta de treinamento para reguladores, a criação de programas de intercâmbio e a provisão de suporte técnico para a implementação de tecnologias de proteção de dados. Essas iniciativas são essenciais para garantir que todos os países possam proteger os direitos de seus cidadãos na era digital.

Conclusão

A análise sobre o impacto da inteligência artificial (IA) e do colonialismo digital nas políticas públicas e na proteção dos direitos fundamentais evidencia a necessidade urgente de um reexame profundo das estruturas legais e sociais vigentes. O avanço tecnológico trouxe benefícios incontestáveis, porém, também revelou e amplificou desigualdades sociais e econômicas, exigindo uma abordagem holística para mitigar seus efeitos adversos.

O conceito de colonialismo digital nos alerta para uma nova forma de dominação e exploração, onde dados pessoais são utilizados como recursos a serem extraídos e monetizados, muitas vezes sem o consentimento adequado dos indivíduos. Essa dinâmica não apenas perpetua desigualdades históricas, mas também cria formas de controle e vigilância que ameaçam a autonomia e a liberdade dos cidadãos. A concentração de poder nas mãos de poucas corporações tecnológicas é um reflexo preocupante dessa nova era, onde o controle de informações se traduz em poder econômico e político.

A proteção de dados e a privacidade emergem como direitos fundamentais indispensáveis para a dignidade humana na era digital. A implementação de regulamentações como a Lei Geral de Proteção de Dados (LGPD) no Brasil e a General Data Protection Regulation (GDPR) na Europa são passos significativos, mas não suficientes. A eficácia dessas leis depende não apenas de sua aplicação rigorosa, mas também da capacidade das instituições de se adaptarem às rápidas mudanças tecnológicas. Além disso, a cooperação internacional é crucial para enfrentar os desafios transfronteiriços impostos pela globalização dos fluxos de dados.

A digitalização e a IA têm o potencial de melhorar o acesso à justiça e aumentar a eficiência em diversos setores, mas também apresentam riscos significativos. A automação de processos decisórios pode perpetuar preconceitos e discriminações se não for acompanhada de transparência e mecanismos de auditoria. A responsabilidade e a equidade nas decisões automatizadas são princípios que devem ser rigorosamente observados para garantir que a tecnologia sirva ao bem comum e não aos interesses de uma minoria privilegiada.

O colonialismo digital e o capitalismo de vigilância exemplificam como as dinâmicas de poder e exploração podem ser replicadas na era digital. Essas práticas representam uma ameaça não apenas à privacidade, mas também

à própria estrutura democrática das sociedades. A manipulação de informações e a criação de "bolhas de filtro" nas plataformas digitais comprometem o debate público e a liberdade de expressão, pilares fundamentais de qualquer democracia saudável. Para enfrentar esses desafios, é essencial que políticas públicas sejam formuladas com base em princípios de transparência, responsabilidade e inclusão.

Notas finais

1 ?

2 "Dados pessoais referem-se a qualquer informação relacionada a uma pessoa natural identificada ou identificável. Uma pessoa natural é considerada identificável quando é possível identificar essa pessoa, direta ou indiretamente, em particular por referência a um identificador como um nome, um número de identificação, dados de localização, identificadores online, ou a um ou mais fatores específicos à identidade física, fisiológica, genética, mental, econômica, cultural ou social dessa pessoa" (MENDES; LEITE, 2019).

Referências bibliográficas

- ACLU Michigan. (2020). *Man Wrongfully Arrested Because Face Recognition Can't Tell Black People Apart*. American Civil Liberties Union. <https://www.aclu.org/press-releases/man-wrongfully-arrested-because-face-recognition-cant-tell-black-people-apart>
- ADC. (2016). *El sistema de protección de datos personales en América Latina: oportunidades y desafíos para los derechos humanos – Vol. I | Asociación por los Derechos Civiles*. Asociación Por Los Derechos Civiles. <https://adc.org.ar/informes/sistema-proteccion-datos-personales-latam>
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). *Machine Bias*. ProPublica. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Baptista Luz Advogados, da Silva, D. R. S., Silva, G. H. L., Ribeiro, N. G., Júnior, O. P., & Braoios, R. R. (2022, December). *Guia América Latina: Legislações de Proteção de Dados | A Year in Privacy #11 - Baptista Luz*. Baptista Luz. <https://baptistaluz.com.br/america-latina-legislacoes-de-protecao-de-dados-a-year-in-privacy-11>
- Bartlett, R. P., Morse, A., Stanton, R., & Wallace, N. (2019). *Consumer Lending Discrimination in the FinTech Era*. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3063448>
- Bioni, B. R. (2018). *Proteção de dados pessoais: a função e os limites do consentimento*. Editora Forense.
- Brennen, S., & Kreiss, D. (2014, September 8). *Digitalization and Digitization – Culture Digitally*. Culturedigitally.org. <https://culturedigitally.org/2014/09/digitalization-and-digitization/>
- Burrell, J. (2015). How the Machine “Thinks:” Understanding Opacity in Machine Learning Algorithms. SSRN Electronic Journal, 3(1). <https://doi.org/10.2139/ssrn.2660674>
- Cassino, J. F., Souza, J., & Amadeu da Silveira, S. (Eds.). (2022). *Colonialismo de dados*. Autonomia Literária.
- Couldry, N., & Mejias, U. A. (2019). *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating it for Capitalism*. Stanford University Press.
- Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.
- de Hert, P., & Papakonstantinou, V. (2016). The new General Data Protection Regulation: Still a sound system for the protection of individuals? *Computer Law & Security Review*, 32(2), 179–194. <https://doi.org/10.1016/j.clsr.2016.02.006>
- de Oliveira, S. R. (2021). Sorria, você está sendo filmado!: repensando direitos na era do reconhecimento facial. *Thomson Reuters, Revista Dos Tribunais*.
- Douek, E. (2020). Governing Online Speech: From “Posts-As-Trumps” to Proportionality and Probability. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3679607>
- Eubanks, V. (2018). *Automating Inequality: how high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Faustino, D., & Lippold, W. (2023). *Colonialismo digital*. Boitempo Editorial.
- Greengard, S. (2021). *The internet of things*. Mit Press.
- Kuner, C. (2013). *Transborder data flows and data privacy law*. Oxford University Press.
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data: a Revolution That Will Transform How We Live, Work and Think*. John Murray.

- Newman, A. (2009). Bennett, Colin. 2008. The Privacy Advocates: Resisting the Spread of Surveillance. Cambridge: MIT Press. *Surveillance & Society*, 6(3), 343–344. <https://doi.org/10.24908/ss.v6i3.3299>
- Nissenbaum, H. F. (2009). *Privacy in context: technology, policy, and the integrity of social life*. Stanford University Press.
- Nunes, P. (2022). *Um Rio de olhos seletivos – uso de reconhecimento facial pela polícia fluminense*. CeSEC–Centro de Estudos de Segurança e Cidadania.
- O’neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Penguin Books.
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations. *Science*, 366(6464), 447–453. <https://doi.org/10.1126/science.aax2342>
- Pariser, E. (2011). *The Filter Bubble: What the internet is hiding from you*. Penguin Press.
- Pasquale, F. (2016). The Black Box Society: The Secret Algorithms That Control Money and Information. *Contemporary Sociology: A Journal of Reviews*, 45(3), 367–368. <https://doi.org/10.1177/0094306116641409c>
- Privacy International. (n.d.). *A world without data exploitation | Privacy International*. Privacyinternational.org. <https://privacyinternational.org/demand/a-world-without-data-exploitation>
- Privacy International. (2020). *2020 is a crucial year to fight for data protection in Africa*. Privacy International. <https://privacyinternational.org/long-read/3390/2020-crucial-year-fight-data-protection-africa>
- R7. (2021, December 15). “Disseram que eu era traficante”, diz pedreiro preso injustamente. Notícias R7. <https://noticias.r7.com/brasil/disseram-que-eu-era-trafficante-diz-pedreiro-preso-injustamente-16122021/>
- Rede de Observatórios de Segurança (Ed.). (2019). *Retratos da Violência: Cinco meses de monitoramento, análises e descobertas*. Centro de Estudos de Segurança e Cidadania (CESeC).
- Rodotà, S., Moraes, M. C. B. de, Doneda, D., & Doneda, L. C. (2008). *A vida na sociedade da vigilância: a privacidade hoje* (p. 381). Renovar.
- Russel, S., & Norvig, P. (2021). *Artificial intelligence: A Modern approach* (4th ed.). Prentice Hall.
- Schneier, B. (2015). *Data and Goliath: the hidden battles to collect your data and control your world*. WW Norton & Company.
- Selwyn, N. (2004). Reconsidering Political and Popular Understandings of the Digital Divide. *New Media & Society*, 6(3), 341–362. <https://doi.org/10.1177/1461444804042519>
- Stevenson, T. (2024). *Navigating Digital Neocolonialism in Africa*. Centre for International Governance Innovation. <https://www.cigionline.org/publications/navigating-digital-neocolonialism-in-africa/>
- Stokes, E., & CBS News. (2020, November 19). *Wrongful arrest exposes racial bias in facial recognition technology*. CBS News. <https://www.cbsnews.com/news/detroit-facial-recognition-surveillance-camera-racial-bias-crime/>
- Sunstein, C. R. (2018). *#Republic : Divided Democracy In The Age Of Social Media*. Princeton University Press.
- Veale, M., & Binns, R. (2017). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society*, 4(2), 205395171774353. <https://doi.org/10.1177/2053951717743530>
- Veloso Meireles, A. (2023). Privacidade no século 21: proteção de dados, democracia e modelos regulatórios. *Revista Brasileira de Ciência Política*, 41. <https://doi.org/10.1590/0103-3352.2023.41.265909>

- Veronese, A., Igreja, R. L., & Silveira, A. (2023). Cultura, privacidade e proteção de dados pessoais na América Latina. *Revista de Estudos Empíricos Em Direito*, 10, 1–44. <https://doi.org/10.19092/reed.v10.766>
- Warren, S. D., & Brandeis, L. D. (1890). The Right to Privacy. *Harvard Law Review*, 4(5), 193–220. <https://doi.org/10.2307/1321160>
- Warschauer, M. (2002). *Technology and social inclusion: Rethinking the digital divide*. MIT Press.
- Wu, T. (2017). *The attention merchants: from the daily newspaper to social media, how our time and attention is harvested and sold*. London Atlantic Books.
- Yeung, K. (2017). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, 12(4), 505–523. <https://doi.org/10.1111/rego.12158>
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Public Affairs.

