

Questões conjunturais sobre a regulação da IA

Conjunctural issues in AI regulation

Cuestiones coyunturales sobre la regulación de la IA

Sérgio Amadeu da Silveira^{1,a}

sergio.amadeu@ufabc.edu.br | <https://orcid.org/0000-0003-1029-9133>

¹ Universidade Federal do ABC, Centro de Engenharia, Modelagem e Ciências Sociais Aplicadas. São Bernardo do Campo, SP, Brasil.

^a Doutorado em Ciência Política pela Universidade de São Paulo.

RESUMO

O texto parte da disputa em torno da regulação da inteligência artificial (IA) realmente existente, dividida entre dois polos, o da proteção da inovação e o da necessidade de mitigar riscos sociais e ambientais. Avança para apontar a colossal concentração de poder computacional e de dados nas mãos dos oligopólios digitais, as *big techs*. Levanta preocupações sobre o impacto dessa concentração na soberania digital e na capacidade do país em controlar suas infraestruturas e dados. Argumenta como a falta de transparência e explicabilidade nos sistemas de IA agrava os riscos de discriminação e exclusão social. Ainda, defende que a regulação da IA deve garantir direitos, salvaguardar a soberania e promover um ecossistema digital autônomo e inclusivo.

Palavras-chave: Regulação da IA; Sistemas automatizados; Transparência; Explicabilidade; Soberania digital e de dados.

ABSTRACT

The text begins by addressing the debate surrounding the regulation of existing artificial intelligence (AI), divided between two poles: the protection of innovation and the need to mitigate social and environmental risks. It progresses to highlight the colossal concentration of computational power and data in the hands of digital oligopolies, the big tech companies. The text raises concerns about the impact of this concentration on digital sovereignty and the country's ability to control its infrastructures and data. It argues that the lack of transparency and explainability in AI systems exacerbates the risks of discrimination and social exclusion. The text advocates that AI regulation should ensure rights, safeguard sovereignty, and promote an autonomous and inclusive digital ecosystem.

Keywords: AI regulation; Automated systems; Transparency; Explainability; Digital and data sovereignty.

RESUMEN

El texto parte de la disputa en torno a la regulación de la inteligencia artificial (IA) existente, dividida entre dos polos: la protección de la innovación y la necesidad de mitigar riesgos sociales y ambientales. Avanza para señalar la colosal concentración de poder computacional y de datos en manos de los oligopolios digitales, las *big tech*. Plantea preocupaciones sobre el impacto de esta concentración en la soberanía digital y en la capacidad del país para controlar sus infraestructuras y datos. Argumenta que la falta de transparencia y explicabilidad en los sistemas de IA agrava los riesgos de discriminación y exclusión social. El texto defiende que la regulación de la IA debe garantizar derechos, salvaguardar la soberanía y promover un ecosistema digital autónomo e inclusivo.

Palabras clave: Regulación de la IA; Sistemas automatizados; Transparencia; Explicabilidad; Soberanía digital y de datos.

INFORMAÇÕES DO ARTIGO

Contribuição dos autores: o autor é responsável por todo o texto.

Declaração de conflito de interesses: não há.

Fontes de financiamento: bolsista produtividade CNPq-2.

Considerações éticas: não há.

Agradecimentos/Contribuições adicionais: ao meu grupo de pesquisa.

Histórico do artigo: submetido: 2 set. 2024 | aceito: 4 set. 2024 | publicado: 30 set. 2024.

Apresentação anterior: não há.

Licença CC BY-NC atribuição não comercial. Com essa licença é permitido acessar, baixar (*download*), copiar, imprimir, compartilhar, reutilizar e distribuir os artigos, desde que para uso não comercial e com a citação da fonte, conferindo os devidos créditos de autoria e menção à Reciis. Nesses casos, nenhuma permissão é necessária por parte dos autores ou dos editores.

