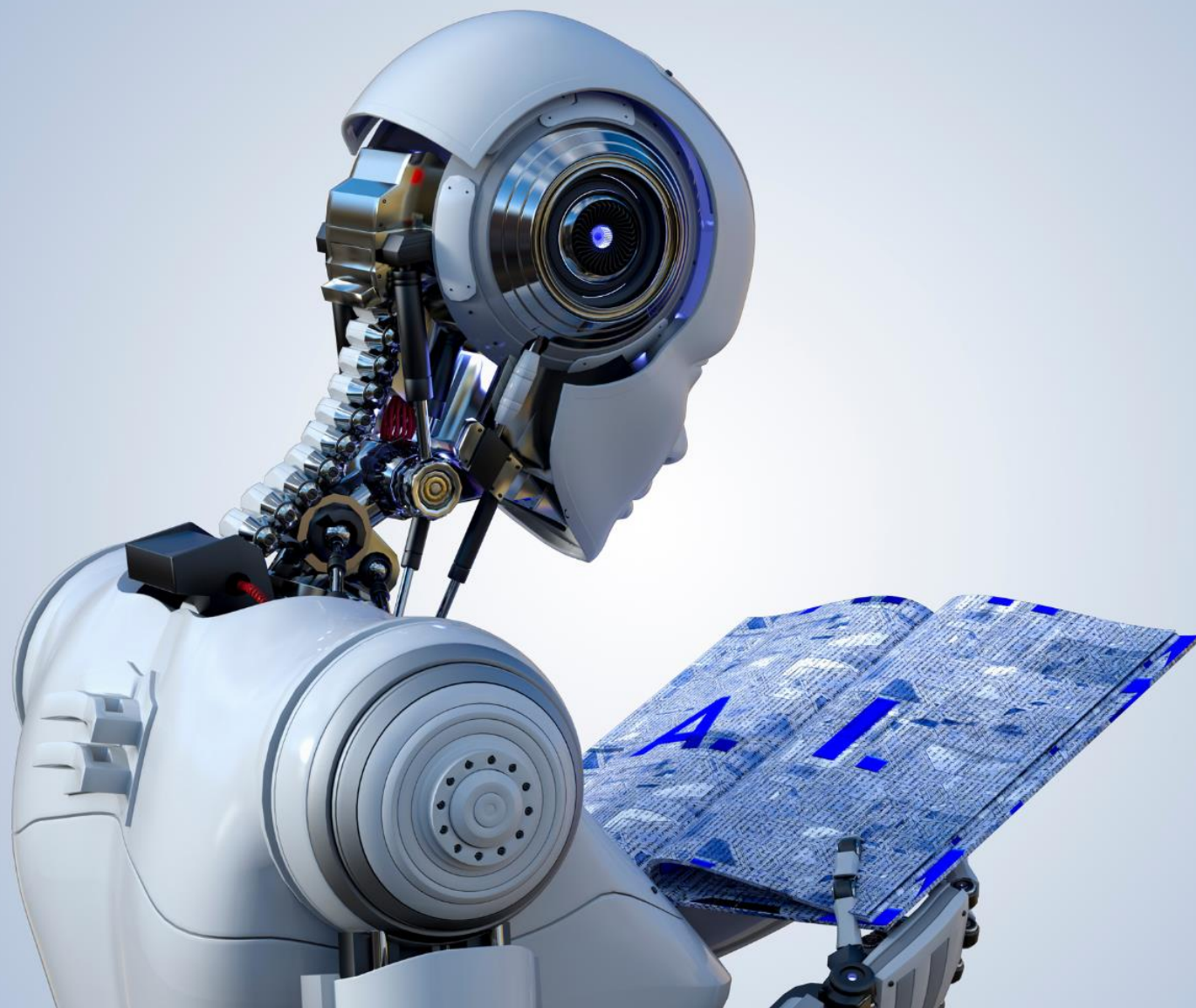




BIBLIOGRAFIA SOBRE INTELIGÊNCIA ARTIFICIAL E DISCRIMINAÇÃO DE GÊNERO E RAÇA

Biblioteca do
Senado Federal



SENADO
FEDERAL



BRASÍLIA – 2025

BIBLIOGRAFIA SOBRE INTELIGÊNCIA ARTIFICIAL E DISCRIMINAÇÃO DE GÊNERO E RAÇA

BRASÍLIA - 2025

BIBLIOTECA ACADÊMICO LUIZ VIANA FILHO

SENADO FEDERAL

BIBLIOGRAFIA SOBRE INTELIGÊNCIA ARTIFICIAL E DISCRIMINAÇÃO DE GÊNERO E RAÇA

Coordenação: Biblioteca Acadêmico Luiz Viana Filho - COBIB/SGIDOC

Foto de capa: Canva

Bibliografia sobre inteligência artificial e discriminação de gênero e raça. -- Brasília : Senado Federal, Biblioteca Acadêmico Luiz Viana Filho, 2025.
9 p.