ARTIGO

Participação Cidadã e Regulação da Inteligência Artificial: Análise da Composição Setorial das Audiências Públicas na CJSUBIA

Juan Manuel García

jm.garcia@derechosdigitales.org

Coordenador de Pesquisa na Derechos Digitales. É Licenciado em Comunicação pela Universidade Nacional de San Martín e possui um Mestrado em Políticas Públicas pela Universidade de Georgetown, além de estudos de pós-graduação sobre governança da internet e de dados.

Jamila Venturini

jamila@derechosdigitales.org

Co-Diretora Executiva da Derechos Digitales. É Jornalista pela Universidade de São Paulo (Brasil) e mestre em Ciências Sociais com foco em Educação pela Faculdade Latino-Americana de Ciências Sociais (Flacso, Argentina), Jamila também é membro da Rede Latino-Americana de Estudos sobre Vigilância, Tecnologia e Sociedade (Lavits).

Participação Cidadã e Regulação da Inteligência Artificial: Análise da Composição Setorial das Audiências Públicas na CJSUBIA

Palavras-chave

Governança de Inteligência Artificial
Participação Cidadã
Audiências públicas
Políticas de tecnologia no Brasil
Inclusão e representatividade

Resumo

Este artigo tem como objetivo analisar o processo de participação no contexto do Comissão de Juristas responsável por subsidiar elaboração de substitutivo sobre inteligência artificial no Brasil (CJSUBIA), focando em duas dimensões principais. Primeiro, examina a continuidade de uma tradição de discussões mais amplas na sociedade sobre políticas de tecnologia no Brasil, incluindo marcos importantes como o Marco Civil da Internet, a Lei Geral de Proteção de Dados e o modelo de governança multissetorial da infraestrutura de internet do país. Em segundo lugar, o artigo explora a composição interna do processo do CJSUBIA para entender quem participa efetivamente desses processos de tomada de decisão. Ao investigar essas dimensões, o estudo busca fornecer insights sobre a inclusão e a representatividade da abordagem do Brasil para a governança da IA.

Civic Participation and AI Regulation: Analysis of the Sectoral Composition of Public Hearings in the CJSUBIA

Keywords

AI governance, civic participation
Public hearings
Technology policies in Brazil
Inclusion and representativeness

Abstract

This paper aims to analyze the participation process within the context of the Commission of Jurists Responsible for Supporting the Drafting of a Substitute Bill on Artificial Intelligence in Brazil, focusing on two key dimensions. First, it examines the continuity of a tradition of broader societal discussions on technology policies in Brazil, including significant milestones such as the Marco Civil of the Internet, the General Data Protection Law, and the multi-stakeholder governance model of the country's internet infrastructure. Secondly, the paper explores the internal composition of the CJSUBIA process to understand who effectively participates in these decision-making processes. By investigating these dimensions, the study seeks to provide insights into the inclusiveness and representativeness of Brazil's approach to AI governance.

1. Introdução

Este artículo analiza la participación de la ciudadanía en el proceso de deliberación parlamentaria para la regulación de la Inteligencia Artificial (IA) en Brasil. Estará centrado en el análisis de los procesos asociados a la Comisión de Juristas encargada de subsidiar la elaboración de un sustituto sobre inteligencia artificial en Brasil (CJSUBIA, por sus siglas en portugués), comisión especial que funcionó en el Senado Federal entre febrero y diciembre de 2022, y que tuvo el objetivo de unificar una serie de proyectos de ley sobre el tema.

La regulación de la IA en Brasil tuvo su primer hito sustantivo en septiembre del año 2021, con el debate y aprobación del Proyecto de Ley (PL) 21/2020¹ en la Cámara de Diputados. Se trata del primer proyecto de ley que superó la instancia de trabajo en comisiones parlamentarias en la región. Este proceso se dio bajo un régimen de urgencia y fue fuertemente criticado por los escasos espacios abiertos a la participación que generó (Coalizão Direitos na Rede, 2021). Su contenido, resumido y basado en un conjunto de principios generales, también sufrió críticas de la sociedad civil y academia.²

Al momento de la media sanción del PL 21/2020, ya habían sido presentados dos proyectos en Senado que buscaban establecer principios y marcos éticos para el uso de la IA en Brasil. Se trata de los PL 5051/2019³ y el 872/2021⁴. De esta manera, dado que el PL 21/2020 sería el tercer proyecto de características similares aguardando a ser tratado, la presidencia de la Casa instauró la mencionada comisión especial, con el propósito de generar un sustitutivo que integre el contenido de todas las iniciativas. Uno de los aspectos centrales de la CJSUBIA consistió en la implementación de diversos espacios de participación: doce paneles de audiencias públicas, un seminario internacional con siete paneles y una consulta

pública online para recepción de contribuciones escritas.

El objetivo del presente trabajo será analizar el proceso de participación que tuvo lugar en el marco de la CJSUBIA, en dos dimensiones. Primero, en tanto continuidad a una tradición de discusiones ampliadas en la sociedad sobre políticas de tecnología en Brasil, que incluye procesos como el Marco Civil de Internet, la Ley General de Protección de Datos e incluso el modelo de gobernanza de múltiples partes interesadas de la infraestructura de internet en el país. Por otro lado, hacia dentro del proceso, analizaremos su composición, para entender la distribución sectorial y regional de la participación en estos procesos.

1.1 La importancia de la participación ciudadana en políticas de IA

Es importante resaltar que la participación ciudadana en procesos decisorios sobre IA ha sido destacada como crucial en distintos documentos e iniciativas internacionales. Eso se debe a que permite asegurar que las tecnologías desarrolladas estén alineadas con los principios y debates de la sociedad, además de garantizar políticas públicas informadas por el contexto y realidades locales. Como veremos a continuación, distintos documentos y recomendaciones internacionales consideran la participación como un mecanismo que puede ayudar a reducir brechas y garantizar mecanismos de gobernanza más inclusivos.

Documento	Órgano	Año	Consideraciones sobre Participación
Recomendación sobre la Ética de la IA	UNESCO	2021	Promover la participación activa de todas las personas o grupos, independientemente de su raza, color, ascendencia, género, edad, idioma, religión, opiniones políticas, origen nacional, étnico o social, condición económica o social de nacimiento, discapacidad o cualquier otro motivo. La participación de las diferentes partes interesadas a lo largo del ciclo de vida de los sistemas de IA es necesaria para garantizar enfoques inclusivos.
Reporte Interino	AIAB, ONU	2023	Garantizar la participación de todos los grupos interesados y de todos los países y regiones en la gobernanza colectiva, la gestión de riesgos y aprovechamiento de oportunidades.
Declaración de Santiago	UNESCO CAF	2023	Proveer la participación de múltiples partes interesadas en el proceso de diseño e implementación de políticas públicas, promover iniciativas y políticas inclusivas, con énfasis en la inclusión, equidad de género y disminución de brechas con perspectiva interseccional.
Resolución 78/265	ONU	2024	Elaborar y apoyar enfoques y marcos normativos y de gobernanza referentes a los sistemas seguros y fiables de inteligencia artificial.
Convención Marco sobre IA y Derechos Humanos	Consejo de Europa	2024	Garantizar que las cuestiones importantes relacionadas con los sistemas de IA se examinen mediante debate público y consultas con las múltiples partes interesadas.

Tabla 1: Menciones sobre participación e IA en marcos de referencia internacionales

A pesar del avance en el reconocimiento de la participación en el ámbito de los documentos internacionales sobre IA, un estudio sobre los espacios de participación en el ámbito de las discusiones sobre estrategias de IA en la región (Hernández et al., 2022) identificó una falta de inclusión efectiva de grupos históricamente marginalizados y personas con discapacidad en los procesos analizados, así como falencias en los mecanismos de transparencia efectivos que permitan disponer de cierta trazabilidad sobre cómo se

gestionó la influencia de los comentarios en las estrategias finales.

En este sentido, la experiencia de la CJSUBIA se muestra como una oportunidad para profundizar en ese aspecto y analizar la participación ciudadana en discusiones de políticas sobre IA. A diferencia del momento en que se realizó el análisis de Hernández et al., ahora contamos con un conjunto más amplio de guías y orientaciones sobre la relevancia de la participación desde el contexto internacional, de modo que es relevante mirar si tales documentos han ejercido influencia en el contexto brasileño.

1.2 Dimensiones de análisis

Este trabajo analizará el proceso de la CJUSBIA en dos dimensiones. La dimensión de análisis histórica busca revisar la actuación de la Comisión en tanto hito en materia de participación dentro de una secuencia de procesos deliberativos en el campo de la regulación de las tecnologías en Brasil. Para analizar esta primera dimensión, realizaremos una revisión bibliográfica que permita recuperar estos antecedentes y una visión ampliada del espacio de estos procesos desde el regreso democrático en Brasil, a fines de los años '80. En segunda instancia, para la dimensión focalizada, analizaremos la composición del proceso participativo de CJUSBIA.

La diversidad en la participación fue uno de los principales objetivos buscados por la CJSUBIA en sus mecanismos de consulta, según el texto del reporte final de la Comisión. El mismo afirma que las audiencias “se organizaron en un formato multisectorial, buscando garantizar la diversidad entre los expositores.” Y continúa: “con esto, fue posible recoger miradas plurales del sector público, la academia, la industria y el tercer sector de diferentes regiones de Brasil con hablantes de múltiples identidades raciales y de género”. Cabe enfatizar que tales aspectos fueron objeto de discusión pública durante el periodo de conformación de la Comisión, debido a que algunos sectores consideraron que su composición no reflejaba la diversidad de la población brasileña al no incluir juristas “negros, negras e indígenas”, como enfatizaba una carta de un grupo de más de 50 organizaciones de sociedad civil a la Presidencia del Senado en febrero de 2022.⁵ Según reportes de la época, la convocatoria de audiencias públicas fue una forma de responder a las críticas y promover mayor participación en las discusiones (Irishb, 2022).

