

**TERMO DE REFERÊNCIA TEMÁTICO:
INTELIGÊNCIA ARTIFICIAL E DESIGUALDADES EM SAÚDE**



Fevereiro de 2021



NETHIS
NÚCLEO DE ESTUDOS SOBRE
BIOÉTICA E DIPLOMACIA EM SAÚDE



Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz
Brasília

Inteligência Artificial e Desigualdades em Saúde¹

Imagine um sistema que pressupõe diferentes graus de propensão ao crime de acordo com a cor de um indivíduo ou que induz ofertas de trabalho de menor remuneração às mulheres em comparação aos homens. Estudos recentes apontam algo nesse sentido que merece atenção, ao demonstrar que sistemas baseados em algoritmos reproduzem e amplificam conceitos e valorações oriundos de seus formuladores, muitas vezes invisíveis até para os gestores que comandam seu funcionamento. É possível, assim, que iniquidades detectadas nesses estudos, em especial as vinculadas ao racismo e ao sexismo, agravem-se com a intensificação do uso da inteligência artificial (IA).

O progressivo avanço da IA, em muitos aspectos da vida, é particularmente notável na gestão administrativa, tanto na esfera pública como na privada. A seleção de pessoal, o filtro de candidatos para vagas em universidades, a categorização de cidadãos como elegíveis para programas sociais e até os cálculos para a previdência são basicamente realizados com bancos de dados e algoritmos de processamento. Boa parte dos seguros (de saúde e outros) usam algum processo automatizado que “decide” os assuntos que serão analisados ou jogados no lixo. O mesmo acontece nos softwares de processamento de relatórios preliminares nos grandes estudos jurídicos e em muitos outros âmbitos. Os sistemas automatizados estão tomando conta, por exemplo, da seleção de estudos científicos para produzir evidências.

Poderia se pensar que tudo isso garante eficiência e imparcialidade, já que esses sistemas executam sequências de tarefas obedecendo critérios de forma impessoal. Mas seria de fato sem a participação de pessoas? Os sistemas automatizados podem condicionar ou agravar desigualdades e injustiças, em razão de determinações incrustadas em sua lógica computacional, com o agravante de que a própria natureza de autômatos impede que suas decisões sejam auditáveis ou sequer discutíveis.

Contudo, tais inovações representam enorme potencial para ampliar benefícios à saúde individual e coletiva, mediante a expansão do acesso e da qualidade desses bens e serviços; podem, inclusive, auxiliar no alcance das metas de Desenvolvimento Sustentável (ODS/ONU). Advertência em contrário é que as vantagens auferidas tendem a privilegiar os que podem pagar, além de representar riscos maiores para os pobres.

O propósito desta linha de reflexões, debates e investigações é apoiar a sistematização, a formulação e a difusão de estratégias e instrumentos de salvaguardas éticas e de garantias de equidade em mecanismos que, rapidamente, vão ficando fora do escrutínio público, quando deveriam operar somente sob orientação e regulação do poder público. Em suma, a presente iniciativa busca contribuir para o desafio: a IA no campo da saúde pode escapar dessa tendência e ajudar a construir um mundo mais justo e humano?

¹ Termo de Referência sobre Inteligência Artificial e Desigualdades em Saúde, do Programa de Pesquisa e Comunicação do Nethis/Fiocruz, compartilhado com o Cepedisa/FSP/Usp.

Referências:

1. Noble S. **Algorithms of Oppression**. NYU Press, 2018;
2. Eubanks V. **Automating Inequality**. St. Martin Press, 2019;
3. American Civil Liberties Union Foundation. **Racial Disparities in Sentencing**. Submitted to the Inter-American Commission on Human Rights. 153rd Session, October 27, 2014. Disponível em: https://www.aclu.org/sites/default/files/assets/141027_iachr_racial_disparities_aclu_submission_0.pdf. Acesso em: 4 fev. 2021;
4. Howard A, Borenstein J. **The Ugly Truth About Ourselves and Our Robot Creations**: The Problem of Bias and Social Inequity. *Social and Engineering Ethics*. October 2018, Volume 24;
5. Angwin J, Larson J, Mattu S, Kirchner L. **Machine Bias**: There's software used across the country to predict future criminals. And it's biased against blacks. *ProPublica*, May 23, 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 4 fev. 2021;
6. Bian L, Leslie SJ, Cimpian A. Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, 355(6323), 389-391, 2017;
7. Bogost I. 'Artificial Intelligence' has become meaningless. *The Atlantic*, March 4, 2017. Disponível em: <https://www.theatlantic.com/technology/archive/2017/03/what-is-artificial-intelligence/518547>. Acesso em: 4 fev. 2021;
8. Bolukbasi T, Chang K-W, Zou J, Saligrama V, Kalai A. Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. 2016. Disponível em: <https://arxiv.org/pdf/1607.06520.pdf>. Acesso em: 4 fev. 2021;
9. Borenstein J, Herkert J, Miller K. Self-driving cars: Ethical responsibilities of design engineers. *IEEE Technology and Society Magazine*, 36(2), 67-75, 2017a;
10. Borenstein J, Howard A, Wagner A. Pediatric robotics and ethics: The robot is ready to see you now but should it be trusted? In: Lin P, Abney K, Bekey G (Eds.). **Robot ethics 2.0**. Oxford: Oxford University Press, 2017b;
11. Brownstein M. (2016). Implicit Bias. In: Zalta EN (Ed.). *The stanford encyclopedia of philosophy*. Disponível em: <https://plato.stanford.edu/archives/win2016/entries/implicit-bias/>. Acesso em: 4 fev. 2021;
12. Caliskan A, Bryson JJ, Narayanan A. Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183-186, 2017;
13. Carpenter J. Google's algorithm shows prestigious job ads to men, but not to women. *Independent*, July 7, 2015. Disponível em: <http://www.independent.co.uk/life-style/gadgets-and-tech/news/googles-algorithm-shows-prestigious-job-ads-to-men-but-not-to-women-10372166.html>. Acesso em: 4 fev. 2021;
14. Carty SS. Many cars tone deaf to women's voices. *Autoblog*, May 31, 2011. Disponível em: <http://www.autoblog.com/2011/05/31/women-voice-command-systems/>. Acesso em: 4 fev. 2021;
15. Castellanos S. Capital One pursues 'explainable AI' to guard against bias in models. *The Wall Street Journal*, December 2, 2016. Disponível em: <http://blogs.wsj.com/cio/2016/12/06/capital-one-pursues-explainable-ai-to-guard-against-bias-in-models/>. Acesso em: 10 fev. 2017;
16. Chavalarias D, Ioannidis JPA. (2010). Science mapping analysis characterizes 235 biases in biomedical research. *Journal of Clinical Epidemiology*, 63(11), 1205-1215, 2010. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0895435610000223>. Acesso em: 4 fev. 2021;
17. Chawla NV, Hall LO, Bowyer KW, Kegelmeyer WP. SMOTE: Synthetic minority oversampling technique. *Journal of Artificial Intelligence Research*, 16, 321-357, 2002. Disponível em: <https://www.jair.org/index.php/jair/article/view/10302>. Acesso em: 4 fev. 2021;
18. Chayes J. How machine learning advances will improve the fairness of algorithms. *Huffington Post*, August 23, 2017. Disponível em: [Google Scholar](#). Acesso em: 25 ago. 2017;
19. Datta A, Tschantz MC., Datta A. Automated experiments on ad privacy settings. *Proceedings on Privacy Enhancing Technologies*, 2015(1), 92-112, 2015. Disponível em: [CrossRefGoogle Scholar](#). Acesso em: 4 fev. 2021.