INTRODUÇÃO ÀS DISPUTAS

Há uma disputa acirrada a respeito da necessidade e importância da regulação da IA entre pelo menos dois polos de pensamento. Um deles indica que a medida pode atrasar e até mesmo bloquear a inovação. Esse argumento é muito utilizado quando legisladores e gestores públicos querem regular as tecnologias da informação, não somente a IA. O outro polo afirma que é preciso atuar preventivamente sobre o desenvolvimento e o uso socialmente perigosos das tecnologias, mitigar suas tendências inaceitáveis, agir sobre os riscos e prejuízos delas decorrentes, prováveis ou conhecidos, para segmentos da sociedade, o meio ambiente ou as democracias.

As origens teóricas dessa disputa são múltiplas. A própria ciência econômica utiliza a expressão **externalidade negativa** para descrever os impactos nefastos, muitas vezes não esperados, de uma atividade econômica, industrial ou não. Arthur Pigou (2017) disseminou a noção de externalidades, que teve efeitos importantes nos debates econômicos e, posteriormente, socioambientais. A poluição e a contaminação da água não são aceitáveis, a necessidade de redução das emissões de CO₂ e outros poluentes tem gerado legislações e ações regulatórias, com restrições e proibições sobre a produção e uso de determinados materiais, entre outros exemplos.

Em geral, os argumentos das grandes empresas contra medidas regulatórias visam, obviamente, assegurar um espaço aberto e irrestrito para que possam experimentar produtos e serviços altamente lucrativos, sem considerar suas implicações sociais e políticas. Quando discutimos a IA, ninguém tem dúvida de seu potencial de automatização e de ampliação da velocidade de atividades cognitivas, mas, para desenvolver modelos, precisamos de dados que são gerados, armazenados e processados em gigantescas infraestruturas com diversos efeitos extremamente preocupantes. Esses dados são tratados por sistemas algorítmicos invisíveis e imperceptíveis para a maioria das pessoas, com finalidades desconhecidas ou parcialmente ofuscadas (Pasquale, 2015).

O cientista político Langdon Winner, ao debater situações em que as tecnologias interferem nas relações de poder, trouxe uma importante perspectiva do historiador Lewis Mumford. Este relatou que, desde tempos neolíticos, dois tipos de tecnologia têm coexistido, as autoritárias e democráticas. As primeiras seriam mais poderosas, centralizadoras e instáveis, e as outras, mais flexíveis, frágeis e duradouras (Mumford, 1964 *apud* Winner, 1986). É preciso destacar que o modo como as *big techs* estão desenvolvendo a chamada IA realmente existente tende a reforçar o alerta de Mumford. Estamos vivendo uma concentração tamanha de insumos e de poder computacional para o desenvolvimento dos modelos da nomeada IA generativa, que é preciso colocar em dúvida seu potencial democrático.

A CONCENTRAÇÃO DE DADOS E DO PODER COMPUTACIONAL

As abordagens do aprendizado de máquina e do aprendizado profundo são dominantes na chamada IA realmente existente. Elas visam realizar classificações, extrair padrões e efetuar previsões a partir de dados. A operação de fazer o modelo algorítmico captar detalhes e minúcias dos dados para extrair padrões e cumprir sua finalidade tem sido chamada de treinamento. Esse processo pode ser refinado, repetido e, ainda, reforçado pela intervenção humana. Bases de dados gigantescas estão sendo preparadas para treinar modelos que utilizam um número grande de computadores com alta capacidade de processamento. O treinamento de um modelo simples pode ser obtido com hardware comum, mas são significativos os obstáculos para a implantação de um LLM (Large Language Models), como o GPT, o Gemini, o LLaMA e outros modelos generativos.

Os dados são o insumo fundamental da IA realmente existente, que pode ser resumida como sistemas automatizados que extraem padrões e realizam classificações com algoritmos estatísticos e probabilísticos

que rodam em máquinas de alto processamento. Bases de dados variadas e colossais exigem também gigantescas infraestruturas computacionais. A Microsoft desenvolveu e lançou, em fevereiro de 2020, o Turing Natural Language Generation (T-NLG), considerado então o maior modelo de linguagem, com 17 bilhões de parâmetros (Sterling, 2020). É preciso destacar que parâmetros são os elementos ajustáveis do modelo que são aprendidos durante o treinamento, incluindo pesos e vieses, por exemplo, se estivermos operando as várias camadas de uma rede neural. Assim, os parâmetros vão determinar como os dados de entrada são transformados em saídas que contêm as previsões.