Es por este motivo que consideramos pertinente el análisis en profundidad del proceso,

a partir de informaciones sobre representación sectorial y distribución geográfica. El análisis estará enfocado en estos dos aspectos dado el escaso nivel de desagregación de los datos publicados por las fuentes oficiales del Senado, lo que se presenta como una primera limitante para confirmar si los referidos objetivos de diversidad y pluralidad a nivel racial y de género fueron cumplidos. Aunque existen estimaciones públicas acerca de la proporción de participantes negras en los debates (Irishb. (2022), por ejemplo, la ausencia de información oficial sobre las participantes o de una contabilización por parte del Senado impide el análisis objetivo de las variables género y raza a posterior. Tal análisis supondría una asunción subjetiva a partir de datos como nombre, por ejemplo, que pueden ser extremadamente frágiles para la indicación de género, además de implicar potenciales sesgos adicionales.

1.3 Consideraciones metodológicas

Para el análisis de los datos de participación, se creó una base con dos fuentes distintas. La primera, que permitió el armado de la base de datos de Audiencias Públicas y del Seminario Internacional, fue la web del Senado titulada “Documentos de Audiências Públicas”⁶. Esta web tiene un listado no estructurado de todas las audiencias públicas, sus paneles y participantes. Para la reconstrucción de la base de datos de las Contribuciones Escritas fue utilizado el listado publicado en la Sección VII, *Apenso IV* del reporte final (*relatorio final*) de la Comisión, publicado en PDF⁷. Este proceso permitió obtener una base de datos con todas las exposiciones y aportes que hubo en cada espacio de participación. A ese listado se agregaron las variables Sector, Intereses, País y Ciudad, a

partir de un proceso de análisis que describiremos a continuación.

Para la variable Sector se utilizaron las categorías que el personal del Senado asignó a las contribuciones escritas. Incluye las categorías “Gobierno”, “Academia”, “Sector Privado”, “Sociedad Civil” e “Individual”. Estas categorías pueden representar una visión clásica de la división de los modelos de gobernanza de múltiples partes interesadas (*multistakeholder*), según identificada por Virgilio Almeida et al. en su estudio (2015) sobre el origen de su implementación en el ámbito de la gobernanza de internet. Este modelo fue tomado para la gestión del CGI.br, así como en los procesos de la Cumbre Mundial sobre la Sociedad de la Información (CMSI). La categoría “Individual” fue asignada por el personal del Senado a algunas contribuciones escritas con un criterio no explicitado. El reporte final de la CJSUBIA sólo indica que éstas fueron “analizadas y organizadas según su origen” (CJSUBIA, 2022, p. 198). La variable Sector fue reconstruida para las audiencias públicas y para el seminario internacional, dado que esos espacios no estaban categorizados de origen. Se asignó alguna de las categorías mencionadas a cada expositor, según la afiliación publicada en la web Documentos de Audiências Públicas, como fue mencionado.

A pesar de tratarse de una clasificación consolidada en distintos documentos y foros, algunas autoras y autores apuntan que esta categorización puede no representar de forma fiel los intereses efectivos de las instituciones que participan en los procesos de gobernanza de internet. Ellas apuntan distorsiones como la existencia de organizaciones de Sociedad Civil vinculadas estrechamente al sector privado, que, argumentan, deberían categorizarse de manera distinta a organizaciones sociales de base, por ejemplo. En ese sentido, y a partir de un análisis de la composición de las mesas de trabajo del Comité de Gobernanza de la Estrategia Brasileña de IA (EBIA), Belli et al.

(2023) generaron una propuesta alternativa, la cual hemos utilizado para la variable “Intereses” y que describiremos a continuación.

La categorización propuesta por Belli et al. mantiene las categorías de Academia, Sector Privado e Industria, pero subdivide Sociedad Civil en “Organizaciones Sin Fines de Lucro”, “Organizaciones Sociales” (las cuales hemos tomado de forma integrada en la primera) y “Organizaciones del Sector Privado” (Belli et al., 2023, pp. 6-7). La primera, incluye ONGs que no son creadas por empresas privadas ni representan intereses privados. La segunda, organizaciones sin fines de lucro que representan o agrupan empresas privadas con agendas compartidas (los autores ponen como ejemplo la Confederación Nacional de la Industria). Además, por otro lado, esta categorización agrega la categoría “Instituciones Internacionales”, en la que incluyen principalmente a organismos técnicos.

La clasificación de Belli et al. representa un intento de superar eventuales desviaciones en cómo distintos grupos sociales se buscan representar en la distribución sectorial clásica. Sin embargo, es difícil, en la práctica, identificar cómo cada organización puede efectivamente representar distintos intereses, una vez que están permeadas por relaciones múltiples y complejas. Eso no es exclusivo del sector de sociedad civil, por lo que puede considerarse una limitación en la categorización.

Por último, las variables de procedencia geográfica “Ciudad” y “País” fueron reconstruidas en su totalidad, dado que no estaban categorizadas en las fuentes oficiales. Ellas están compuestas por nombres de ciudades y países y fueron asignadas según la procedencia de la institución a la que están afiliados los y las expositores. Los datos fueron, asignados según la dirección publicada en el sitio web oficial de dichas instituciones. Esto permitió categorizar la totalidad de las instituciones participantes en las audiencias públicas y el seminario

internacional. Para las “contribuciones escritas”, dado los textos de las contribuciones fueron publicados en la web del Senado⁸, se utilizó la ciudad mencionada en el encabezado de esas misivas. Para los casos donde las instituciones representan una agrupación, cámara o federación, se utilizó la categoría “Múltiple”. Para los casos donde no fue posible reconstruir el dato a partir de información certera, se utilizó la categoría “sin datos”.

A continuación, compartimos los resultados de esta investigación, para analizar en términos de distribución qué tan efectivamente pudo este proceso cumplir con su misión de pluralidad en términos de representación sectorial y geográfica.

2. Participación Ciudadana y Gobernanza de Internet en Brasil

Con el objetivo de ofrecer una visión ampliada de los espacios de participación en la CJSUBIA, abordaremos dos vías de análisis: primero, los espacios de participación como instrumentos para ampliar la democracia en Brasil y, segundo, la posible correlación entre el avance democrático y la creación de organismos e instrumentos de participación que fortalecieron una comunidad de múltiples partes interesadas en la gobernanza de internet.

La pregunta por la regulación de Internet y el rol del gobierno, así como de la academia, la sociedad civil y la industria, centró gran parte de la conversación sobre el desarrollo de internet a comienzo de la década de 1990. En Brasil, según los investigadores Carlos Affonso de Souza y Christian Perrone (2023), el desarrollo de internet fue un esfuerzo colectivo multisectorial que tuvo uno de sus principales hitos en 1991, con la instalación de una línea

fija permanente entre la red brasileña y la red troncal de Internet de Estados Unidos. El hito se dió sólo tres años después de sancionada la nueva Constitución. De hecho, los autores resaltan que “el desarrollo de Internet en Brasil coincidió con la democratización del país y su proceso de apertura de su mercado interno para productos y capitales extranjeros.” (p. 221).

Según resaltan De Souza y Perrone, la creación, en 1995, del Comité de Gestión de Internet de Brasil (CGI.br), representó el inicio de un organismo clave de consolidación de una comunidad multisectorial involucrada en la toma de decisiones sobre la agenda. El Comité está compuesto por 21 miembros: nueve de organizaciones gubernamentales, cuatro de sociedad civil, cuatro del sector privado y cuatro de la academia y la comunidad técnica. Como resaltan Almeida et al. (2015) la composición de la junta directiva de CGI.br refleja la naturaleza de múltiples partes interesadas de Internet. La creación de este organismo representa un antecedente clave para el análisis de los procesos participativos actuales en materia de gobernanza de IA en Brasil, una vez que consolida institucionalmente un espacio donde la academia, la industria y la sociedad civil toman parte, junto al gobierno, en la gestión de una agenda directamente relacionada, como lo es la gobernanza de internet. Por otro lado, contribuye con la creación de capacidades específicas a partir de una comunidad que tiene lugar en ambos espacios. Eso es visible, por ejemplo, en el hecho de que integrantes actuales y pasados de CGI.br participaron en las audiencias públicas de la CJSUBIA como expertos.

Un antecedente clave relacionado con el CGI.br es el proceso de debate sobre la Ley del Marco Civil de Internet (Ley 12.965), esencial para la regulación del ecosistema de internet en Brasil. La discusión de esta ley comenzó con una consulta pública en línea entre 2009 y 2010, que generó un anteproyecto enviado al Congreso en 2011 y, tras años de debate, fue

sancionada en 2014. Según Carlos Affonso Souza et al. (2017, p. 3), este proceso involucró a una amplia gama de actores de la sociedad civil, el sector privado, la comunidad técnica, la academia e incluso usuarios individuales. Fue la primera vez que se utilizó un sistema de consulta pública en línea en un proceso de este tipo, mecanismo luego replicado en otras discusiones, como la de la ley general de protección de datos.

La consulta online tuvo dos etapas. En la primera, se recibieron más de 800 comentarios de todos los sectores, revisados por un grupo de expertos multisectorial creado por el Ministerio de Justicia. En la segunda, el anteproyecto generado fue abierto a consulta pública por 45 días, permitiendo a los participantes comentar artículo por artículo y ver los aportes recibidos, los cuales podían valorarse mediante un sistema de pulgar arriba/pulgar abajo (Affonso Souza et al., 2017, p. 9). Según analizan Bragatto et al. (2014) y Nicolás et al. (2017), en sus estudios sobre ambas fases del proceso de consulta, durante estas instancias, los participantes procuraron entablar intercambios discursivos justificando sus posturas mediante narraciones o perspectivas técnicas.

El uso de tecnologías de la comunicación en procesos de participación ciudadana puede ser sujeto a críticas por distintos factores, incluso el potencial uso para manipular o reforzar la hegemonía de sectores de poder sobre los ciudadanos comunes. Sin embargo, como rescatan Martínez Stefani y Vaz (2016), un análisis “moderadamente optimista” de esta experiencia puede considerar que las tecnologías de participación basadas en internet son resultado de una construcción social, donde los elementos tecnológicos dan forma al cambio social, a la vez que responden a procesos sociales más amplios. De esta manera, afirman los autores, la experiencia del Marco Civil muestra que el empleo de tecnología para impulsar la participación ciudadana es una posibilidad real y que los

gobiernos están dispuestos a implementarlas.

