

Desafios da IA responsável na pesquisa científica

Virgílio Almeida

Elen Nas

resumo

A inteligência artificial está cada vez mais presente em praticamente todas as áreas da ciência e tecnologia. O uso de tecnologias de IA tem o potencial de acelerar a produtividade dos pesquisadores e reduzir custos da pesquisa científica. Apesar dos benefícios, preocupações surgem em relação à opacidade das ferramentas de IA, que podem levar a consequências negativas, como discriminação, falsidades e desinformação científica, comprometendo a confiança pública na ciência. O artigo explora a implementação dos princípios de IA responsável nas organizações de pesquisa visando minimizar possíveis impactos negativos da IA na ciência.

Palavras-chave: IA responsável; transparência; inteligência artificial; IA ética; centralidade do humano.

abstract

Artificial intelligence is gaining prominence across every sector of science and technology. Leveraging AI technologies holds promise for boosting researchers' productivity and cutting down on scientific research costs. However, concerns have emerged regarding the lack of transparency in AI tools, which may result in adverse outcomes like discrimination, scientific misinformation, and eroded public trust in science. This article delves into how research organizations are adopting responsible AI principles to mitigate potential negative effects on science.

Keywords: *responsible AI; artificial intelligence; transparency; ethical AI; centrality of the human.*

O

destino das tecnologias de inteligência artificial não está previamente determinado; há um elevado grau de incerteza e imprevisibilidade em relação ao seu uso e desenvolvimento. O futuro está em aberto e depende em boa medida do que o país for capaz de formular e construir. É essencial compreender o potencial de utilização e progresso da IA no Brasil, especialmente considerando a desigualdade econômica e o baixo nível de escolaridade, que representam desafios para a implementação das novas tecnologias no país. Ao analisar o papel da IA no contexto brasileiro, é preciso entender como essas tecnologias afetam a produtividade, o mercado de trabalho

e a inovação. Mas é também fundamental avaliar os riscos e as consequências da IA em países em desenvolvimento, como o Brasil.

Inteligência artificial, com seus sistemas algorítmicos e robôs, tem sido usada de muitas maneiras socialmente benéficas, como, por exemplo, antecipar necessidades de saúde, auxiliar gerência de tráfego em centros urbanos, otimizar o uso de recursos escassos, facilitar decisões financeiras e aumentar a eficiência de serviços governamentais. Algoritmos têm sido também usados para tomar decisões críticas

VIRGÍLIO ALMEIDA é professor emérito do Departamento de Ciência da Computação da UFMG e professor associado ao Berkman Klein Center da Universidade de Harvard.

ELEN NAS é pesquisadora da Cátedra Oscar Sala do Instituto de Estudos Avançados (IEA) da USP.

para indivíduos, com impacto direto na vida das pessoas. Baseados em grandes massas de dados, algoritmos operam as plataformas de mídia social, controlando o fluxo de informações para diferentes setores da sociedade, com impactos na formação social, cultural e política dos indivíduos e no comportamento coletivo da sociedade. Spotify, Twitter, YouTube, Waze, Netflix, Facebook, TikTok, Amazon são algumas das plataformas algorítmicas que impactam bilhões de vidas em todo o mundo, influenciando hábitos, culturas e valores compartilhados por diferentes grupos sociais. Há, no entanto, uma preocupação crescente em relação ao papel dos algoritmos na sociedade, em particular com a tomada de decisão algorítmica sobre indivíduos e grupos sociais. Por exemplo, algoritmos projetados para filtrar, selecionar e exibir a grande quantidade de informações disponíveis online, combinados com a tendência das pessoas de buscar ambientes sociais similares, podem induzir a vieses em grupos da sociedade, alimentando racismo, preconceitos, discriminação e desinformação. Inteligência artificial, robôs e demais máquinas operadas por algoritmos mediam cada vez mais nossas interações profissionais, sociais, culturais, econômicas e políticas. Compreender o comportamento dos algoritmos e sistemas de inteligência artificial é essencial para nossa capacidade de governar suas ações, colher benefícios e minimizar seus impactos negativos (Mendonça; Filgueiras; Almeida, 2023).