O GPT-3, lançado em junho de 2020, foi treinado com 175 bilhões de parâmetros. Seu treinamento pode ter sido realizado por diversas semanas, com computadores que possuem alto poder computacional.

O treinamento do GPT-3 da Open AI custou mais de US\$ 4 milhões, enquanto o treinamento do modelo LLaMA da Meta levou 1 milhão de horas de GPU e custou mais de US\$ 2,4 milhões. Treinar um modelo de IA consome energia; pode gastar mais eletricidade do que 100 residências típicas dos EUA durante um ano inteiro. O treinamento do GPT3 da Open AI consumiu 1.287 MWh de eletricidade e emitiu mais de 550 toneladas de dióxido de carbono (MIT, 2023, p. 11, tradução nossa).

O desenvolvimento de modelos chamados LLMs (Large Language Models) está se tornando cada vez mais caro. Conforme as estimativas do AI index 2024, os custos de treinamento de modelos de IA de última geração são uma barreira quase intransponível para quem não for do seleto grupo das *big techs*. O treinamento do GPT-4, da OpenAI, foi estimado em US\$ 78 milhões em computação, enquanto somente o treinamento em petaflop/s-day¹ do Gemini Ultra, do Google, custou US\$ 191 milhões (Maslej *et al.*, 2024, p. 5).

A infraestrutura para armazenar dados, para treinar e depois rodar os modelos da IA generativa ou de outros tipos da chamada IA realmente existente estão concentradas em poucas empresas gigantescas, os oligopólios digitais das *big techs*. Um problema não menos grave vem do controle que esses oligopólios estão obtendo no terreno da pesquisa científica da IA. Em 2023, as empresas produziram 51 dos modelos notáveis² de aprendizado de máquina. Já as universidades desenvolveram apenas 15 modelos. As parcerias entre empresas e academia foram responsáveis por 21 modelos notáveis. Quando observamos os países que sediam as organizações que lideraram, em 2023, o desenvolvimento desses modelos, chegamos ao seguinte cenário: “61 modelos notáveis de IA se originaram de instituições sediadas nos EUA, superando em muito os 21 da União Europeia e os 15 da China” (Maslej *et al.*, 2024, p. 30).

A INDISSOCIAÇÃO DA TRANSPARÊNCIA E DA EXPLICABILIDADE

Há um controle desproporcional sobre a criação, coleta e análise de dados. Os grupos Alphabet e Meta possuem mais informações sobre as sociedades e nações do que qualquer instituição governamental. O controle dos dados, a concentração dos negócios de provimento de nuvem e o enorme poder computacional das *big techs* gera o efeito de ampliação da concentração de dados, que implica, por sua vez, a concentração de poder econômico e computacional. Desse modo, Frank Pasquale (2015) alertou para o fato de que essas empresas estão construindo um poder de moldar opiniões, modular as atenções e interferir nas decisões políticas.

1 Flops (Floating Point Operations Per Second) é uma medida que indica a quantidade de operações matemáticas que um computador ou sistema pode realizar por segundo. Quando falamos de petaflops, estamos nos referindo a quadrilhões (10^{15}) de operações por segundo. Assim, a expressão petaflop/s-day se refere ao total de operações que um sistema pode realizar ao longo de um dia, considerando uma taxa constante de petaflops por segundo. Em outras palavras, ele mede a quantidade total de trabalho computacional (em termos de operações matemáticas) que uma máquina ou conjunto de máquinas realiza durante 24 horas.

2 Modelos notáveis de aprendizado de máquina, conforme descrito no relatório, referem-se a modelos influentes no ecossistema de IA/aprendizado de máquina. São selecionados com base em critérios como avanços no estado da arte, significância histórica ou altas taxas de citação. Esses modelos representam o que há de mais avançado em pesquisa e desenvolvimento de IA, e são frequentemente considerados por introduzir novas arquiteturas significativas, demonstrar capacidades impressionantes ou obter resultados de ponta.

