

Apresentação do dossiê ética da inteligência artificial

Denis Coitinho

<https://orcid.org/0000-0002-2592-5590>

Universidade do Vale do Rio dos Sinos - Unisinos, Programa de Pós-Graduação em Filosofia, São Leopoldo, RS, Brasil. Email: deniscs@unisinos.br

Marcelo de Araujo

<https://orcid.org/0000-0001-8934-6195>

Universidade do Estado do Rio de Janeiro - UERJ, Departamento de Filosofia, e Universidade Federal do Rio de Janeiro – UFRJ, Faculdade de Direito. Rio de Janeiro, RJ, Brasil. Email: marceloaraujo@direito.ufrj.br

Apresentação

A tecnologia orientada por Inteligência Artificial (IA), entendida como a capacidade de um sistema – como um software ou incorporada em um aparelho – para executar tarefas comumente associadas a seres inteligentes, já é uma constante em nossas vidas, seja pelas recorrentes recomendações de filmes que recebemos pelos serviços de streaming, tal como Netflix e Amazon Prime, seja pelos filtros de e-mails que usamos, ou pela utilização de algum aplicativo de navegação por GPS, como o Waze, e ela não parece problemática em muitas áreas. Ao contrário, ela parece facilitar nossa vida. Entretanto, esta tecnologia está se estendendo progressivamente para certos domínios nos quais, provavelmente, terá um impacto maior, assim como decidir em circunstâncias de risco, estabelecer prioridades entre pessoas e fazer julgamentos complexos e, portanto, terá que tomar decisões que já se enquadram no domínio moral. Por isso, parece importante pensar sobre os algoritmos que alimentam esses produtos. Os carros autônomos, por exemplo, precisarão tomar decisões sobre como distribuir o risco entre os passageiros, pedestres e ciclistas, isto é, entre os que utilizam as vias públicas. As armas autônomas letais, terão que identificar e selecionar os alvos humanos que serão eliminados. Por sua vez, algoritmos que já estão em uso no sistema de saúde e judiciário em alguns países, estabelecem a prioridade de quem receberá um transplante de órgão, bem como aconselham os juízes sobre quem deve obter liberdade condicional ou uma sentença maior de prisão.

A promessa da IA é que ela melhore nossa vida, pois esses algoritmos podem, teoricamente, decidir sem vieses de gênero, raça ou classe, de forma lógica e racional, evitando qualquer discriminação. Entretanto, pode ser que isso não ocorra tal como esperado, uma vez que todas essas decisões referidas anteriormente, inevitavelmente, parecem envolver padrões éticos e avaliações morais complexas, o que pode resultar em injustiças. Por exemplo, os carros autônomos devem sempre se esforçar para minimizar as baixas, mesmo que às vezes isso signifique sacrificar seus próprios passageiros para um bem maior? As armas autônomas letais devem sempre objetivar a vitória, mesmo ao custo de eliminar um alvo civil ou um soldado ferido, o que colocaria em risco a dignidade humana? As crianças devem sempre ter prioridade para transplantes de órgãos, mesmo quando um paciente mais velho é uma combinação genética melhor para um órgão disponível? Os algoritmos usados em tribunais devem sempre procurar reduzir a reincidência, mesmo resultando em uma discriminação injusta para os réus negros?

O problema que nos deparemos atualmente é que alguns usos da IA utilizados em diferentes contextos, como no trânsito, educação, saúde, judiciário, política etc., a exemplo dos programas de reconhecimento facial, avaliação de currículos, classificação de fotografias por assunto e programas que influenciam as decisões judiciais no âmbito penal, tal como o COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), têm revelado disparidades preocupantes de desempenho com base na raça e no gênero das pessoas. Essas disparidades levantam questões urgentes sobre como o uso da IA pode funcionar para promover a justiça ou consolidar a injustiça, podendo revelar, por exemplo, discriminação institucional, injustiça estrutural, como racismo, sexismo e preconceito de classe.

Com essa urgência em mente, é com satisfação que apresentamos à comunidade filosófica brasileira um dossiê sobre os desafios éticos trazidos pela IA. O primeiro artigo, de Celso Azambuja e Gabriel Ferreira, trata dos efeitos da IA na educação, enfatizando a necessidade de repensar as estratégias pedagógicas nas universidades, explorando as transformações trazidas pela IA, incluindo a automação e o potencial de personalização do ensino. O artigo seguinte, por sua vez, aborda a importante questão da necessidade de regulação ética e legal do uso da IA, considerando a perspectiva humana de seu uso e os impactos sobre a humanidade e é de autoria de Delamar Dutra. O terceiro artigo, de André Olivier e Denis Coitinho, investiga o problema da injustiça algorítmica, tendo por foco central a defesa dos direitos humanos e a proteção aos mais vulneráveis. O quarto artigo é de Dora Kaufman, importante especialista no tema em tela, que reflete sobre a necessidade de maximizar os benefícios e minimizar os possíveis danos sociais trazidos pela IA. Na sequência, o próximo texto propõe a uma análise centrada no lapso emocional da IA, enfocando sua influência em dois aspectos cruciais. Os autores, Evandro Barbosa e Thais Costa, exploram o impacto ontológico do lapso emocional na inabilidade das máquinas em desenvolver um senso moral (Gert, 2016; Kauppinen, 2022), bem como tematizam a questão prática que surge nas relações humano-máquina e seu impacto na relação humana de amizade. Em seguida, Murilo Vilaça, Isabella Pederneira e Mariza Ferro, apresentam abordagens multi/interdisciplinares, delimitadas, realistas e ponderadas sobre os desafios da IA. O sétimo artigo é de Murilo Karasinski e Kleber Candioto que problematiza a questão da caixa preta da IA (opacidade) (Adadi, Berrada, 2018) e a supremacia dos padrões. Já o oitavo artigo, de Nythamar de Oliveira, defende uma abordagem decolonial (De Oliveira, 1996) da teoria crítica para tratar do problema da IA, defendendo um igualitarismo interseccional. O Dossiê se encerra com uma importante entrevista feita com Luciano Floridi, especialista de renome internacional sobre a ética da IA, conduzida por Murilo Karasinski, Murilo Vilaça e Kleber Candioto.

Esperamos que esse dossiê possa auxiliar nesse novo campo de investigação em ética, de forma a fomentar o debate na comunidade filosófica brasileira, destacando a importância de refletirmos no tempo presente e de pensarmos em alternativas aos possíveis problemas que concernem a toda a sociedade.

Referências

- ADADI, A.; BERRADA, M. 2018. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE access*, **6**: p. 52138-52160, 2018.
- DE OLIVEIRA, N. 1996. The Worldhood of the World in Heidegger's Reading of Heraclitus. *Manuscrito*, XIX (2): p. 200-223. <https://periodicos.sbu.unicamp.br/ojs/index.php/manuscrito/article/view/8669093/28413>.
- GERT, J. 2016. Moral sentimentalism. *Routledge Encyclopedia of Philosophy*. Routledge. DOI.org (Crossref), <https://doi.org/10.4324/9780415249126-L3578-1>.
- KAUPPINEN, A. 2022. Moral Sentimentalism. *The Stanford Encyclopedia of Philosophy*, organizado por Edward N. Zalta, Spring 2022, Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/spr2022/entries/moral-sentimentalism/>.