1. Inteligência artificial, aspectos morais e éticos, bibliografia. 2. Discriminação racial, tecnologia, bibliografia. I. Brasil. Congresso Nacional. Senado Federal. Secretaria de Gestão de Informação e Documentação. Coordenação de Biblioteca.

CDD 006.3

Ficha catalográfica preparada por Alessandra Marinho da Silva CRB-1 2102

Senado Federal

Praça do Três Poderes s/nº

Brasília DF

CEP 70165-900

Introdução

Esta bibliografia, compilada pela Biblioteca do Senado, foi elaborada para dar suporte informacional aos estudos, às discussões e proposições normativas sobre inteligência artificial (IA), tendo como foco os desafios éticos relacionados à discriminação de gênero e raça. Atendendo à demanda do Comitê pela Promoção da Igualdade de Gênero e Raça do Senado Federal, o objetivo é oferecer referências para entender como os algoritmos podem reproduzir ou amplificar preconceitos sociais.

O que é viés algorítmico?

Viés algorítmico ocorre quando sistemas de IA apresentam decisões ou resultados que favorecem ou prejudicam determinados grupos de forma injusta. Isso pode acontecer por várias razões, como dados enviesados usados no treinamento do sistema ou pela forma como os algoritmos foram projetados. Esses vieses podem resultar em discriminação em áreas como contratações, crédito e até mesmo no sistema de justiça.

Por que isso é importante?

A discriminação algorítmica viola o princípio da igualdade ao oferecer tratamento desfavorável a pessoas ou grupos sem justificativa razoável. Em um mundo cada vez mais orientado por tecnologia, é fundamental que as políticas públicas levem em conta tanto os avanços tecnológicos quanto os princípios éticos, garantindo que ninguém seja excluído ou prejudicado.

Aspectos éticos em IA

A ética desempenha um papel central no desenvolvimento e na aplicação da IA. Além do viés algorítmico, questões como responsabilidade, privacidade e transparência precisam ser consideradas. A presença crescente da IA em sistemas do dia a dia traz implicações profundas para o futuro do trabalho, as relações sociais e até mesmo a democracia. Embora a IA possa oferecer soluções inovadoras, é um equívoco acreditar que ela resolverá crises globais, como desigualdades sociais ou mudanças climáticas, sem intervenção humana e políticas públicas bem estruturadas.

Desafios éticos principais incluem:

- **Responsabilidade:** Quem é responsável por decisões tomadas por IA?
- **Privacidade:** Como proteger dados pessoais em sistemas que dependem de grandes volumes de informação?

- **Viés e discriminação:** Como garantir que algoritmos não perpetuem ou amplifiquem preconceitos?
- **Regulamentação:** Quais políticas públicas são necessárias para mitigar riscos e promover justiça?

Foco da bibliografia

A seleção dos textos seguiu três principais temas:

1. **Estudos sobre viés algorítmico e seus impactos** – como algoritmos perpetuam desigualdades.
2. **Políticas e regulamentações para mitigar discriminação em IA** – explorando soluções e boas práticas.
3. **Perspectivas críticas e éticas sobre IA** – reflexões sobre como garantir justiça, responsabilidade e inclusão em decisões automatizadas.

Esta bibliografia visa fomentar discussões e promover ações que assegurem o uso ético, responsável e igualitário da inteligência artificial, contribuindo para um futuro mais justo e inclusivo.

Referências

ALBAROUDI, Elham; MANSOURI, Taha; ALAMEER, Ali. A comprehensive review of AI techniques for addressing algorithmic bias in job hiring. **AI**, Basel, v. 5, n. 1, p. 383–404, Mar. 2024. Disponível em: <https://www.mdpi.com/2673-2688/5/1/19>. Acesso em: 6 jan. 2025.

ALLEN, Robin; MASTERS, Dee (org.). **Regulating for an equal AI**: a new role for equality bodies. Brussels: European Network of Equality Bodies, 2020. Disponível em: https://equineteurope.org/wp-content/uploads/2020/06/ai_report_digital.pdf. Acesso em: 6 jan. 2025.