Atualmente, não existe um só país em que as redes de relacionamento on-line não sejam relevantes. Essas plataformas de interação e comunicação são operadas por sistemas algorítmicos, em geral, por redes neurais artificiais. Trata-se de um tipo específico de aprendizado profundo, que coleta dados de cada ação de seus usuários e prepara o ambiente de atração de seu olhar para anúncios ou conteúdos definidos por modelos de IA, que são treinados para dar lucros para as plataformas digitais. Esses sistemas automatizados que gerenciam as redes sociais on-line são completamente invisíveis para as pessoas que as utilizam.

As buscas feitas no Google, as perguntas ao ChatGPT e ao Gemini geram dados e metadados que são analisados pelos controladores desses sistemas e plataformas. Eles alimentam ou retroalimentam os modelos e permitem seu aprimoramento. Na realidade, esses sistemas não transparentes são regidos por determinações específicas e executam procedimentos de cujas finalidades não temos a menor ideia, para além do jargão “queremos melhorar sua experiência”.

No cenário digital, cada vez mais operado pela IA realmente existente, temos inúmeras dificuldades para garantir a transparência e a explicabilidade dos sistemas algorítmicos. Por mais que a transparência apareça como um princípio nos documentos e legislações sobre IA, ela não tem sido executada. A explicabilidade – o princípio que reivindica que os passos e fundamentos das decisões algorítmicas sejam conhecidos e compreensíveis para as pessoas – está bem longe de ser alcançada.

O GPT-4 tem seu código fechado, já o PaLM2 (Alphabet) e o Llama2 (Meta) têm o código aberto. O número de projetos relacionados à IA no GitHub³ saltou de 845, em 2011, para aproximadamente 1,8 milhões, em 2023 (Maslej *et al.*, 2024, p. 30). Isso significa que a maioria dos projetos de chats e modelos de linguagem são de código-fonte aberto. Entretanto, quando são utilizados com novas bases de dados para seu treinamento, eles criam um novo modelo, que, em geral, é fechado. Além disso, quando se trata de uma rede neural artificial, os passos dados para chegar a uma predição não são conhecidos nem mesmo por quem organizou o experimento. As *big techs* têm utilizado modelos abertos de base do aprendizado profundo para atrair os desenvolvedores e o desenvolvimento para suas infraestruturas digitais.

O debate sobre a necessidade e a possibilidade da transparência e explicabilidade na IA é complexo. O professor de informática biomédica Nigam Shah, da Universidade de Stanford, defende que os modelos de IA não precisam ser interpretáveis se forem úteis e testados: “Dos 4.900 medicamentos prescritos rotineiramente, não sabemos completamente como a maioria deles realmente funciona” (Miller, 2021). Porém, Shah declarou que, em contextos nos quais direitos são decididos, tais como empréstimos, seleção de empregos, programas de assistência médica ou moradia, a interpretabilidade é indispensável (Miller, 2021).

As democracias necessitam de transparência. A gestão democrática e republicana do Estado também, mas, diante de um avanço no aprendizado profundo dos sistemas automatizados no setor público, seria importante garantir que os procedimentos que incidem e definem os direitos dos cidadãos e cidadãs sejam transparentes e, além disso, explicáveis. Saber os motivos pelos quais um benefício foi ou não concedido é imprescindível. Não existe democracia baseada na opacidade. Dada a complexidade dos sistemas automatizados, muitas vezes o direito à transparência e à explicabilidade deve vir acompanhado da ideia de auditabilidade plena, ou seja, de os grupos sociais recorrerem a especialistas que possam dizer o que um sistema busca, como e com que dados age.

3 O GitHub é uma plataforma on-line que permite o armazenamento, o compartilhamento e o gerenciamento de código-fonte de projetos de software. Utiliza o sistema de controle de versão Git, que permite que múltiplos desenvolvedores trabalhem simultaneamente em um projeto, acompanhem mudanças no código, colaborem e integrem suas contribuições de maneira organizada. Em resumo, o GitHub é utilizado para desenvolver softwares de forma colaborativa.