La reflexión apunta a la importancia de encontrar un medio término entre un extremo pesimismo relacionado a la preocupación con control y manipulación, y un optimismo ciego que confía en las tecnologías como mejor (o único) mecanismo para la participación significativa. En ese mismo sentido, como Derechos Digitales ha manifestado en distintas ocasiones (Velsasco, 2021), la participación en línea favorece la inclusión de más actores en ciertos debates, pero debe siempre ser acompañada de otros mecanismos de consulta, una vez que los países de América Latina siguen permeados por brechas digitales significativas. Por ejemplo, en 2010, fecha de la segunda consulta realizada para el Marco Civil de Internet, solamente 27% de la población brasileña contaba con acceso a la red en sus hogares y 41% era usuaria de internet según datos de CETIC.br (2010, p. 136). Los datos evidencian algo que los números absolutos de participantes o comentarios en las consultas no muestran, que es la potencial desigualdad en la participación desde ese proceso.

Más allá de la discusión sobre la representatividad de la participación, estos dos eventos, es decir, la configuración de un sistema de múltiples partes interesadas con CGI.br y el debate multisectorial de una ley marco para la gobernanza de internet, representan hitos que ayudaron a la creación de una comunidad especializada en una serie de agendas emergentes, vinculadas a las problemáticas del uso de internet: especialistas en ciberseguridad, privacidad, copyright, protección de datos personales, aspectos técnicos de internet y las telecomunicaciones, libertad de expresión, acceso abierto, entre tantos otros. Las implicancias de la formación de esa masa crítica de personas e instituciones pueden verse en el grado de organización de la sociedad civil brasileña para incidir en discusiones clave sobre políticas de tecnologías. Un ejemplo es la Coalición de Derechos en la Red (CDR), un grupo de más de 50 organizaciones

de todo Brasil, impulsoras del respeto por los derechos humanos en la red. La CDR surge en junio de 2016⁹ durante las discusiones sobre la Ley General de Protección de Datos (LGPD) y en reacción a un conjunto de proyectos de ley que buscaban dar otra dirección al consenso alcanzado a raíz del Marco Civil. Como resaltan Bioni y Marques Rielli (2021, p. 29) la Coalición emerge entre organizaciones que vivieron la experiencia de articulación para la aprobación del Marco Civil de Internet. El proceso de debate parlamentario de la LGPD fue otra de las instancias decisivas en materia de participación ciudadana, y la CDR dio inicio a un trabajo que continúa aún hoy de incidencia legislativa, con gran intervención en el proceso de regulación de la IA en Brasil.

En 2025 se cumplirán 40 años del regreso a la democracia en Brasil, cuando comenzó el proceso de pluralidad e inclusión de la sociedad en procesos más participativos para la generación de políticas públicas. También se cumplirán 30 años de la creación de CGI.br, fundada en mayo de 1995, una instancia clave para la conformación de una discusión multi-sectorial. Este año, además, se cumplen 10 años de la sanción del Marco Civil de Internet, sancionado en abril de 2014, una de las instancias definitivas para la consolidación de una masa crítica de investigadores, activistas y técnicos interesados en la participación en políticas públicas, además de una instancia clave para la innovación en los mecanismos de participación. A raíz de múltiples reflexiones sobre las fortalezas y debilidades de estos procesos democráticos, tensionados en particular durante el gobierno de extrema derecha de Jair Bolsonaro (2019-2022) y un intento fallido de golpe de Estado, es relevante observarlos en el marco de la regulación de la IA en Brasil para conocer si, efectivamente, las experiencias mencionadas generaron aprendizajes que permitan fortalecer aún más la discusión sobre políticas de tecnología en el país. Además, habiendo

revisado brevemente las particularidades de Brasil en cuanto un país donde se ha consolidado una comunidad alrededor de temas tenidos como técnicos, el proceso brasileño puede servir como inspiración, considerando que los debates sobre regulación de IA a nivel latinoamericano avanzan rápidamente, pero con posibilidades de participación pública aún limitadas (Lara, 2024).

3. La discusión por la regulación de la IA en Brasil

Antes de llegar a la discusión en el marco de la CJSUBIA, es necesario describir el comienzo del proceso de debate sobre la regulación de IA en Brasil. Proceso relativamente pionero en la región, aunque, como veremos, marcado por graves problemas sustantivos y de formato, en particular en lo que tiene que ver con la participación pública.

Podemos encontrar un primer antecedente en la adhesión del Gobierno Federal, en mayo de 2019, a la Recomendación del Consejo de IA generada por la Organización para la Cooperación y el Desarrollo Económicos (OCDE)¹⁰. Se trata de un conjunto de directrices que tienen el objetivo de promover el desarrollo y la implementación responsable de la IA, basadas en cinco principios: "crecimiento inclusivo, desarrollo sostenible y bienestar", "respeto del Estado de Derecho, los derechos humanos y los valores democráticos, incluidas la imparcialidad y la privacidad.", "transparencia y explicabilidad", "robustez, seguridad y protección", y "rendición de cuentas". El texto, además, genera una serie de recomendaciones a ser implementadas por sus adherentes, entre las que destaca la de generar un esquema de políticas y gobernanza interoperable para la IA.

Esta última prerrogativa generó distintas iniciativas entre funcionarios y legisladores. Algunos meses más tarde, en septiembre, el Senador Styvenson Valentim (PODEMOS/RN), presentó el Proyecto de Ley (PL) 5051 de 2019, con el objetivo de “establecer los principios para el uso de la Inteligencia Artificial en Brasil”. Este fue el primer proyecto de ley dedicado específicamente a esta temática, basado en una serie de principios similares a los impulsados por la OCDE. Este proyecto no fue tratado en comisiones sino hasta su tratamiento conjunto en la CJSUBIA.

Una segunda acción, en diciembre del mismo año, 2019, se refiere a la apertura de una consulta pública para definir la Estrategia Brasileña de Inteligencia Artificial (EBIA), generada por el Ministerio de Ciencia, Tecnología, Innovaciones y Comunicaciones (MCTIC), con el objetivo de orientar el desarrollo y uso de IA en el país. La primera versión de la EBIA, luego de las mesas de trabajo, fue publicada solamente en abril de 2021 – después de la presentación de otros proyectos de ley en el ámbito parlamentario. El texto refiere a la recomendación de la OCDE como fundamento para la constitución de una serie de ejes transversales que centran su trabajo¹¹. Establece nueve ejes temáticos que incluyen la investigación, la innovación y el desarrollo de soluciones de IA, además del uso consciente y ético de la tecnología¹².

Durante el mismo año de la aprobación de la EBIA, pero en el mes de septiembre, el parlamento de Brasil dio lugar a uno de los hitos de mayor trascendencia para la regulación de la inteligencia artificial, no sólo para Brasil, sino para América Latina. Se trata de la aprobación, por la Cámara de Diputados, en primer turno, del PL 21/2020¹³ que, según su título, establecería las “Bases, Principios y Directrices para el desarrollo y aplicación de la Inteligencia Artificial en Brasil”. Se trató de la primera iniciativa de la región en alcanzar la instancia de tratamiento y votación en el plenario de la

Cámara, con una propuesta de alcance integral para este tipo de sistemas y caso siguiera tal ritmo, podría anticiparse incluso a la normativa europea, entonces en etapa de discusión. Sin embargo, su camino posterior encontraría ramificaciones que dificultaron la continuidad de su tratamiento. El PL 21/2020 fue criticado por distintos sectores de la sociedad, principalmente por tres aspectos: su tratamiento expeditivo, la falta de espacios de participación y su carácter eminentemente basado en principios o *principiológico* (CDR, 2021). De hecho, más allá de la ambición indicada en el título, el texto aprobado tenía sólo un total de 10 artículos.

El proyecto, presentado por el Diputado Eduardo Bismarck, fue tratado en un régimen de urgencia, un formato de tramitación excepcional que permite prescindir de “requisitos, intersticios y formalidades reglamentarias”, para que una propuesta pueda ser tratada por la Cámara con mayor celeridad (Portal da Câmara dos Deputados, 2021). Contrariando el histórico brasileño en estos temas y los cuidados con que se venía tratando la regulación de la IA en contextos como el europeo, la celeridad del tratamiento no dio lugar a un debate multisectorial que permitiera alcanzar un proyecto más representativo de las demandas e intereses de los diversos actores que componen la agenda.

La CDR publicó un texto muy crítico de este proceso, previo a la votación en diputados, donde justamente apunta que el debate avanzó a un paso desproporcionadamente más rápido que otras discusiones sobre temas de internet y tecnología digital. Cita como ejemplo los cuatro años de debate del Marco Civil contra los pocos meses de debate de este proyecto. En lo sustantivo afirma que el proyecto no cuenta con “instrumentos concretos para abordar aspectos relevantes de la tecnología, dada la posibilidad de que la aplicación de algunos de estos sistemas ponga a las personas en riesgo” (Coalizão Direitos na Rede, 2021). A este punto se suma el carácter cerrado del trabajo en comisiones con un

escaso proceso participativo: fueron solamente tres audiencias públicas, según figura en el reporte de la relatora de la Comisión de Ciencia y Tecnología, Comunicación e Informática, la Dip. Luisa Canziani (PTB-PR).¹⁴

3.1. La Creación de la Comisión de Juristas

Ya en Senado, el proceso de discusión de este proyecto de ley cambiaría diametralmente su ritmo de aceleración con la creación por su presidente de una Comisión de Juristas encargada redactar un proyecto sustitutivo sobre Inteligencia Artificial en Brasil, la CJSUBIA. Sustitutivo, dado que buscaría generar una versión integrada de los proyectos 5.051/2019, 21/2020 y 872/2021.