ANTONOPOULOU, Caterina. Algorithmic bias in anthropomorphic artificial intelligence: critical perspectives through the practice of women media artists and designers. **Technoetic Arts**: a journal of speculative research, v. 21, n. 2, p. 157–174, 24 Jan. 2024. Disponível em: https://doi.org/10.1386/tear_00109_1. Acesso em: 6 jan. 2025.

BAER, Tobias. **Understand, manage, and prevent algorithmic bias**: a guide for business users and data scientists. Berkele: Apress, 2019.

BAEZA-YATES, Ricardo; MURGAI, Leena; DUDY, Shiran. Bias in retrieval systems. In: ALONSO, Omar; BAEZA-YATES, Ricardo (ed.) **Information retrieval**: advanced topics and techniques. New York: Association for Computing Machinery, 2024. p. 389–444. Disponível em: <https://doi.org/10.1145/3674127.3674138>. Acesso em: 8 jan. 2025.

BAROCAS, Solon; SELBST, Andrew D. Big data's disparate impact. **California Law Review**, Berkeley, v. 104, n. 3, p. 671–732, June 2016. Disponível em: <https://heinonline.org/HOL/P?h=hein.journals/calr104&i=695>. Acesso em: 8 jan. 2025.

BLACKMAN, Reid; SURGEK, Rafael. **Máquinas éticas**: seu guia conciso para uma IA totalmente imparcial, transparente e respeitosa. Rio de Janeiro: Alta Books, 2024.

BORGESIUUS, Frederik Zuiderveen. **Discrimination, artificial intelligence and algorithmic decision-making**. Strasbourg: Council of Europe, 2018. Disponível em: <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>. Acesso em: 8 jan. 2025.

BORGESIUUS, Frederik Zuiderveen. Price discrimination, algorithmic decision-making, and european non-discrimination law. **European Business Law Review**, London, v. 31, n. 3, p. 401–422, May 2020. Disponível em: <https://www.proquest.com/scholarly-journals/price-discrimination-algorithmic-decision-making/docview/3125960151/se-2?accountid=44316>. Acesso em: 8 jan. 2025.

BROUSSARD, Meredith. **More than a glitch**: confronting race, gender, and ability bias in tech. Cambridge: The MIT Press, 2023.

BUOLAMWINI, Joy. **Unmasking AI**: my mission to protect what is human in a world of machines. New York: Random House, 2023.

COECKELBERGH, Mark. **Ética na inteligência artificial**. Rio de Janeiro: PUC-Rio, 2023. (Coleção Exit).

COITINHO, Denis; SILVA, André Luiz Olivier da. Algorithmic injustice and human rights. **Filosofia Unisinos**, v. 25, p. e25109, 14 June 2024. Disponível em: <https://www.scielo.br/j/fun/a/QgphbcgsqMzvPFbXVTRpJNz/>. Acesso em: 8 jan. 2025.

COMISSÃO EUROPEIA. **Livro branco sobre a inteligência artificial**: uma abordagem europeia virada para a excelência e a confiança. Bruxelas: Comissão Europeia, 2020. Disponível em: <https://op.europa.eu/pt/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1>. Acesso em: 8 jan. 2025.

CRIADO PEREZ, Caroline. **Mulheres invisíveis**: o viés dos dados em um mundo projetado para homens. Rio de Janeiro: Intrínseca, 2022.

CUSTERS, Bart. Data dilemmas in the information society. *In*: CUSTERS, Bart; CALDERS, Toon; SCHERMER, Bart; ZARSKY, Tal (org.). **Discrimination and privacy in the information society**: data mining and profiling in large databases. Berlin: Springer, 2013. Disponível em: <https://papers.ssrn.com/abstract=3047756>. Acesso em: 8 jan. 2025.

DAS, Sunit; DUBBER, Markus Dirk; PASQUALE, Frank (org.). **The Oxford handbook of ethics of AI**. New York, NY: Oxford University Press, 2021.

DUBAL, Veena. On algorithmic wage discrimination. **Columbia Law Review**, New York, v. 123, n. 7, p. 1929–1992, Nov. 2023. Disponível em: <https://research.ebsco.com/linkprocessor/plink?id=b4976901-fb39-31b6-8882-bd361a38bc6d>. Acesso em: 8 jan. 2025.

FEFERBAUM, Marina; SILVA, Alexandre Pacheco da (org.). **Ética, governança e inteligência artificial**. São Paulo: Almedina, 2023. (Coleção Obras coletivas).

FIORAVANTI, Carlos; VASCONCELOS, Yuri. Graça Nunes: IA sem ilusões. **Pesquisa FAPESP**, São Paulo, n. 345, nov. 2024. Disponível em: <https://revistapesquisa.fapesp.br/graca-nunes-ia-sem-ilusoes/>. Acesso em: 8 jan. 2025.