A DISCRIMINAÇÃO ÉTNICO-RACIAL, DE GÊNERO E CLASSE

Os sistemas automatizados têm um grande poder discriminatório. Poucos pesquisadores ainda duvidam disso. Algoritmos opacos com definições e finalidades segregacionistas, bases de dados racializadas, podem gerar treinamentos de modelos que não apenas reproduzem preconceitos, mas também os ampliam. Os pesquisadores Matteo Pasquinelli e Vladan Joler (2020) alertam para a ideia de que as limitações da IA são frequentemente discutidas em termos de vieses, como os de gênero, raça, classe e outras discriminações. Eles demonstram que os sistemas algorítmicos são construções sociotécnicas que deixam evidente que a neutralidade está longe de ser encontrada.

Pasquinelli e Joler diferenciam o viés histórico, o de conjuntos de dados, e o algorítmico. Esses vieses ocorrem em diferentes fases do processamento de informações. O histórico refere-se às desigualdades, preconceitos e discriminações presentes na sociedade antes mesmo de qualquer intervenção tecnológica. Os pesquisadores argumentam que, quando esses vieses são integrados em tecnologias que aparentam ser neutras, como algoritmos de IA, eles podem ser naturalizados e perpetuados.

Por sua vez, o viés de conjuntos de dados ocorre no processo de treinamento dos algoritmos de aprendizado de máquina. A base de dados pode não representar de forma adequada a diversidade e a complexidade da realidade que se pretende modelar. O viés pode ser introduzido de várias maneiras, como na escolha dos dados e na forma como eles são coletados, organizados ou rotulados pelos operadores humanos.

Já o viés algorítmico amplifica os outros dois, devido à compressão de informações pelos algoritmos, o que pode resultar em discriminação e perda de diversidade cultural. Além disso, Pasquinelli e Joler (2020) tratam do problema da “caixa-preta”, segundo o qual a complexidade dos algoritmos torna suas operações opacas, alimentando a percepção de que “a IA é um sistema alienígena e fora de controle, o que contribui para narrativas conspiratórias” (Pasquinelli; Joler, 2020).

O racismo e outras discriminações podem ser mantidos e incluídos nas operações da IA, devido a sua ausência de transparência e explicabilidade. O pesquisador Tarcízio Silva deixou isso bem evidente ao afirmar que

o princípio de “explicabilidade” pode ser visto como uma prerrogativa essencial de combate ao racismo algorítmico se encarado como pertinente não apenas às linhas de código, mas também aos processos de planejamento, implementação e definição sobre quem os sistemas beneficiam ou excluem (Silva, 2022, p. 206).

Além disso, nos segmentos financeiro e bancário, as discriminações de classe e renda são constitutivas do próprio sistema. Analistas de risco, junto com equipes multidisciplinares, desenvolvem e ajustam os dados, algoritmos e modelos para atender ao interesse do sistema financeiro. É muito comum a elevação das taxas de risco e até mesmo a negação de empréstimos para pessoas que residem em determinados endereços, ou seja, os algoritmos são ajustados para dar peso negativo a determinados códigos de endereçamento postal. Em seleções de currículos, também é muito comum a definição de quais escolas e endereços na cidade não devem ser classificados como aptos à realização de testes e entrevistas. As discriminações e exclusões que antes eram feitas em reuniões, agora são praticadas por algoritmos.

O DOMÍNIO DO FUTURO E A PRECARIZAÇÃO DAS ESCOLAS

Atualmente, a IA generativa, principalmente os chats, estão sendo utilizados nas escolas de modo preocupante. A princípio, esses sistemas prometem a personalização do ensino, mas promovem a homogeneização de realidades completamente heterogêneas. As bases de dados a partir das quais foram treinados foram obtidas, principalmente, nos países do Norte Global. Nesse sentido, portam traços culturais que não são os nossos. O processo de ensino deve considerar o contexto e a realidade de cada turma. Isso

exige maior autonomia da professora ou professor, não sua submissão aos ditames e condicionamentos de um sistema automatizado.