La Comisión fue presidida por el Ministro del Superior Tribunal de Justiça (STJ) Ricardo Villas Bôas Cueva y contó con la profesora de la Universidad de Brasília Laura Schertel Ferreira Mendes como relatora. Tuvo 16 integrantes:

1. Ana de Oliveira Frazão
2. Bruno Ricardo Bioni
3. Danilo Cesar Maganhoto Doneda
4. Fabrício de Mota Alves
5. Miriam Wimmer
6. Wederson Advincula Siquiera
7. Claudia Lima Marques
8. Juliano Souza de Albuquerque Maranhão
9. Thiago Luís Santos Sombra
10. Georges Abboud
11. Frederico Quadros D'Almeida
12. Victor Marcel Pinheiro
13. Estela Aranha

14. Clara Iglesias Keller
15. Mariana Giorgetti Valente
16. Filipe José Medon Affonso

La CJSUBIA fue instaurada el 17 de febrero de 2022 y tuvo un plazo de trabajo original de 120 días, que fue extendido por 120 días más, lo que da un total de 240 días –o seis meses– de trabajo, hasta el 7 de diciembre de 2022. Todos los trabajos de la comisión fueron documentados y publicados en la web del Senado Federal de Brasil¹⁵. Mantuvo once encuentros ordinarios, en cumplimiento de ese plan de trabajo. Durante las reuniones 2, 3, 4 y 5, tuvieron lugar las Audiencias Públicas, divididas en 12 paneles, de las que participaron 61 panelistas. Durante las reuniones 7 y 8, tuvo lugar un seminario internacional, en un esquema similar al de audiencias públicas, pero compuesto por participantes del extranjero. En paralelo a estos procesos, la Comisión abrió un espacio para el envío de contribuciones escritas que recibió 103 aportes de individuos y diversos sectores de la sociedad.

Según consta en el reporte final, las audiencias públicas “se organizaron en un formato multi-sectorial, buscando garantizar la diversidad entre los expositores”. A pesar de no presentar datos consolidados que corroboren la conclusión, el documento continúa: “con esto, fue posible recoger miradas plurales del sector público, la academia, la industria y el tercer sector de diferentes regiones de Brasil con hablantes de múltiples identidades raciales y de género”. Si bien estos objetivos son deseables y oportunos, es necesario hacer un análisis más detallado sobre las personas que integraron estos espacios y conocer mejor qué sectores representan y qué procedencia geográfica tienen para entender su alcance.

3.2. Análisis de las instancias de participación en la CJSUBIA

Las audiencias públicas tuvieron lugar los días 28 y 29 de abril y 12 y 13 de mayo de 2022. Fueron divididas en 12 paneles, de los que participaron 61 representantes de distintos sectores (ver Tabla 2).

Fecha	Panel	Cantidad Oradores
28/04/2022	Panel 1: Inteligencia artificial y regulación: objeto a ser regulado y aspectos socio-técnicos	5
	Panel 2: Inteligencia artificial y regulación: modelos de regulación y enfoque	5
	Panel 3: Inteligencia artificial y regulación: fundamentos y principios	5
29/04/2022	Panel 4: Contexto económico-social y beneficios: desarrollo sostenible y bienestar; competencia e innovación; consumo y marketing; investigación y desarrollo de IA; bases de datos, derechos de autor y minería.	6
	Panel 5: Experiencias sectoriales: Seguridad pública, ejecución de políticas públicas, innovación en la industria, transformación digital en las empresas, protección a la infancia.	5
	Panel 6: Inteligencia artificial y riesgos: grados de riesgo; hipótesis de riesgos inaceptables y principio de precaución.	5

Fecha	Panel	Cantidad Oradores
12/05/2022	Panel 7: Inteligencia artificial y riesgos: sesgos y discriminación.	4
	Panel 8: Atributos del diseño socio-técnico de confiabilidad de la IA: seguridad, precisión, transparencia, rastreabilidad y monitoreo.	5
	Panel 9: Derechos y deberes: transparencia y explicabilidad; revisión y derecho a la intervención humana; corrección de sesgos.	4
13/05/2022	Panel 10: Regímenes de responsabilidad civil.	4
	Panel 11: Estructuras institucionales de supervisión: comando y control, regulación responsiva y debate sobre órganos reguladores.	6
	Panel 12: Instrumentos regulatorios para la innovación: códigos éticos y mejores prácticas; evaluaciones de impacto; sandboxes y otros.	7

Tabla 2: Paneles en las Audiencias Públicas CJSUBIA

Fuente: elaboración propia

El Seminario Inter-nacional tuvo lugar los días 9 y 10 de junio de 2022. Durante esas dos jornadas, participaron 26 expositores divididos en 7 paneles (ver Tabla 3).

Fecha	Panel	Cantidad Oradores
09/06 /2022	Panel 1: Democracia y derechos fundamentales: fundamentos de la regulación de Inteligencia Artificial.	4
	Panel 2: Desafíos de la regulación de Inteligencia Artificial: experiencias comparadas.	4
	Panel 3: Transparencia, sesgo y debido proceso en la toma de decisiones automatizada.	5
10/06 /2022	Panel 4: Protección de datos y regulación de I.A.	4
	Panel 5: Técnicas regulatorias y enfoque basado en riesgo.	3
	Panel 6: Mercados y experiencia: Perspectivas sectoriales en I.A.	3

Tabla 3: Paneles en el Seminario Internacional CJSUBIA

Fuente: elaboración propia

Ambas instancias tuvieron carácter híbrido, con participaciones presenciales y remotas. Además, tanto las audiencias como el seminario fueron transmitidos en vivo por streaming, vía Youtube, por el canal oficial del Senado, “TV Senado”. TV Senado es la emisora oficial del Senado Federal, con alcance nacional. Además de transmitir las actividades diarias del Senado, la emisora produce periodismo diario, informativos, entrevistas, programas culturales, documentales, productos digitales e institucionales¹⁶. Es posible acceder a sus transmisiones a través de la televisión abierta y por su canal de Youtube, que cuenta con 1,51 millones de suscriptores, lo que constituye un potencial alcance no despreciable.

Espacio	Jornada	Vistas
Audiencias Públicas	1 ^{ra} jornada (paneles 1, 2 y 3)	3,454 vistas
	2 ^{da} jornada (paneles 4, 5 y 6)	2,901 vistas
	3 ^{ra} jornada (paneles 7, 8 y 9)	2,017 vistas
	4 ^{ta} jornada (paneles 10, 11 y 12)	1,226 vistas
Seminario Internacional	1 ^{ra} jornada (paneles 1, 2 y 3)	629 vistas
	2 ^{da} jornada (paneles 4, 5, 6 y 7)	831 vistas

Tabla 4: Cantidad de vistas en los videos de las sesiones publicados en el canal de Youtube TV Senado, divididos por jornada.

Fuente: elaboración propia

Cabe destacar el uso de una plataforma como Youtube tanto para la transmisión como para el archivo de sesiones, habilitando la disponibilidad para personas de todas latitudes. Puede ser otra oportunidad para ser “moderadamente optimistas” desde un punto de vista de acceso a la información y participación pública.

3.4 Datos generales de participación según sector, intereses y procedencia geográfica

A continuación, se presentan los totales generales de participación en todas las instancias de consulta (audiencias públicas, seminario internacional y aportes escritos), según la categorización clásica sectorial utilizada en la mayoría de los sistemas de múltiples partes interesadas (Tabla 5). Luego, presentamos la categorización en base a los intereses que representan esas organizaciones, en base a Belli et al. (2023) (Tabla 6).

Sector	Cantidad	Porcentaje
Privado	61	31.44
Academia	50	25.77
Sociedad Civil	38	19.59
Individual	30	15.46
Gobierno	15	7.73

Tabla 5: totales generales de participación por sector.

Fuente: elaboración propia

Intereses	Cantidad	Porcentaje
Academia	50	25.77
Asociaciones del Sector Privado	40	20.62
Asociaciones Civiles Sin Fines de Lucro	38	19.59
Individual	30	15.46
Sector Privado	21	10.82
Gobierno	13	6.70
Instituciones Internacionales	2	1.03

Tabla 6: totales generales de participación por intereses representados.

Fuente: elaboración propia

Además de las categorías “Individual” y “Academia”, podemos apreciar un desdoblamiento desde la categoría del sector privado hacia las categorías Sector Privado y Asociaciones del Sector Privado. Sociedad Civil pasa a llamarse específicamente “Asociaciones Civiles Sin Fines de Lucro”, para hacer hincapié en su carácter sin fines de lucro, y Gobierno se desdobra entre las categorías “Gobierno” e “Instituciones Internacionales”. El Gráfico 1 a continuación puede ser clarificador en este sentido.

Con los totales generales a la vista, es propio realizar algunos análisis desagregados de los resultados, para conocer mejor la distribución según los espacios de participación y la procedencia geográfica de los participantes.

Espacio	Sector	Cantidad	Porcentaje	Porcentaje Global
Aportes Escritos	Privado	37	35.24	19.07
Aportes Escritos	Individual	30	28.57	15.46
Aportes Escritos	Sociedad civil	19	18.10	9.79
Aportes Escritos	Academia	13	12.38	6.70
Aportes Escritos	Gobierno	6	5.71	03.09
Audiencia Pública	Academia	26	41.27	13.40
Audiencia Pública	Privado	19	30.16	9.79
Audiencia Pública	Sociedad Civil	13	20.63	6.70
Audiencia Pública	Gobierno	5	7.94	2.58
Seminario Internacional	Academia	11	42.31	5.67
Seminario Internacional	Sociedad Civil	6	23.08	03.09
Seminario Internacional	Privado	5	19.23	2.58
Seminario Internacional	Gobierno	4	15.38	02.06

Tabla 7: Cruce de Sector por Espacio de Participación.

Fuente: elaboración propia

La Tabla 7 realiza un cruce de datos entre representación sectorial, según la variable “Sector”, discriminada por espacio de participación. Contiene los datos de todos los espacios y fueron contemplados independientemente los casos donde hubo más de un sector por expositor. Por ejemplo, donde un expositor se presentó como afiliado a una organización de la sociedad civil y otra de la academia.

En los aportes escritos, el sector privado domina con un 35.24% de las contribuciones,

seguido por contribuciones individuales (28.57%) y de la sociedad civil (18.10%). La academia y el gobierno tienen una representación menor, con 12.38% y 5.71% respectivamente. En las audiencias públicas, la academia sobresale con un 41.27% de participación, destacando su rol en el debate. El sector privado también tiene una fuerte presencia con un 30.16%, seguido por la sociedad civil (20.63%) y el gobierno (7.94%). En los seminarios internacionales, la academia nuevamente lidera con un 42.31%, seguida por la sociedad civil (23.08%) y el sector privado (19.23%). La participación del gobierno es más significativa en este espacio, con un 15.38%.