FRANKLIN, Gillian; STEPHENS, Rachel; PIRACHA, Muhammad; TIOSANO, Shmuel; LEHOULLIER, Frank; KOPPEL, Ross; ELKIN, Peter L. The sociodemographic biases in machine learning algorithms: a biomedical informatics perspective. **Life**, Basel, v. 14, n. 6,

p. 652, 1 June 2024. Disponível em: <https://doi.org/10.3390/life14060652>. Acesso em: 8 jan. 2025.

HALL, Paula; ELLIS, Debbie. A systematic review of socio-technical gender bias in AI algorithms. **Online Information Review**, Leeds, v. 47, n. 7, p. 1264–1279, Nov. 2023. Disponível em: <https://doi.org/10.1108/OIR-08-2021-0452>. Acesso em: 8 jan. 2025.

HERVIEUX, Sandy; WHEATLY, Amanda (org.). **The rise of AI: implications and applications of artificial intelligence in academic libraries**. Chicago: Association of College and Research Libraries, 2022. (ACRL publications in librarianship, n. 78).

KARASINSKI, Murilo; CANDIOTTO, Kleber Bez Birolo. AI's black box and the supremacy of standards. **Filosofia Unisinos**, São Leopoldo, RS, v. 25, n. 1, p. e25113, jun. 2024. Disponível em: <https://www.scielo.br/j/fun/a/YbzcpgB8gGvLS3rBQVrNpRs/>. Acesso em: 8 jan. 2025.

KAUFMAN, Dora. Logic and foundations of artificial intelligence and society's reactions to maximize benefits and mitigate harm. **Filosofia Unisinos**, São Leopoldo, RS, v. 25, n. 1, p. e25110, 14 jun. 2024. Disponível em: <https://www.scielo.br/j/fun/a/LqLKT488FQMbRxdw4sxFLKF/>. Acesso em: 8 jan. 2025.

LIEBIG, Laura; JOBIN, Anna; GÜTTEL, Licinia; KATZENBACH, Christian. Situating AI policy: controversies covered and the normalisation of AI. **Big Data & Society**, London, v. 11, n. 4, Dec. 2024. Disponível em: 10.1177/20539517241299725. Acesso em: 8 jan. 2025.

LIN, Shu; PANDIT, Saket; TRITSCH, Tara; LEVY, Arkene; SHOJA, Mohammadali M. What goes in, must come out: generative artificial intelligence does not present algorithmic bias across race and gender in medical residency specialties. **Cureus**, United States, v. 16, n. 2, p. e54448, 19 Feb. 2024. Disponível em: <https://doi.org/10.7759/cureus.54448>. Acesso em: 8 jan. 2025.

MACHADO SEGUNDO, Hugo de Brito. **Direito e inteligência artificial: o que os algoritmos têm a ensinar sobre interpretação, valores e justiça**. Indaiatuba: Foco, 2023.

MAVROGIORGOS, Konstantinos; KIOURTIS, Athanasios; MAVROGIORGOU, Argyro; MENYCHTAS, Andreas; KYRIAZIS, Dimosthenis. Bias in machine learning: a literature review. **Applied Sciences**, v. 14, n. 19, p. 8860, Oct. 2024. Disponível em: <https://doi.org/10.3390/app14198860>. Acesso em: 8 jan. 2025.

OLIVEIRA, Nythamar H. de. A decolonial critical theory of artificial intelligence: intersectional egalitarianism, moral alignment, and AI governance. **Filosofia Unisinos**, São Leopoldo, RS, v. 25, n. 1, p. e25114, 14 jun. 2024. Disponível em:

<https://www.scielo.br/j/fun/a/g5MwkPjfQfZ4m3Mf59CNJnk/?lang=en>. Acesso em: 8 jan. 2025.

PÜTTEN, Astrid Marieke Rosenthal-von der; SACH, Alexandra. Michael is better than Mehmet: exploring the perils of algorithmic biases and selective adherence to advice from automated decision support systems in hiring. **Frontiers in Psychology**, Yunnan, v. 15, Sept. 2024. Disponível em: <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2024.1416504/full>. Acesso em: 8 jan. 2025.

RAGHAVAN, Manish. Mitigating bias in algorithmic hiring: evaluating claims and practices. In: RAGHAVAN, Manish. **The societal impacts of algorithmic decision-making**. New York: Association for Computing Machinery, 2023a. Disponível em: <https://doi.org/10.1145/3603195.3603203>. Acesso em: 8 jan. 2025.

