



Resenha: Gouveia, Steven S. *Thinking the New World: Conversations on Artificial Intelligence*. Braga, Portugal: Independently published, 2022

Romeu Ivoleta Neto

Mestre (2019) e investigador de doutoramento no MLAG do Instituto Filosofia da Universidade do Porto. Possui graduação em tecnologia em análise e desenvolvimento de sistemas web (2008) e especialização em história da filosofia pela PUC-SP (2014). Atua como consultor de sistemas web e inteligência artificial para startups brasileiras. Atualmente trabalha na pesquisa de sua tese de doutorado acerca dos limites hermenêuticos na IA via softwares. Mais informações: <http://romeuivolela.com.br>

RESUMO:

A resenha de *Thinking the New World* explora o campo em rápido crescimento da ética em Inteligência Artificial (IA). Ao abordar temas como armas autônomas, ética de máquinas e impactos sociais, Steven S. Gouveia, mediante treze entrevistas, esclarece tanto os benefícios quanto os perigos dessa tecnologia. Com essas discussões, o livro expõe a necessidade urgente de estabelecer padrões éticos no desenvolvimento e uso da IA. Os pesquisadores oferecem perspectivas diversas sobre os temas abordados, tornando-se um recurso valioso para qualquer pessoa que busque uma compreensão mais profunda das complexas questões éticas envolvendo a IA no mundo atual.

PALAVRAS-CHAVE:

Ética; Filosofia; Inteligência Artificial; Machine Learning.

ABSTRACT:

The review of *Thinking the New World* explores the rapidly growing field of ethics in Artificial Intelligence (AI). By addressing topics such as autonomous weapons, machine ethics, and social impacts, Steven S. Gouveia, through thirteen interviews, elucidates both the benefits and dangers of this technology. Through these discussions, the book exposes the urgent need to establish ethical standards in the development and use of AI. Researchers offer diverse perspectives on the topics discussed, making it a valuable resource for anyone seeking a more profound understanding of the complex ethical issues involving AI in the current world.

KEYWORDS:

Artificial Intelligence; Ethics; Machine Learning; Philosophy.



INTRODUÇÃO

Thinking the New World, editada pelo Dr. em (Neuro)filosofia da mente Steven S. Gouveia (2022), explora as implicações filosóficas e éticas da inteligência artificial (IA) por meio de discussões informais com treze especialistas no campo. Baseada no documentário internacional *The age of artificial intelligence: the documentary*¹⁰⁷, é uma introdução que mergulha o leitor nos desafios éticos contemporâneos provocados pelo avanço da IA. As conversas abordam uma ampla gama de tópicos: sexualidade, privacidade, segurança, relacionamentos pessoais, renda mínima, uso de armas e carros autônomos, IA na medicina, consciência na IA, para mencionar apenas alguns.

CONSIDERAÇÕES ÉTICAS NA IA: REFLEXÕES E DESAFIOS

Na primeira entrevista, Wulf Loh, Professor Assistente na Universidade de Tübingen (GOUVEIA, 2022, p. 25–46), argumenta que os algoritmos que operam sistemas de *Big Data*¹⁰⁸ e *Machine Learning*¹⁰⁹ são pervasivos e a atribuição de decisão a eles um tema sensível, já que basicamente conduzem a dilemas éticos (GOUVEIA, 2022, p. 26–31). Esses dilemas na IA exigirão a cooperação entre filosofia e dados empíricos, a partir de uma abordagem cooperativa teórica e

¹⁰⁷ “The Age of Artificial Intelligence: The Documentary”. 2020. Disponível em: <https://www.youtube.com/watch?v=zMrt6ALaio0>. Acessado em 2023.

¹⁰⁸ *Big Data* são conjuntos de dados grandes produzidos em formato digital e que podem ser analisados por meio de ferramentas computacionais. Portanto, as duas características mais comumente associadas ao *Big Data* são volume e velocidade. Volume refere-se ao tamanho dos arquivos usados para arquivar e distribuir dados. Velocidade refere-se à rapidez com que os dados são gerados e processados. No caso da produção científica, o corpo de dados digitais criado por pesquisas está crescendo a uma velocidade vertiginosa e de maneiras que são discutivelmente impossíveis para o sistema cognitivo humano compreender e, portanto, exigem alguma forma de análise automatizada (LEONELLI, 2020, p. 3-8, grifo do autor, tradução nossa).