As tecnologias de inteligência artificial já afetam vários aspectos da ciência e tecnologia. São tecnologias que podem

analisar grandes volumes de dados com o objetivo de classificar, categorizar, identificar padrões, prever e tomar decisões baseadas em evidências. O desenvolvimento e aplicação da IA ocorre em diferentes aspectos das ciências, incluindo biologia, física, matemática e ciências sociais. O uso das tecnologias de IA na ciência, como aprendizado de máquina, aprendizado profundo, processamento de linguagem natural e tecnologias de IA generativa tem o potencial de acelerar a produtividade de pesquisadores, reduzir custos de experimentos e levar qualidade e segurança à pesquisa científica. Por exemplo, formular e testar hipóteses são etapas fundamentais do método científico, que demandam tempo e conhecimento nas pesquisas científicas. Os métodos de IA podem contribuir em várias fases do processo de geração de hipóteses, a partir da identificação de padrões e expressões simbólicas em grandes massas de dados experimentais. Essas tecnologias de IA são capazes de acelerar a identificação de contraexemplos que contradizem uma conjectura teórica (Wang, 2023). Há a expectativa de que a IA possa levar a um aumento de produtividade necessária para a ciência avançar na compreensão dos grandes desafios globais da humanidade, como mudanças climáticas, poluição, escassez de recursos e doenças infecciosas.

Os avanços da inteligência artificial na pesquisa são evidenciados pelo aumento de publicações científicas que mencionam IA nos títulos e resumos. Um levantamento da revista *Nature* mostra que a proporção de artigos na base de dados Scopus mencionando termos referentes

à IA nos títulos ou resumos agora é de 8%, um aumento de 2% em relação a uma década atrás. A pesquisa da *Nature* coletou opiniões sobre o papel da IA ao entrevistar mais de 1.600 pesquisadores de todo o mundo. Os pesquisadores identificaram várias formas positivas de uso das ferramentas de IA na pesquisa científica. Dois terços dos entrevistados observaram que a IA acelera o processamento de grandes massas de dados, 58% informaram que a IA permite a execução de cálculos que anteriormente não eram viáveis e 55% mencionaram economia de tempo e recursos nos projetos de pesquisa (Van Noorden, 2023).

Ao mesmo tempo que a maioria dos pesquisadores mostrou aspectos positivos no uso de IA na pesquisa científica, foram também identificadas preocupações com essas tecnologias, que incluem principalmente a opacidade das ferramentas de IA, consideradas por muitos “caixas-pretas”, das quais não é possível entender as razões subjacentes aos resultados gerados por IA. As preocupações também abrangem a existência de vieses, bem como as possibilidades de estudos com resultados falsos e propagação de desinformação científica. A consequência desses aspectos potencialmente negativos é o comprometimento da confiança pública na ciência e nos resultados científicos.

Este artigo explora as possibilidades de se adequar os fundamentos e práticas da inteligência artificial responsável (IAR) ao contexto das organizações de pesquisa, principalmente dentro de universidades, pois no Brasil mais de 95% da produção científica nas bases internacionais deve-se à capacidade de pes-

quisa de universidades públicas, federais e estaduais (Moura, 2019).

ENTENDENDO A IA RESPONSÁVEL

A IA responsável visa ao desenvolvimento de sistemas algorítmicos que beneficiem indivíduos, sociedades e o meio ambiente. Abrange os aspectos éticos, legais e técnicos do desenvolvimento e implementação de tecnologias de IA benéficas. Isso inclui garantir que o sistema de IA não cause danos, não interfira na agência humana, não discrimine e não desperdice recursos. Assim, a IA responsável é uma área de pesquisa multidisciplinar, em que cientistas da computação, éticos, juristas, sociólogos, psicólogos e muitos outros especialistas trabalham juntos para moldar o futuro da tecnologia, visando ao bem-estar humano como principal objetivo (Baeza-Yates, 2022).