Espacio	Interés	Cantidad	Porcentaje	Porcentaje Global
Aportes Escritos	Asociaciones del Sector Privado	32	30.48	16.49
Aportes Escritos	Individual	30	28.57	15.46
Aportes Escritos	Asociaciones Civiles Sin Fines de Lucro	19	18.10	9.79
Aportes Escritos	Academia	13	12.38	6.70
Aportes Escritos	Gobierno	6	5.71	03.09
Aportes Escritos	Sector Privado	5	4.76	2.58
Audiencia Pública	Academia	26	41.27	13.40
Audiencia Pública	Asociaciones Civiles Sin Fines de Lucro	13	20.63	6.70
Audiencia Pública	Sector Privado	12	19.05	6.19
Audiencia Pública	Asociaciones del Sector Privado	7	11.11	3.61
Audiencia Pública	Gobierno	5	7.94	2.58
Seminario Internacional	Academia	11	42.31	5.67
Seminario Internacional	Asociaciones Civiles Sin Fines de Lucro	6	23.08	03.09
Seminario Internacional	Sector Privado	4	15.38	02.06
Seminario Internacional	Gobierno	2	7.69	01.03
Seminario Internacional	Instituciones Internacionales	2	7.69	01.03
Seminario Internacional	Asociaciones del Sector Privado	1	3.85	0.52

Tabla 8: Cruce Intereses por espacio de participación.

Fuente: elaboración propia

La Tabla 8 realiza un cruce de datos entre representación sectorial, según la variable

“Intereses”, discriminada por espacio de participación. Contiene los datos de todos los espacios y fueron contemplados independientemente los casos donde hubo más de un valor declarado para la variable por expositor, al igual que en el cuadro anterior.

En los aportes escritos, las asociaciones del sector privado representan la mayor proporción (30.48%), seguidas por contribuciones individuales (28.57%) y asociaciones civiles sin fines de lucro (18.10%). La academia también tiene una presencia significativa (12.38%). En las audiencias públicas, la academia destaca con un 41.27% de participación, lo que muestra el rol que se le buscó dar en las exposiciones organizadas por el Senado. Las asociaciones civiles y el sector privado también tienen una representación considerable (20.63% y 19.05%, respectivamente). Por otro lado, el seminario internacional muestra una mayor participación de la academia (42.31%) y asociaciones civiles (23.08%), con una presencia menor del sector privado y el gobierno.

Ciudad	Cantidad	Porcentaje
São Paulo	20	40
Brasília	10	20
Rio de Janeiro	10	20
Belo Horizonte	2	4
Porto Alegre	2	4
Recife	2	4
Goiás	1	2
Maceió	1	2
Manaus	1	2
Viçosa	1	2

Tabla 9: Integrantes de las Audiencias Públicas discriminados por ciudad (sólo para Brasil)

Fuente: elaboración propia

Por otro lado, respecto de la distribución geográfica, podemos observar la Tabla 9, acerca del lugar de origen de las instituciones integrantes de las Audiencias y de los Aportes Escritos.

La Tabla 10 muestra las observaciones discriminadas por Ciudad, sólo para los participantes de Brasil. Reflejan una tendencia similar a la de las Audiencias Públicas, con São Paulo liderando (38.30%) y seguida por Brasília y Rio de Janeiro (13.83% cada una).

Ciudad	Cantidad	Porcentaje
São Paulo	36	38.30
Múltiple	14	14.89
Brasília	13	13.83
Rio de Janeiro	13	13.83
Sin datos de lugar	5	5.32
Curitiba	2	2.13
Paraná	2	2.13
Recife	2	2.13
Andaraí	1	01.06
Belo Horizonte	1	01.06
Fortaleza	1	01.06
Joaçaba	1	01.06
Porto Alegre	1	01.06
Río Grande del Norte	1	01.06
Teresina	1	01.06

Tabla 10: Contribuciones escritas discriminadas por ciudad (sólo para Brasil)

Fuente: elaboración propia 4.

4. Discusión

Durante las dos secciones anteriores analizamos, por un lado, el proceso participativo propiciado por la CJSUBIA en tanto su ubicación dentro de un proceso más amplio de debate sobre regulación de tecnología en Brasil y, por otro lado, su composición específica en tanto diversidad sectorial y geográfica. Ese análisis nos permite identificar potenciales lecciones aprendidas y desafíos persistentes para futuras instancias similares de participación. Un análisis más detallado de los datos presentados puede echar luz sobre este punto.

En ambas categorizaciones sectoriales, las contribuciones escritas del sector privado tienen la mayor preponderancia, hayan sido enviadas por las empresas mismas o a través de las cámaras y alianzas de colaboración sectorial. Del total de 105 registros para la variable, el 35% corresponde a ese sector. Esto es llamativo por el hecho de que el espacio de las contribuciones escritas es de orden abierto y proactivo por parte de la ciudadanía, a diferencia de las audiencias y el seminario que fueron resultado de una invitación y un programa armado por la Comisión. Los datos muestran, de esta manera, la alta proactividad por parte de las cámaras empresarias por manifestar su punto de vista para el proceso deliberativo para la regulación de la IA en Brasil.

Además, estos datos dan indicios de una significativa disponibilidad y uso de recursos dedicados por el sector privado a la participación en estos procesos, una vez que la formulación de aportes escritos requiere planificación y movilización interna. Aunque no contamos con datos más específicos en este análisis, sí podemos asumir que el sector dispone de recursos humanos y/o financieros para participar en consultas como la promovida por la CJSUBIA de forma diferenciada al sector académico o de sociedad civil. Unida a la movilización de agrupaciones

sectoriales, la evidencia de una inversión en participar en las discusiones regulatorias sobre el tema es un indicativa del peso estratégico que ve el sector privado en el caso de Brasil. Esto es visible para una multiplicidad de sectores de la economía identificable a partir de los autores de las contribuciones, entre otros:

- a) bancos,
- b) empresas fintech,
- c) representantes de la industria,
- d) grandes empresas de tecnología como Microsoft e IBM,
- e) empresas del software,
- f) cámaras del comercio,
- g) integrantes de estudios jurídicos,
- h) empresas aseguradoras, de las telecomunicaciones y del sector de la industria de la salud.

Respecto de la participación de la sociedad civil, podemos inferir que la situación es inversa. De los 103 aportes escritos recibidos, sólo 19 corresponden a la categoría Sociedad Civil. De ese total, 4 corresponden a la CDR como coalición, 3 a organizaciones internacionales y sólo 4 tienen un foco que excede la agenda de tecnología: el Instituto Igarapé, enfocado en temas de seguridad y desarrollo; el Instituto Alana, enfocado en la agenda medioambiental; la asociación de juristas negras y el Instituto Brasileiro de Defesa del Consumidor (IDEC).

Más allá de que exista un espacio abierto para las contribuciones escritas, no siempre es posible movilizarse para participar dada la limitación de recursos y la excesiva exposición a distintos tipos de urgencias en materia de protección de derechos a que suelen estar expuestas. La existencia de fondos de emergencia, como el sostenido por Derechos Digitales¹⁷, busca mitigar este tipo de desbalance al ofrecer

fondos de manera ágil para que las organizaciones se involucren de manera más significativa en estas discusiones. Sin embargo, son aún limitados y restringidos a un grupo de organizaciones más cercanas a la agenda digital. En el caso de organizaciones de base, organizaciones de derechos humanos y movimientos sociales, otras barreras, como el mismo conocimiento de estos espacios y procesos, pueden aplicarse, así como la existencia de múltiples agendas que compiten con la de regulación de IA, más allá de los impactos que estas tecnologías generan para las poblaciones y en las temáticas que ellas atienden. Estudios apoyados por Derechos Digitales han apuntado, en el caso de Brasil, el impacto del uso de IA en programas relacionados al empleo¹⁸ y la asistencia social¹⁹, pero no se ha visto mayor movilización de organizaciones involucradas en estos debates en la consulta de CJSUBIA.

Llamamos la atención también sobre el rol de las agrupaciones sectoriales para la participación en este proceso. Para el sector privado, el rol de las federaciones, asociaciones y cámaras tuvo el mayor peso: de los 61 registros de participación del sector privado, 40 fueron realizados a través de sus agrupaciones. De ese número, que representa más del 65% del total, 32 fueron presentadas a través de aportes escritos, vía email. Este dato puede mostrar la relevancia de los escritos, en tanto pueden permitir generar textos consensuados que representen la mirada de subsectores en sí mismos, y no de una empresa en particular. Cabe destacar la ausencia de escritos por parte de agrupaciones de trabajadores, sea en gremios, sindicatos o asociaciones, con una posible excepción de la asociación de juristas negras, en tanto agrupación profesional. Otra excepción, desde la sociedad civil, son los envíos de la CDR. Nuevamente, la capacidad de organización y movilización de las asociaciones del sector privado pueden tener que ver con el acceso diferencial a los recursos necesarios para viabilizar contribuciones conjuntas, además del

acceso a la información oportuna sobre el proceso y los mecanismos de participación.

Respecto de las audiencias públicas, puede verse una predominancia del sector académico (26 observaciones), con mayoría de instituciones de Sao Paulo (8) y de Río de Janeiro (6). Del total, la mayoría son instituciones universitarias (17), con menor representación de institutos y fundaciones. Segundo en el orden de representación sectorial, en las audiencias públicas, está el sector privado, con 19 representantes, la sociedad civil con 13 y el gobierno con sólo 5. Cabe destacar que dentro del sector privado permanecen las cámaras y asociaciones. Si desdoblamos la categorización en la propuesta de Belli et al., 12 participantes corresponden al sector privado y 7 a sus asociaciones. En este caso, en las audiencias, las cámaras empresariales no tienen tanto peso como en las contribuciones escritas.

Por último, respecto del seminario internacional, la preponderancia es de los representantes de la academia, con 11 observaciones, seguido por sociedad civil (6), sector privado (5) y gobierno (4). De los 26 expositores en el seminario internacional, la mitad provienen de Estados Unidos, 11 de países europeos y tan solo una observación para otros países de América Latina y otra para África. Mientras que los representantes de los Estados Unidos muestran una distribución equilibrada en términos sectoriales, los representantes de países europeos provienen principalmente de universidades (7 de 11), seguidos por representantes de la Comisión Europea (2). Este aspecto es de particular relevancia dada la influencia que tuvo en la CJSUBIA la discusión por la Ley de Inteligencia Artificial en el parlamento de la Unión Europea²⁰, que se dio en paralelo al proceso analizado. La perspectiva africana, si bien estuvo representada por una sola organización, se trata de la Africa Digital Rights' Hub, que cuenta con una mirada regional. El seminario no contó con ningún expositor del continente asiático.