Os governos do Paraná e de São Paulo tentam solucionar problemas educacionais sem considerar o contexto social. Partem do pressuposto de que a gamificação poderá resolver questões que estão longe do problema do interesse, pois a falta de atenção dos estudantes pode advir de dificuldades que não estão na escola, sendo ligadas à pobreza e a outras dimensões do cotidiano. Acreditam que a educação é um problema da ordem da produtividade. Sem dúvida, sistemas algorítmicos podem elevar a produtividade de uma empresa, por exemplo, de um escritório de advocacia que passa a produzir peças jurídicas com soluções de IA generativa. Já a substituição de cinco estagiários por apenas um revisor para os textos produzidos por máquina é sinônimo de redução de custos.

Como um resumo de um livro de Machado de Assis feito pelo ChatGPT pode beneficiar a capacidade de leitura das crianças e adolescentes? De que forma testes curiosos podem melhorar a compreensão de uma equação, se não forem feitos com o acompanhamento de um educador, que perceberá a dificuldade da criança? Sem dúvida, professores e professoras podem preparar aulas e exemplos utilizando a IA generativa, a fim de testar maiores possibilidades de envolvimento de uma turma na escola, mas seu uso não é indispensável.

Vivemos, hoje, três grandes problemas na educação, diante dos sistemas automatizados, da IA realmente existente. O primeiro diz respeito ao fato de que esses sistemas podem ser utilizados para reduzir a autonomia do professor em sala de aula. Em segundo lugar, muitas vezes servem à precarização do trabalho do educador, pois geram sistemas de aula em vídeo que reduzem custos dos empreendimentos educacionais, aumentando seus lucros e reduzindo a qualidade do ensino. Já o terceiro problema é que, principalmente nos ensinos médio e superior, os chats estão sendo utilizados pelos estudantes para deixar de ler e interpretar livros, bem como escrever redações. Trata-se de um processo de redução de atividades cognitivas dos adolescentes, o que é grave. Aprender a ler não é uma atividade possível de se realizar com um chat da OpenAI.

O debate ético nas escolas é cada vez mais importante. A desmistificação da IA realmente existente e a conversa sobre as possibilidades de gerar acomodação, emburrecimento e despreparo para quem precisa aprender são, agora, fundamentais. Todavia, a doutrina neoliberal se entrelaça à crise fiscal dos Estados e a visões obscuras e reacionárias da sociedade. Assim, professores são desvalorizados, salários de educadores são reduzidos e sistemas automatizados são apresentados como solução de algo que a automatização nunca poderá fazer.

Por fim, cabe dizer que os sistemas de aprendizado profundo serviram para treinar modelos de reconhecimento facial no estado do Paraná para substituir a chamada de presença em sala de aula (Tecnologia [...], 2023). Para esse fim, dados biométricos – considerados sensíveis por quase todas as legislações do mundo, inclusive a LGPD brasileira – das crianças e adolescentes estão sendo coletados nas escolas. Tudo isso porque a equipe do governo paranaense não confia nas chamadas realizadas pelos professores. Um gasto inútil e altamente perigoso que não motivou as autoridades de proteção de dados a anular tamanha aberração.

A NEGAÇÃO DA SOBERANIA NOS DEBATES REGULATÓRIOS

A soberania é um conceito importante e controverso na teoria política. Tal noção sofreu profundas alterações ao longo do tempo e dos estudos na área, conforme mostra a historiografia de Quentin Skinner (2010). Por isso, dificilmente podemos ter uma visão essencialista da soberania. Em cada comunidade linguística, sua definição e uso envolveram um emaranhado de concepções sobre o que seria o poder, sua extensão, sua origem, sua justificativa moral, sua finalidade, entre outras acepções similares ao que ocorreu com a palavra Estado. O mais importante a destacar é que soberania é um termo em disputa.