No ambiente de pesquisa científica, há necessidade de formular estratégias visando apoiar as equipes de pesquisa científica para o desenvolvimento e implementação de IA responsável, com fundamentos éticos, responsabilidades morais e sociais, que estabeleçam limites de proteção para sociedade e meio ambiente. As organizações de pesquisa científica que estão envolvidas no desenvolvimento, implementação e uso de ferramentas de IA têm observado potenciais conflitos entre avanço tecnológico e preocupações éticas e sociais com o impacto dessas novas tecnologias. A tendência em direção à IA responsável nessas instituições destaca um crescente foco na incorporação de considerações éticas no processo de *design* e

desenvolvimento de ferramentas de IA, promovendo valores fundamentais como *responsibilities*, transparência, justiça e inclusão social.

Os fundamentos básicos da IA responsável buscam garantir que os sistemas de IA sejam seguros, aceitos e confiáveis. Para isso, o sistema deve ser projetado levando em consideração princípios éticos e as consequências morais de suas ações e decisões, de maneira responsável e transparente (Dignum, 2017). Diversas organizações internacionais têm elaborado conjuntos de princípios fundamentais para a adoção responsável da inteligência artificial. Estes princípios visam mitigar os riscos específicos associados ao uso dessas tecnologias, tais como equidade, viés, autonomia, privacidade, proporcionalidade e responsabilidade. A minimização dos riscos da inteligência artificial requer o estabelecimento de normas e princípios relacionados à transparência, interpretabilidade, explicabilidade e auditabilidade, entre outros aspectos. Tais princípios servem como diretrizes para orientar o desenvolvimento de aplicações de inteligência artificial em diversos domínios, como ciência, saúde, educação, finanças e mídias sociais, apenas para mencionar algumas áreas impactadas por essa tecnologia. É fundamental que os princípios de inteligência artificial responsável sejam complementados por uma governança adequada dos dados e modelos de IA, além do treinamento das pessoas capacitadas para sua implementação.

O Conselho de Políticas Tecnológicas da ACM (Association for Computing Machinery) elaborou uma Declaração de Princípios para Sistemas Algorítmicos

Responsáveis (Council, 2022), dado que os algoritmos de IA estão sendo cada vez mais utilizados por empresas, governos e cidadãos, com impactos significativos sobre indivíduos e sociedade. O documento da ACM apresenta um conjunto de princípios fundamentais para a criação de sistemas de IA responsáveis. Esse conjunto abrange os seguintes princípios: 1) legitimidade e competência; 2) minimização de danos; 3) segurança e privacidade; 4) transparência; 5) interpretabilidade e explicabilidade; 6) manutenibilidade; 7) contestabilidade e auditabilidade; 8) responsabilidade (*accountability*); e 9) limitação de impactos ambientais (Council, 2022).

Implementar diretrizes de IA responsável e ética requer transformar princípios abstratos em medidas concretas, procedimentos e práticas que possam ser incorporados aos sistemas, ferramentas e fluxos de trabalho de uma organização. Uma implementação eficiente dos princípios conduz a uma maior transparência e confiança entre os usuários e as partes interessadas, diminuindo a distância entre os conceitos teóricos e sua aplicação no mundo real.

Os países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) adotaram a Recomendação do Conselho sobre Inteligência Artificial. Os princípios da OCDE sobre Inteligência Artificial promovem uma IA inovadora, confiável e que respeite os direitos humanos e os valores democráticos. Esses princípios se aplicam aos atores de IA e a todo o governo, incluindo a implementação de IA (Yeung, 2020).