El análisis conjunto de la prevalencia de participación académica en las audiencias y el seminario, en este caso, puede apuntar a un intento desde la Comisión de priorizar visiones “expertas”. En su amplia mayoría, los representantes del sector académico provienen del ámbito del Derecho. De las 37 exposiciones y contribuciones de la academia en estos espacios, 28, más del 75%, provienen de profesionales especializados en estudios de innovación, tecnología y las comunicaciones, con excepciones en especialistas en temas de diversidad, estudios raciales y otras ramas de las ciencias sociales, como la ciencia política. Dentro de las ramas del derecho sí podemos encontrar una moderada diversidad, con especialistas en temas como derecho administrativo, regulatorio, derecho comparado, civil, autoral, entre otros.

En términos de las contribuciones escritas en tanto espacio de participación online, es elocuente el peso de las categorizadas como “individuales”, una potencial forma de participación más allá de la afiliación institucional. Estas contribuciones fueron enviadas por correo electrónico, según instrucciones del Senado, y luego publicadas en la web de la CJSUBIA. La categorización surge del trabajo de clasificación realizado por el personal de la Comisión. Una mirada hacia dentro de las contribuciones individuales muestra, de forma contraria a lo esperado, una mayoría de envíos de personas afiliadas a instituciones de distintos sectores, lo que puede derivar en un sobregistro de esta categoría.

4.1 Paralelismos con el Comité de Gobernanza de la EBIA

Es propicio comparar la diversidad sectorial de los espacios participativos de la CJSUBIA con los que tuvieron lugar en el

marco de las reuniones del comité del Comité de Gobernanza de la Estrategia Brasileña de Inteligencia Artificial (EBIA). Según los artículos 2 y 3 de su reglamento interno²¹, el Comité de Gobernanza de EBIA está compuesto por un representante y un suplente de cada uno de los siguientes organismos e instituciones: Ministerio de Ciencia, Tecnología e Innovación (MCTI), la Red de Innovación en Inteligencia Artificial MCTI/EMBRAPII; e "Instituciones invitadas"; y es presidido por el Secretario de Emprendimiento e Innovación del MCTI. Tiene la función de implementar los ejes de trabajo mencionados y gestionar todo lo referido a las actividades enmarcadas en la Estrategia. Como menciona Eduardo Magrani (2021, p. 10), el texto de la Resolución 4617/2021, que le da entidad, no incluye información sobre estructuras de gobernanza concretas, aunque existía un arreglo institucional en ese sentido. El Comité de Gobernanza se crea en una segunda instancia. En el mencionado artículo de Belli et al., sus autores analizan 4 de las reuniones que tuvieron lugar durante 2021, comenzando en mayo con el primer encuentro. Esta primera reunión contó con una mayoría significativa de representantes del gobierno, con un 42% (34 personas), seguida por un 24,7% de representantes de la categoría Asociaciones del Sector Privado, con 20 personas.

Los autores de esta investigación afirman que la tendencia de esa primera reunión se mantiene en el resto, y que puede detectarse una “representación desigual” de sectores importantes como la academia y organizaciones de la sociedad civil no ligadas a corporaciones (op. cit., p.8-9). Luego, afirman que esta mayoría no es sorprendente dado que estos encuentros eran organizados por el Ministerio de Ciencia, Tecnología e Innovación del Gobierno Federal, quienes presiden y gestionan el Comité. Además, afirman que la “consistente mayoría” de la categoría Asociaciones del Sector Privado, por sobre representantes directos del sector

privado, puede traducirse en una mirada que evite el análisis más profundizado de los desafíos de la implementación de la IA.

La preponderancia de Asociaciones del Sector Privado en la CJSUBIA sólo se manifiesta, en cambio, en los aportes escritos, como fue mencionado. Los espacios organizados por el Senado, tanto las audiencias como el seminario, tienen mayoría de representantes de la academia y de la categoría “Asociaciones Civiles Sin Fines de Lucro”. Esto marca una diferencia sustantiva con el análisis de Belli et al., donde el gobierno, organizador del evento, tuvo una marcada mayoría de representantes. Representantes del sector gubernamental, e incluso del Estado en su conjunto, fueron una marcada minoría durante el proceso de la CJSUBIA lo que parece indicar una efectiva disposición de escucha social por parte de quienes lideraron el proceso.

5. Conclusiones

En esta última sección, señalaremos algunas conclusiones que pueden surgir del proceso de la CJSUBIA, que esperamos puedan ser útiles para otros países de la región.

Respecto de la diversidad sectorial en la participación, podemos ver que, aún con más de una década de procesos participativos, una diversidad plena sigue pendiente. Los datos muestran un desbalance entre los distintos sectores, con una manifiesta subrepresentación de las organizaciones de la sociedad civil, sean organizaciones sin fines de lucro, organizaciones sociales de base, sindicatos, asociaciones profesionales, entre otras. Pero también con una ausencia de ciudadanos que no integran ninguna organización y puedan aportar elementos que permitan “romper la burbuja” del pensamiento de la comunidad de trabajo sobre tecnología y sociedad. Es que, como pudimos ver, el de

la CJSUBIA fue un proceso integrado mayoritariamente por participantes de la comunidad del estudio y desarrollo de la tecnología y las comunicaciones.

Más que sobre la ausencia de espacios, el caso brasileño de CJSUBIA reflexionar sobre las condiciones habilitantes para la participación de sociedad civil en este tipo de procesos en la actualidad. En términos de recursos, podemos hablar de una condición necesaria pero no suficiente para garantizar la participación. Los fondos y los recursos humanos limitados de, por ejemplo, las organizaciones sin fines de lucro implican un obstáculo de suma importancia a la hora de participar de procesos que, en muchos casos, surgen de manera inesperada y sin una anticipación que permita una planificación acorde. Sin embargo, contar con los recursos no alcanza para garantizar una participación plural. Es necesario disponer de una difusión amplia que dé cuenta de la existencia de estos procesos y, por otro lado, el impulso a la participación de grupos que pueden verse afectados de forma directa por este tipo de políticas. Como fue mencionado, este punto es de atención a la hora de convocar organizaciones de base, organizaciones de derechos humanos y movimientos sociales, que no necesariamente

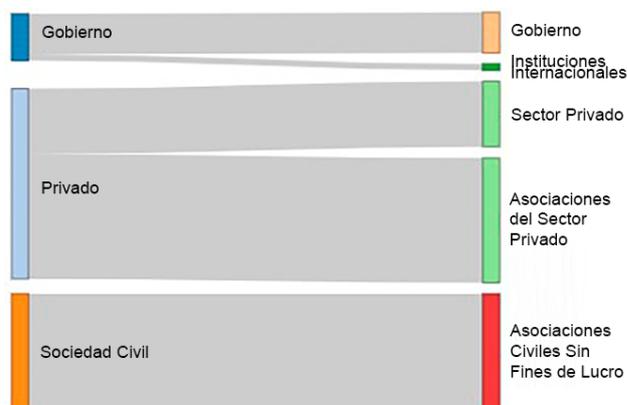


Gráfico 1: Derivación de observaciones de la variable “Sector” a la variable “Intereses”

Fonte: Elaborado pela autora (2024)

participan o siguen estrictamente el desarrollo de la agenda de la tecnología y las comunicaciones, pero pueden realizar aportes sustantivos para las estrategias de regulación.

Los puntos de mejora mencionados nos permiten trazar aprendizajes a partir de la experiencia y el diseño de la CJSUBIA en tanto espacio de participación, principalmente a la luz de su manifiesta vocación de diversidad. La mera apertura de los espacios no fue suficiente para incluir una distribución equitativa de actores y sectores que pueden ser impactados por la implementación de tecnologías basadas en IA. Otro elemento para tomarse en consideración es relativo a la publicidad de la información relacionada con la CJSUBIA. Si bien la información es abundante, es propicia la implementación de formatos abiertos con datos estructurados que permitan su reprocesamiento. Esto, además de una desagregación que permita comprender mejor la procedencia de las personas y organizaciones participantes. Disponer de información desagregada y accesible para su análisis es necesario para poder comprender mejor estos procesos y generar investigaciones basadas en datos e información oficial. En la actualidad, hay múltiples puntos de análisis que necesitan ser profundizados, lo que abre espacios de oportunidad para la investigación futura.

Referências bibliográficas

- AAVV. (2023). *Declaración de Santiago para promover una Inteligencia Artificial (IA) ética en América Latina y el Caribe*. https://minciencia.gob.cl/uploads/filer_public/40/2a/402a35a0-1222-4dab-b090-5c81bbf34237/declaracion_de_santiago.pdf
- Affonso Souza, C., Steibel, F., & Lemos, R. (2017). Notes on the creation and impacts of Brazil's Internet Bill of Rights. *The Theory and Practice of Legislation*, 5(1), 73-94. <https://doi.org/10.1080/20508840.2016.1264677>
- Almeida, V., Getschko, D., & Afonso, C. (2015). The Origin and Evolution of Multistakeholder Models. *IEEE Internet Computing*, 19(1), 74-79. *IEEE Internet Computing*. <https://doi.org/10.1109/MIC.2015.15>
- Avritzer, L. (with Negrete Jiménez, M., & Sánchez Fuentes, D. L.). (2010). *Las instituciones participativas en el Brasil democrático* (Primera edición, UV, Xalapa, 27 de octubre de 2010). : Universidad Veracruzana.
- Belli, L., Curzi, Y., & Gaspar, W. B. (2023). AI regulation in Brazil: Advancements, flows, and need to learn from the data protection experience. *Computer Law & Security Review*, 48, 105767. <https://doi.org/10.1016/j.clsr.2022.105767>
- Bioni, B., & Marques Rielli, M. (2021). A CONSTRUÇÃO MULTISSETORIAL DA LGPD: história e aprendizados. *En Proteção de dados: Contexto, narrativas e elementos fundantes*.
- Bragatto, R. C., Sampaio, R. C., & Nicolás, M. A. (2015). A segunda fase da consulta do marco civil da internet: como foi construída, quem participou e quais os impactos? *Eptic Online: Revista Eletrônica Internacional de Economia Política da Informação, da Comunicação e da Cultura*, 17(1), 236-255.
- CAF. (2023). *CAF promueve iniciativa de gobernanza de Inteligencia Artificial*. <https://www.caf.com/es/actualidad/noticias/2023/10/caf-promueve-iniciativa-de-gobernanza-de-inteligencia-artificial-en-america-latina-y-el-caribe/>
- CETIC.br. (2010). *TIC Domicílios e Empresas 2010 Pesquisa sobre o Uso das Tecnologias de Informação e Comunicação no Brasil*. <https://cetic.br/media/docs/publicacoes/2/tic-domicilios-e-empresas-2010.pdf>
- CJSUBIA. (2022). *Relatorio Final*.
- Coalizão Direitos na Rede. (2021). *Inteligência Artificial não pode ser regulada a toque de caixa*. <https://direitosnarede.org.br/2021/09/23/inteligencia-artificial-nao-pode-ser-regulada-a-toque-de-caixa/>
- COE. (2024a). Council of Europe adopts first international treaty on artificial intelligence—Portal—Www.coe.int. Portal. <https://www.coe.int/en/web/portal/-/council-of-europe-adopts-first-international-treaty-on-artificial-intelligence>
- COE. (2024b). *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*. [https://search.coe.int/cm/#{%22CoEObjectId%22:\[%220900001680afb11f%22\],%22sort%22:\[%22CoEValidationDate%20Descending%22}}](https://search.coe.int/cm/#{%22CoEObjectId%22:[%220900001680afb11f%22],%22sort%22:[%22CoEValidationDate%20Descending%22}})