Os embates mais contemporâneos expressam um deslocamento radical dessa noção. Se, até certo período, ela estava restrita à dimensão e extensão do poder de Estado, passou a ser aplicada como ideia de autonomia decisória da sociedade, de grupos sociais ou de indivíduos sobre aspectos da vida. É a partir dessas alterações

de sentido e emprego que a soberania se encontrou com o digital. A intensa digitalização da sociedade e sua dataficação (Mayer-Schönberger, Cukier, 2013), inaugurando a primazia da plataformização (Srnicek, 2017; Van Dijck; Poell; De Waal, 2018), todas controladas pelas gigantescas empresas chamadas de *big techs* (Morozov, 2018), bem como a extração massiva e constante de dados de usuários (Van Dijck; Poell; De Waal, 2018; Zuboff, 2019) abriram um terreno para o questionamento de como as tecnologias informacionais estariam alterando relações de poder, reforçando ou deslocando domínios econômicos e políticos (Bratton, 2016; Kwet, 2019). Assim, a caracterização de um novo colonialismo, tecnologicamente situado, dará sentido às oposições desse conceito, a partir das reivindicações de soberania digital e soberania de dados.

A soberania digital refere-se ao controle e governança que a sociedade e seu Estado possuem sobre suas infraestruturas digitais, enquanto a soberania de dados diz respeito ao controle sobre os dados gerados, armazenados e processados da população. A soberania é crucial para proteger a privacidade dos cidadãos, garantir a segurança dos dados, mas é decisiva para o desenvolvimento econômico de um país, uma vez que a economia é cada vez mais digital e dataficada. A autodeterminação de nosso país envolve o exercício e a ampliação de nossa capacidade de inventar, desenvolver e regular os sistemas de IA.

Uma das principais questões das soberanias digital e de dados passa pelo debate sobre a localização dessas informações e a construção de infraestruturas soberanas para elas e a IA. Lobistas das *big techs* atuam para manter a livre extração de dados nos diversos países em que atuam. No Brasil, por exemplo, as universidades continuam hospedando os dados de suas atividades, inclusive de pesquisa, nas infraestruturas dessas empresas e o MEC entrega os dados do desempenho escolar dos adolescentes brasileiros que se registram no Sisu para a Microsoft (Microsoft [...], 2020). Ainda, durante quase toda a gestão Bolsonaro, o governo brasileiro enviava dados dos servidores e pensionistas federais que utilizavam o chat do SouGov para treinar o modelo de linguagem chamado IBM-Watson, localizado nos Estados Unidos.

Nos debates regulatórios brasileiros, a questão da soberania é pouco discutida. Isso ocorre por diversas causas, dentre as quais as mais relevantes são: a supremacia da visão neoliberal, avessa à ideia de desenvolvimento nacional com maior independência; o lobby das *big techs*, que busca assegurar seus mercados de extração de dados e valor; bem como o desconhecimento ou negação do atual cenário internacional, em que os países estão buscando interiorizar elementos estratégicos da cadeia de produção de tecnologias fundamentais.

CONCLUSÃO

As questões aqui abordadas sobre a regulação da IA indicam que os dados, sistemas algorítmicos, modelos e infraestruturas de seu armazenamento, processamento e análise devem ser tratados, para garantir alguns fins fundamentais. São eles: ampliar direitos individuais e coletivos; salvaguardar a autodeterminação de nosso país, nossos povos e nossa cultura; defender nossa diversidade e criar condições favoráveis a um ecossistema digital em que não sejamos meros consumidores e usuários de produtos e serviços desenvolvidos com a extração de dados de nossa população. A regulação deve apostar na consolidação da soberania digital e de dados, a fim de que possamos liberar nossa inventividade tecnológica e ampliar nossa governança tecnopolítica sobre os sistemas automatizados e a inteligência artificial realmente existente.