A OCDE estabeleceu os seguintes princípios para a administração responsável de IA confiável:

- crescimento inclusivo, desenvolvimento sustentável e bem-estar, visando engajar os agentes proativamente na busca responsável de uma IA confiável em busca de resultados benéficos para as pessoas e o planeta;
- valores centrados no ser humano e justiça, que devem respeitar os direitos humanos e os valores democráticos ao longo do ciclo de vida dos sistemas de IA;
- transparência e explicabilidade, que buscam oferecer uma compreensão geral dos sistemas de IA, para permitir que aqueles afetados por um sistema de IA entendam os resultados que lhes afetam;
- robustez e segurança, que devem garantir sistemas de IA seguros em relação aos dados e decisões tomadas por esses sistemas;
- responsabilidade, visando tornar os agentes de IA responsáveis pelo funcionamento adequado dos sistemas de IA.

Esses princípios da OCDE ressaltam pontos que são comuns a vários outros documentos sobre IA responsável, que são os pilares fundamentais: transparência, responsabilização, explicabilidade e justiça. Além dos princípios, há também ênfase na importância de monitoramento permanente das aplicações de IA e na necessidade de investimentos públicos e privados em pesquisa e desenvolvimento para estimular a inovação em IA responsável.

Cada vez mais, IA está sendo implantada em ambientes de ensino superior para

aprimorar a eficiência dos sistemas, refinar tomadas de decisão e melhorar a prestação de serviços. As universidades estão começando a adotar ferramentas habilitadas por IA para tarefas que variam desde administração até ensino e pesquisa. No entanto, o uso de IA pode acarretar riscos significativos. A Universidade da Califórnia (UC), como uma grande universidade pública, desenvolveu uma estratégia para orientar a implantação das tecnologias de IA, de maneira a proteger a comunidade e refletir os valores da instituição, que incluem não discriminação, segurança e privacidade. Assim, a UC adotou um conjunto de recomendações para orientar a implantação segura e responsável da inteligência artificial e estabeleceu os seguintes princípios para aquisição, desenvolvimento, implementação e monitoramento de ferramentas de IA dentro dos diversos *campi* da universidade (University of California, 2021):

- avaliação cuidadosa dos benefícios e riscos da IA, considerando que as necessidades das pessoas afetadas são essenciais para decidir sobre sua aplicação;
- transparência é fundamental; as pessoas devem ser informadas sobre o uso de ferramentas de IA e capazes de entender os resultados e contestá-los, se necessário;
- as ferramentas de IA devem ser eficazes, precisas e seguras para seu propósito, ao longo de sua vida útil;
- é crucial avaliar e mitigar preconceitos e discriminação nas ferramentas de IA, adotando procedimentos para remediar danos;
- privacidade e segurança dos dados devem ser priorizadas no *design* das ferramentas habilitadas por IA;

- desenvolvimento e uso de IA devem respeitar valores humanos e direitos civis e humanos, observando as leis e princípios aplicáveis;
- as ferramentas de IA devem promover benefícios equitativos para todos, incluindo aspectos sociais, econômicos e ambientais;
- a Universidade da Califórnia deve ser responsável pelo desenvolvimento e uso ético de sistemas de IA em seus serviços, de acordo com os princípios acima.

GOVERNANÇA DE IA RESPONSÁVEL

A existência de conjuntos de princípios destinados a orientar pesquisadores e professores no desenvolvimento e na implementação de tecnologias de IA é um ponto de partida útil, porém insuficiente, pois não estão associados a explicações sobre como os princípios poderiam ser implementados na prática. A questão fundamental a ser explorada é: que tipos de estruturas de governança precisam estar em vigor para garantir o uso responsável de IA na pesquisa científica?