¹⁰⁹ O campo da pesquisa em ciência da computação que se concentra no desenvolvimento e avaliação de algoritmos que permitem que computadores aprendam com a experiência é chamado de *machine learning*. Geralmente, o conceito de experiência é representado como um conjunto de dados de eventos históricos, e a aprendizagem envolve identificar e extrair padrões úteis de um conjunto de dados. Um algoritmo de *machine learning* recebe um conjunto de dados como entrada e retorna um modelo que codifica os padrões que o algoritmo extrai (ou aprendeu) dos dados (KELLEHER, 2019, p. 253, grifo do autor, tradução nossa).



prática, resultando em ações que possam orientar regulações governamentais e de mercado, como a operação dos carros autônomos (GOUVEIA, 2022, p. 38–42).

Na sequência, Fabio Fossa, Pesquisador Pós-Doutorado no Politécnico de Milão (GOUVEIA, 2022, p. 47–62), a conversa explora a democratização na discussão de ética de máquinas e os desafios de integrar sua autonomia com o julgamento humano, bem como a importância de regular sistemas de IA para a implementação de valores éticos (GOUVEIA, 2022, p. 48). Fossa enfatiza a importância das considerações contextuais na implementação de valores éticos na IA, a qual deve ser abordada de forma democrática, entretanto, exigindo especialização teórica e técnica na sua implementação (GOUVEIA, 2022, p. 51–54).

Na terceira entrevista, Shawn Kaplan, Professor de Filosofia na Universidade de Adelphi (GOUVEIA, 2022, p. 25–46), a conversa aprofunda as complexidades éticas impostas pelo uso de armas autônomas, tratando particularmente o dilema filosófico do *responsibility gap* na IA (GOUVEIA, 2022, p. 65–66). O dilema evidencia a complexidade em determinar a responsabilidade moral e legal em casos de danos injustificados a civis, devido a erros de direcionamento ou situações semelhantes causadas pela IA, que ocorrem em um contexto impreciso de aprendizado de máquina e de manipulação de grandes quantidades de informações (GOUVEIA, 2022, p. 66–67).

Com Radu Uszkai, Professor Assistente na Universidade de Estudos Econômicos de Bucareste (GOUVEIA, 2022, p. 69–96), a entrevista mergulha em uma série de preocupações éticas prementes relacionadas ao desenvolvimento e implementação de robôs sexuais (GOUVEIA, 2022, p. 70). Existem perspectivas que sugerem que esses robôs podem atenuar alguns dos aspectos negativos do comércio sexual (GOUVEIA, 2022, p. 71). Uszkai também especula que, à medida que a tecnologia progride, mais indivíduos podem optar por relacionamentos



íntimos com andróides em vez de alternativas ilegais, como abuso e exploração sexual (GOUVEIA, 2022, p. 73).

Na quinta entrevista, Joshua Jowitt, professor na Universidade de Newcastle (GOUVEIA, 2022, p. 97–114), a entrevista explora uma variedade de tópicos provocativos na interseção da ética, do direito e da IA. O interesse de Jowitt pelo conceito de personalidade jurídica e moral para agentes artificiais leva a uma discussão sobre a conexão intrínseca entre o direito e a moral. Ele enfatiza que a personalidade jurídica não pode existir sem uma base moral fundamental, a qual se aplica à agentes naturais e artificiais (GOUVEIA, 2022, p. 98–108).

Na sexta entrevista, Francesca Minerva, Pesquisadora da Universidade de Milão (GOUVEIA, 2022, p. 115 – 134), aponta para o desafio que a sociedade enfrenta em influenciar o desenvolvimento da IA à medida que esta se torna cada vez mais inteligente (GOUVEIA, 2022, p. 117). Ela também explora a possibilidade de integração de seres humanos com a IA para uma vida prolongada, incluindo a substituição de órgãos biológicos por componentes biônicos (GOUVEIA, 2022, p. 118–124).