RAGHAVAN, Manish. **The societal impacts of algorithmic decision-making**. New York: Association for Computing Machinery, 2023b.

RICHARDSON, Rashida. Racial segregation and the data-driven society: how our failure to reckon with root causes perpetuates separate and unequal realities. **Berkeley Technology Law Journal**, Berkeley, v. 36, n. 3, p. 1051–1090, 2022. Disponível em: <https://research.ebsco.com/linkprocessor/plink?id=35fdae7-4513-3121-9a66-3088cbe46077>. Acesso em: 8 jan. 2025.

RHOADES, Gary. Ghosts in the machine: how past and present biases haunt algorithmic tenant screening systems. **Human Rights**, v. 49, n. 4, p. 13–15, June 2024. Disponível em: https://www.americanbar.org/groups/crsj/publications/human_rights_magazine_home/technology-and-the-law/algorithmic-tenant-screening-systems/. Acesso em: 8 jan. 2025.

RODRIGUES, Priscila Lauande. Procedimentos de tomada de decisão, preconceito e discriminação. In: RODRIGUES, Priscila Lauande. **Inteligência artificial nas relações de trabalho**. Brasília: Venturoli, 2024.

ROSA, Alexandre Moraes da; GUASQUE, Bárbara. Inteligência artificial, vieses algorítmicos e racismo: o lado desconhecido da justiça algorítmica. **Opinión Jurídica**, Medellín, v. 23, n. 50, p. 1–23, jul./dic. 2024. Disponível em: <https://revistas.udem.edu.co/index.php/opinion/article/view/4046/3797>. Acesso em: 8 jan. 2025.

ROSSETTI, Regina; ANGELUCI, Alan. Ética Algorítmica: questões e desafios éticos do avanço tecnológico da sociedade da informação. **Galáxia**, São Paulo, p. e50301, 16 jul. 2021. Disponível em: <https://doi.org/10.1590/1982-2553202150301>. Acesso em: 8 jan.

2025.

SANTOS, Rodrigo Coimbra; GRAMINHO, Vivian Maria Caxambu. Algorithmic discrimination in labor relations and principles of the General Data Protection Law. **Sequência**, Florianópolis, v. 45, p. e96294, 14 out. 2024. Disponível em: <https://www.scielo.br/j/seq/a/9FGhvr75HYSdGMRfKnJ9FZd/abstract/?lang=en>. Acesso em: 8 jan. 2025.

SCHWEDE, Matheus Antes. **Inteligência artificial tendenciosa**: discriminação racial pelo viés algorítmico. Londrina: Thoth, 2023.

SHNEIDERMAN, Ben. **Human-centered AI**. New York: Oxford University Press, 2022.

TALIA, Domenico. Algorithm biases and values. In: TALIA, Domenico. **From algorithms to thinking machines**: the new digital power. New York: Association for Computing Machinery, 2023. Disponível em: <https://doi.org/10.1145/3603178.3603186>. Acesso em: 8 jan. 2025.

UKANWA, Kalinda. Algorithmic bias: social science research integration through the 3-D Dependable AI Framework. **Current Opinion in Psychology**, v. 58, 2024. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2352250X24000496>. Acesso em: 8 jan. 2025.

VALENTE, Mariana. **Misoginia na internet**: uma década de disputas por direitos. São Paulo: Fósforo, 2023.

VIDINA, Milla (org.). **Regulating for european AI that protects and advances equality**. Brussels: Equinet, European Network of Equality Bodies, 2021. Disponível em: https://www.defenseurdesdroits.fr/sites/default/files/atoms/files/avis_equinet_sur_intelligence_artificielle_et_legalite.pdf. Acesso em: 8 jan. 2025.

WANG, Xukang; WU, Ying Cheng; JI, Xueliang; FU, Hongpeng. Algorithmic discrimination: examining its types and regulatory measures with emphasis on US legal practices. **Frontiers in Artificial Intelligence**, v. 7, May 2024. Disponível em: <https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2024.1320277/full>. Acesso em: 8 jan. 2025.

ŽIVKOVIĆ, Patricia; DUCATO, Rossana. Algorithmic discrimination: a blueprint for a legal analysis. **EU & Comparative Law Issues & Challenges Series (ECLIC)**, Osijek, v. 7, p. 202–226, Jan. 2023. Disponível em: <https://research.ebsco.com/linkprocessor/plink?id=fd8d630b-a3e1-39f9-b3ba-0cc931c6e4b7>. Acesso em: 8 jan. 2025.