A IA responsável ajudará a minimizar consequências indesejáveis e não intencionais, como danos à sociedade, ao meio ambiente e à confiança que a sociedade tem nos resultados científicos. À medida que o uso de IA permeia uma parcela maior de laboratórios e organizações de pesquisa, é necessário gerenciar o risco do uso dessas tecnologias emergentes. Os princípios são importantes, mas só proporcionarão o que as organizações de pesquisa precisam se forem combinados com práticas de governança que ajudem

a guiar o uso de ferramentas de IA em projetos de pesquisa científica, desde o *design* até a produção e publicação de resultados de pesquisa.

Um estudo publicado pelas Academias Nacionais dos Estados Unidos, que inclui as academias de Medicina, Engenharia e Ciências, explora questões éticas e de governança relacionadas às consequências individuais e sociais da pesquisa em computação e suas aplicações, que incluem inteligência artificial (National Academies of Sciences, 2022). Quando os resultados da pesquisa em computação, como IA, são incorporados em sistemas com potencial impacto social, várias dimensões da responsabilidade social precisam ser abordadas por meio de mecanismos de governança e por órgãos reguladores.

Para evoluir em direção ao uso responsável de IA, as instituições, sejam elas laboratórios ou universidades, devem fornecer orientações sobre estratégias para identificar e avaliar os riscos de ferramentas habilitadas por IA e estabelecer mecanismos de governança apropriados para a instituição se adequar às políticas e normas para uso de IA (Bano, 2023).

TREINAMENTO

É essencial avaliar a compreensão dos conceitos de ética e responsabilidade em IA dentro de uma organização de pesquisa, para orientar estratégias personalizadas, modelos de governança e programas de treinamento adaptados às necessidades específicas do grupo de pesquisadores. Um estudo conduzido junto a pesquisadores do CSIRO (Commonwealth Scientific

and Industrial Research Organization), na Austrália, oferece informações importantes sobre a percepção de pesquisadores e professores referentes a ética e responsabilidades em projetos de pesquisa que fazem uso das tecnologias de IA. Os projetos avaliados, usando tecnologias de IA, situam-se em vários domínios da ciência.

Os resultados apontam para a necessidade de maior conscientização e apoio para abordar questões éticas nessas áreas científicas. As respostas indicam que 70% dos entrevistados relataram não ter usado qualquer documento de ética em IA, enquanto 17% expressaram desconhecimento da existência de documentação para análise de aspectos éticos das ferramentas de IA. Em uma outra análise, foi observado que 83% dos entrevistados admitiram não ter utilizado recursos adicionais para garantir a concepção ou desenvolvimento ético da inteligência artificial (Bano, 2023). Em ambientes de pesquisa, como universidades e laboratórios, são essenciais iniciativas de treinamento e conscientização sobre os impactos éticos do uso das tecnologias de IA em pesquisas científicas.

NÍVEIS DE GOVERNANÇA DE IA

Há uma crescente demanda para implementar regulamentações de IA visando minimizar os riscos para a segurança e preservar os direitos humanos, de forma a garantir que seus benefícios contribuam para a solução de desafios globais, como desenvolvimento econômico, desigualdade e mudança climática. Em resumo, muitos países estão discutindo modelos de

governança de IA que criam estruturas, regras, normas e leis para governar dispositivos, serviços e aplicações de IA. No mundo digital, a governança deriva das ações de diferentes atores que definem políticas e constroem um complexo conjunto de normas formais e informais que organizam as diferentes camadas e processos envolvidos na construção de infraestruturas digitais, plataformas, serviços e aplicações usadas pela sociedade.

Há também uma necessidade premente de desenvolver estruturas de governança para pesquisas científicas que fazem uso de IA, de forma a minimizar consequências que decorrem das atividades de pesquisa que fazem uso das ferramentas de IA (Batool, 2023). A governança da IA busca reconhecer a intrincada interconexão entre ciência, tecnologia e sociedade para compreender e supervisionar as extensas repercussões dos algoritmos de IA sobre indivíduos, comunidades e instituições.