REFERÊNCIAS

- BRATTON, Benjamin H. **The stack**: on software and sovereignty. Cambridge, MA: MIT Press, 2016.
- COULDRY, Nick; MEJIAS, Ulises A. **The costs of connection**: how data is colonizing human life and appropriating it for capitalism. Redwood City: Stanford University Press, 2020.
- KWET, Michael. Digital colonialism: US empire and the new imperialism in the Global South. **Race & Class**, [s. l.], v. 60, n. 4, p. 3-26, 2019. DOI: <https://dx.doi.org/10.2139/ssrn.3232297>. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3232297. Disponível em: 10 set. 2024.
- MASLEJ, Nestor *et al.* **The AI index 2024 annual report**. Stanford: AI Index Steering Committee; Institute for Human-Centered AI; Stanford University, 2024.
- MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. **Big data**: a revolution that will transform how we live, work, and think. Boston: Houghton Mifflin Harcourt, 2013.
- MICROSOFT destaca Sisu em nuvem como case de sucesso. **Portal do MEC**, Brasília, DF, 23 mar. 2020. Disponível em: <http://portal.mec.gov.br/ultimas-noticias/410-sisu-535874847/86661-microsoft-destaca-sisu-em-nuvem-como-case-de-sucesso#:~:text=A%20Microsoft%20destacou%20o%20Sistema,aumentar%20a%20capacidade%20de%20acessos>. Acesso em: 26 ago. 2024.
- MILLER, Katharine. Should AI models be explainable? That depends. **HAI**, Stanford, 16 mar. 2021. Disponível em: <https://hai.stanford.edu/news/should-ai-models-be-explainable-depends>. Acesso em: 8 set. 2024.
- MIT TECHNOLOGY REVIEW INSIGHTS. **The great acceleration**: CIO perspectives on generative AI. São Francisco: Databricks, 2023.
- MOROZOV, Evgeny. **Big tech**: a ascensão dos dados e a morte da política. São Paulo: Ubu, 2018.
- PASQUALE, Frank. **The black box society**: the secret algorithms that control money and information. Cambridge, MA: Harvard University Press, 2015.
- PASQUINELLI, Matteo; JOLER, Vladan. O manifesto nooscópio: inteligência artificial como instrumento de extrativismo do conhecimento. Tradução: Leandro Módolo e Thais Pimentel. **Lavits**, [s. l.], 30 jul. 2020. Disponível em: <https://lavits.org/o-manifesto-nooscopio-inteligencia-artificial-como-instrumento-de-extrativismo-do-conhecimento/>. Acesso em: 8 set. 2024.
- PIGOU, Arthur. **The economics of welfare**. Abingdon: Routledge, 2017.
- SILVA, Tarcízio. **Racismo algorítmico**: inteligência artificial e discriminação nas redes digitais. São Paulo: Edições Sesc, 2022.
- SKINNER, Quentin. The sovereign state. In: KALMO, Hent; SKINNER, Quentin (ed.). **A genealogy**. Sovereignty in fragments: The past, present and future of a contested concept. Cambridge, UK: Cambridge University Press, 2010.
- SRNICEK, Nick. **Platform capitalism**. Hoboken: John Wiley & Sons, 2017.
- STERLING, Bruce. Web semantics: Microsoft Project Turing introduces Turing Natural Language Generation (T-NLG). **Wired**, San Francisco, 13 feb. 2020.
- TECNOLOGIA de reconhecimento facial na chamada chega a 1,6 mil colégios da rede estadual. **Agência Estadual de Notícias**, Curitiba, 17 maio 2023. Disponível em: <https://www.aen.pr.gov.br/Noticia/Tecnologia-de-reconhecimento-facial-na-chamada-chega-16-mil-colegios-da-rede-estadual>. Acesso em: 26 ago. 2024.
- VAN DIJCK, José; POELL, Thomas; DE WAAL, Martijn. **The platform society**: public values in a connective world. Oxford: Oxford University Press, 2018.
- WINNER, Langdon. Do artifacts have politics? In: WINNER, Langdon. **The whale and the reactor**: a search for limits in an age of high technology. Chicago: The University of Chicago Press. 1986.
- ZUBOFF, Shoshana. **The age of surveillance capitalism**: the fight for a human future at the new frontier of power. London: Profile Books, 2019.