- De Souza, C. A., & Perrone, C. (2023). Brazil and multistakeholder diplomacy for the Internet: Past achievements, current challenges and the road ahead. En I. Johnstone, A. Sukumar, & J. Trachtman (Eds.), *Building an International Cybersecurity Regime* (pp. 220-237). Edward Elgar Publishing. <https://doi.org/10.4337/9781035301546.00017>
- Dilhac, M.-A., Abrassart, C., & Voarino, N. (2018). *Report of the Montréal Declaration for a responsible development of artificial intelligence*. <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/27795>
- Hernández, L., Canales, M. P., & De Souza, M. R. (2022). *Inteligencia Artificial y participación en América Latina: Las estrategias nacionales de IA*. Derechos Digitales.
- Intervezes. (2019, abril 15). Nota: Bolsonaro liquida conselhos e reduz participação social. <https://intervezes.org.br/nota-bolsonaro-liquida-conselhos-e-reduz-participacao-social/>
- Irisbh. (2022). *Inteligência artificial no Brasil: O episódio da Comissão de Juristas*. <https://irisbh.com.br/inteligencia-artificial-no-brasil-o-episodio-da-comissao-de-jurist>
- Lara, J.C. (2024) “Inteligencia Artificial en América Latina: Regulación no significa protección”. Disponible en: <https://www.derechosdigitales.org/23866/inteligencia-artificial-en-america-latina-regulacion-no-significa-proteccion/>
- Magrani, E., Pragmático, M., & Gaspar, W. (2021). *Estrategia Brasileña de Inteligencia Artificial: Comentarios sobre la Resolución 4.617/2021 del Ministerio de Ciencia, Tecnología e Innovación (MCTI)*. <https://www.empatia.la/wp-content/uploads/2021/12/ART.-DE-OPINION-Brasil-ES-final.pdf>
- Martinez Stefan, A. G., & Vaz, J. C. (2016). The Brazilian Civil Rights Framework For The Internet And The Lessons Learned About The Capability Of Brazilian Public Authorities To Promote Citizen Participation Through The Internet. En *TIC Governo Eletrônico 2015 Pesquisa sobre o Uso das Tecnologias de Informação e Comunicação no Setor Público Brasileiro*.
- Nicolás, M. A., Sampaio, R. C., Bragatto, R., & Konopacki, M. A. (2017). A primeira fase da consulta pública da regulamentação do Marco Civil da Internet: estrutura comunicativa, limites e contribuições. *Contemporanea*, 15(2). <https://doi.org/10.9771/contemporanea.v15i2.22161>
- ONU. (2023). *AI Advisory Body Interim Report: Governing AI for Humanity*. United Nations; United Nations. <https://www.un.org/en/ai-advisory-body>
- ONU. (2024). *Resolución 78/265. Aprovechar las oportunidades de sistemas seguros y fiables de inteligencia artificial para el desarrollo sostenible*. <https://documents.un.org/doc/undoc/gen/n24/087/86/pdf/n2408786.pdf?token=YqVivWfXBQFg0mFkYJ&fe=true>
- Portal da Câmara dos Deputados. (2021). Câmara aprova regime de urgência para projeto sobre inteligência artificial. *Portal da Câmara dos Deputados*. <https://www.camara.leg.br/noticias/782074-camara-aprova-regime-de-urgencia-para-projeto-sobre-inteligencia-artificial/>
- Velasco, P. (2021) “La política nacional de Inteligencia Artificial chilena y su proceso de participación ciudadana”. *Derechos Digitales* <https://www.derechosdigitales.org/17010/la-politica-nacional-de-inteligencia-artificial-chilena-y-su-proceso-de-participacion-ciudadana/>
- UNESCO. (2021). *Recomendación sobre la Ética de la Inteligencia Artificial*.

Notas finais

1 Ver <https://www25.senado.leg.br/web/atividade-materia/151547>. Visitado en julio 2024.

2 Ver <https://direitosnarede.org.br/2021/09/23/inteligencia-artificial-nao-pode-ser-regulada-a-toque-de-caixa/>. Visitado en julio 2024.

3 Ver <https://www25.senado.leg.br/web/atividade-materias/-/materia/138790>. Visitado en julio 2024.

4 Ver <https://www25.senado.leg.br/web/atividade-materias/-/materia/147434>. Visitado en julio 2024.

5 Ver <https://direitosnarede.org.br/carta-aberta-regulacao-ia/>. Visitado en junio de 2024.

6 Ver <https://legis.senado.leg.br/comissoes/audiencias?codcol=2504>. Visitado en julio 2024.

7 Ver <https://legis.senado.leg.br/comissoes/mnas?codcol=2504&tp=4>. Visitado en julio 2024.

8 Disponibles en: <https://legis.senado.leg.br/comissoes/arquivos?ap=6916&codcol=2504>. Consultado en Julio 2024.

9 Ver <https://direitosnarede.org.br/quem-somos>. Visitado en julio 2024.

10 Ver <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#adherents>. Visitado en julio 2024.

11 Ver https://antigo.mctic.gov.br/mctic/open-cms/legislacao/portarias/Portaria_MCTI_n_4617_de_06042021.html. Visitado en julio 2024.

12 Ver <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/inteligencia-artificial>. Visitado en julio 2024.

13 Ver <https://www25.senado.leg.br/web/atividade-materias/-/materia/151547>. Visitado en julio 2024.

14 Disponible en https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=2068016&filename=Tramitacao-PL%2021/2020. Consultado en julio 2024.

15 Ver <https://legis.senado.leg.br/comissoes/comissao?codcol=2504>. Consultado en julio 2024.

16 Para más información: <https://www12.senado.leg.br/tv/quem-somos>. Consultado en Julio de 2024.

17 Ver <https://www.derechosdigitales.org/19245/fondo-de-respuesta-rapida-tres-anos-de-asistencia-inmediata-para-la-defensa-y-proteccion-de-los-derechos-digitales>.

18 Ver <https://ia.derechosdigitales.org/casos/brasil-sistema-nacional-de-empleo>.

19 Ver <https://ia.derechosdigitales.org/casos/brasil-auxilio-de-emergencia>.

20 Ver <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>.

21 Disponible en https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia_regimento_interno_2021_09_30.pdf.

PRODUÇÃO ARTÍSTICA

A imagem na inteligência artificial

Joana Resek

joanaresek@gmail.com

Graduada em Design Gráfico pela
Escola de Design da UEMG em Belo Horizonte.
Possui dois títulos de Mestrado em Arte,
Design e Novas Tecnologias —
um pela Universidade Paris 8 Saint Denis,
em Paris, França, e outro pelo
Media Center of Art/ESDi,
em Barcelona, Espanha.

Técnica

Imagem digital

Ano

2024

Resumo

O tema central escolhido para a busca da imagem foi "*crowd of people walking*". A imagem foi obtida de um banco de imagens de uso gratuito e serviu como ponto de partida para a composição da capa, explorando a interação entre a sociedade urbana e a inteligência artificial. Em um segundo momento, a imagem foi manipulada utilizando um programa de inteligência artificial, com o comando "*technical risk for system collapse*". Essa frase foi extraída de um dos artigos publicados nesta edição da revista. O processo de manipulação gerou diversas variações de cores e estilos. Por fim, foi realizada uma montagem, a fim de alcançar uma imagem que refletisse o conceito e o ritmo da edição. Este texto foi elaborado com o apoio do ChatGPT.



Imagem 1: Imagem inicial para o desenvolvimento do projeto de capa.

Créditos: Banco de imagens gratuito

Imagem 2: Variações de imagens geradas por inteligência artificial, tendo a imagem 1 como referência.

Créditos: Inteligência artificial

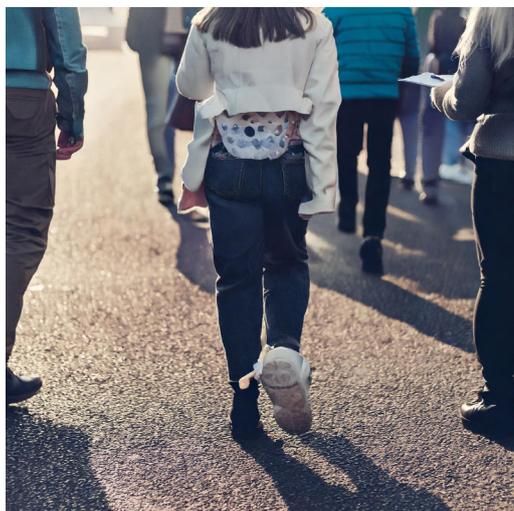




Imagem 4: Montagem e ajuste final com manipulação de elementos.

Fuente: elaboração propia.