A governança é um conceito com muitos significados, que pode ser implementada através de diferentes arranjos institucionais (Filgueiras, 2022). Um arranjo abrangente para diferentes formas de regulação pode ser estruturado em uma hierarquia de níveis de governança. No nível mais alto da estrutura de governança de IA estão legislações nacionais, que estabelecem o contexto para as regulações de IA em instituições. Por exemplo, o “AI Act” aprovado pela União Europeia define regras para promover uma IA confiável na Europa e garantir que os sistemas de IA respeitem os direitos fundamentais, a segurança e os princípios éticos (European Parliament, 2023). A lei

atribui aplicações de IA a três categorias de risco. Primeiro, as aplicações e sistemas que criam um risco inaceitável. Aplicações de IA para calcular pontuação social a ser usada por governos são proibidas. Segundo, as aplicações de alto risco, como uma ferramenta de análise de currículos que classifica os candidatos a emprego, estão sujeitas a requisitos legais específicos. Por fim, as aplicações não explicitamente proibidas ou listadas como de alto risco são em grande parte deixadas sem regulamentação.

No nível intermediário de regulação de IA estão princípios e boas práticas para o desenvolvimento de sistemas responsáveis de IA. Um bom exemplo de iniciativas de governança nesse nível intermediário é a Ordem Executiva sobre o Desenvolvimento Seguro, Responsável e Confiável da Inteligência Artificial do governo americano, publicada em outubro de 2023 (Biden, 2023), que estabelece princípios e prioridades para orientar o governo sobre o desenvolvimento e usos da IA. A Ordem Executiva atribui a mais alta urgência à governança do desenvolvimento e uso da IA de forma segura e responsável, e promove uma abordagem coordenando diferentes setores do governo para alcançar o objetivo de uma IA global.

Em uma abordagem *bottom-up* para governança, o nível inicial da hierarquia de governança de IA começa com regras, normas e políticas para uso de ferramentas de IA dentro de cada instituição. É importante ressaltar que a governança de IA em uma instituição deve estar alinhada com os princípios, práticas e legislações especificadas nos níveis superiores, ou seja, legislação nacional determinada por

Poder Executivo ou Legislativo. É justamente o nível inicial de regulação que laboratórios, universidades e institutos de pesquisa devem adotar para implantação das tecnologias de IA. Nesse primeiro nível, as organizações de pesquisa têm oportunidade de estabelecer regras e normas, quando ainda não se tem no país legislações aprovadas por Congresso e governo, que usualmente necessitam de mais tempo para formar arranjos políticos requeridos para aprovação de legislações. Esse deve ser o ponto de partida das instituições de pesquisa e ensino superior no Brasil para avançar responsavelmente no uso e aplicação das tecnologias de IA.

CONCLUSÃO

Existem muitas definições de governança digital, sem consenso na literatura acadêmica. Conforme proposto por Filgueiras e Almeida (2020), a governança digital é a capacidade de instituições polí-cêntricas no mundo digital de governar (de maneira legítima, inclusiva e segura) o uso dos bens digitais comuns para produzir serviços sustentáveis e políticas públicas implementadas por governos e empresas de maneira não territorial e orientada por resultados. IA responsável é um indicativo de medidas preventivas que buscam compor uma estrutura de governança com múltiplas instituições, princípios, políticas, ferramentas e processos, de modo a garantir que os sistemas de IA sejam desenvolvidos em conformidade com as leis e princípios éticos fundamentais. A governança da IA responsável deve fortalecer na prática valores essenciais como

transparência, justiça, explicabilidade, privacidade e segurança.

A inteligência artificial responsável ajudará a minimizar consequências indesejáveis e não intencionais na pesquisa científica, como, por exemplo, danos à imagem e credibilidade da ciência. À medida que a IA permeia várias áreas das ciências, gerenciar os riscos se torna cada vez mais importante. Princípios para uma IA responsável são um ponto de partida, mas só fornecerão o que as instituições precisam se forem combinados com práticas de governança que ajudem efetivamente a guiar o desenvolvimento e uso de aplicações de IA.

O avanço da IA nas instituições de ensino superior no Brasil deverá ser acompanhado de um processo de governança, incluindo documentação e ferramentas de apoio aos pesquisadores e professores, de modo a ser eticamente responsável. Operacionalizar diretrizes de IA responsável significa transformar princípios abstratos em ações concretas, processos e práticas que podem ser incorporados aos sistemas nos laboratórios e universidades. Isso aumenta a transparência e a confiança entre os usuários e partes interessadas, diminuindo a distância entre os conceitos teóricos e sua aplicação prática na IA.

REFERÊNCIAS

- BAEZA-YATES, R.; VILLOSLADA, P. *Human vs. artificial intelligence*. IEEE 4th International Conference on Cognitive Machine Intelligence (CogMI). 14/dez./2022, pp. 40-8.
- BANO, M. et al. "Investigating responsible AI for scientific research: an empirical study". arXiv preprint arXiv:2312.09561, 2023.
- BATOOL, A.; ZOWGHI, D.; BANO, M. "Responsible AI governance: a systematic literature review". arXiv preprint arXiv:2401.10896, 2023.
- BIDEN, J. R. *Executive order on the safe, secure, and trustworthy development and use of artificial intelligence*. 2023.
- COUNCIL, US Public Policy. "Statement on principles for responsible algorithmic systems". *Communications of the ACM*, 2022.
- DIGNUM, V. "Responsible autonomy". arXiv preprint arXiv:1706.02513, 2017, pp. 4.698-704.
- EITEL-PORTER, R. "Beyond the promise: implementing ethical AI". *AI and Ethics*, v. 1, 2021, pp. 73-80.

- EUROPEAN PARLIAMENT. *Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI*. 2023. Disponível em: <https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>. Acesso em: 3/mar./2024.
- FILGUEIRAS, F.; ALMEIDA, V. "Governance for the digital world: neither more state nor more market". *Springer Nature*, 2020.
- KITANO, H. "Nobel Turing Challenge: creating the engine for scientific discovery". *npj Systems Biology and Applications*, v. 7, n. 1, 2021, p. 29.
- MENDONÇA, R. F.; ALMEIDA, V.; FILGUEIRAS, F. *Algorithmic institutionalism: the changing rules of social and political life*. Oxford, Oxford University Press, 2023.
- MOURA, M. "Universidades públicas respondem por mais de 95% da produção científica do Brasil". *Academia Brasileira de Ciências*. Online. 2019. Disponível em: <https://www.abc.org.br/2019/04/15/universidades-publicas-respodem-por-mais-de-95-da-producao-cientifica-do-brasil/>. Acesso em: 3/mar./2024.
- NATIONAL ACADEMIES OF SCIENCES, ENGINEERING, AND MEDICINE et al. *Fostering responsible computing research: Foundations and practices*. 2022.
- RAKOVA, B. et al. "Where responsible AI meets reality: practitioner perspectives on enablers for shifting organizational practices". *Proceedings of the ACM on Human-Computer Interaction*, v. 5, n. CSCW1, 2021, pp. 1-23.
- UNIVERSITY OF CALIFORNIA. *Recommendations to guide the University of California's artificial intelligence strategy*. Presidential Working Group on AI, Final Report, out./2021.
- VAN NOORDEN, R.; PERKEL, J. M. "AI and science: what 1,600 researchers think". *Nature*, v. 621, n. 7980, 2023, pp. 672-5.
- WANG, H. et al. "Scientific discovery in the age of artificial intelligence". *Nature*, v. 620, n. 7972, 2023, pp. 47-60.
- YEUNG, K. "Recommendation of the Council on Artificial Intelligence (OECD)". *International Legal Materials*, v. 59, n. 1, 2020, pp. 27-34.