Com o renomado Peter Singer, Professor Ira W. DeCamp de Bioética na Universidade de Princeton (GOUVEIA, 2022, p. 135–146), a entrevista adentra no potencial da tecnologia para aperfeiçoar as capacidades humanas, defendendo a melhoria do bem-estar da sociedade antes de considerar aprimoramentos para níveis *supernormais* do homem (GOUVEIA, 2022, p. 139–140). Olhando para o futuro, Singer imagina o melhor resultado da IA no fornecimento de ferramentas mais eficientes para a resolução de problemas. Ao mesmo tempo, ele reconhece os desafios apresentados pelo desemprego em massa, destacando a importância de oferecer às pessoas não apenas apoio financeiro, mas também um senso de propósito em suas vidas (GOUVEIA, 2022, p. 144–145).

Na oitava entrevista, Shoji Nagataki, Professor na Universidade de Chukyo (GOUVEIA, 2022, p. 147–162), contempla a elusiva questão de identificar a



consciência artificial e o desafio de alcançar empatia com tais sistemas (GOUVEIA, 2022, p. 148–149). Nagataki também especula sobre a possibilidade da IA e robôs inteligentes terem os seus próprios direitos artificiais reconhecidos na sociedade, influenciando potencialmente na criação de novas leis que atendam aos seus interesses (GOUVEIA, 2022, p. 156).

Na conversa com Paul Thagard, Professor Emérito da Universidade de Waterloo (GOUVEIA, 2022, p. 163–194), é discutido o papel dos governos na supervisão do desenvolvimento da IA para assegurar a sua transparência e alinhamento com as necessidades humanas (GOUVEIA, 2022, p. 164–167). A conversa ainda trata dos aspectos positivos e negativos das armas autônomas, abordando a importância da empatia em cenários de conflitos militares, característica ausente nos robôs, e incidindo em questões éticas no uso dessa tecnologia (GOUVEIA, 2022, p. 170).

Em Hajo Greif, Professor da Universidade de Tecnologia de Varsóvia (GOUVEIA, 2022, p. 195–208), o entrevistado mostra que a mera emulação da estrutura e funções do cérebro humano não são suficientes para alcançar uma IA no sentido forte ¹¹⁰. Destaca que este tipo de programa deve considerar o papel do corpo e do ambiente no processo cognitivo (GOUVEIA, 2022, p. 196). Greif também destaca a importância do termo *transparência* na filosofia da simulação computacional, referindo-se à sua clareza epistêmica, diferente do significado atribuído na tecnologia de vigilância. A ausência de transparência epistêmica acarreta desafios na explicação das ocorrências em sistemas de IA e na compreensão de seu impacto nos seres humanos (GOUVEIA, 2022, p. 206–207).

Com David Harris Smith, Professor associado na Universidade McMaster (GOUVEIA, 2022, p. 209–226), a entrevista enfatiza a importância das simulações

¹¹⁰ Conforme a distinção de John Searle entre *IA Forte* e *IA Fraca*, a *IA Forte* visa desenvolver pessoas artificiais com capacidades mentais completas, incluindo consciência, enquanto a *IA Fraca* concentra-se em construir máquinas de processamento de informações que simulam habilidades mentais humanas (SEARLE, 1997).



computacionais na visualização e planejamento público, destacando como essas simulações auxiliam na compreensão de mudanças e demonstração de projetos urbanos (GOUVEIA, 2022, p. 210). Ele aborda questões filosóficas sobre como os comportamentos e as intenções podem mudar dependendo da relação humana com um mundo virtual, tratando-o como separado ou contíguo (GOUVEIA, 2022, p. 211).

Na penúltima entrevista, Pii Telakivi, Pesquisadora Pós-Doutorada na Universidade de Helsinki (GOUVEIA, 2022, p. 227–240), enfatiza que as máquinas de hoje não são inteligentes ou conscientes no sentido estrito desses conceitos, e por isto o uso do termo *inteligência artificial* é usado equivocadamente (GOUVEIA, 2022, p. 228). Ela também explora a relação complexa entre a condição necessária, mas não suficiente, da consciência para a moralidade na agência artificial (GOUVEIA, 2022, p. 229). Telakivi rejeita a ideia de que a biologia seja o único caminho para criar a consciência, inclinando-se em vez disso para o enativismo e a teoria da mente incorporada, defendendo um foco nas interações sensorimotoras entre agentes e seu ambiente (GOUVEIA, 2022, p. 230).

Na última entrevista com Sabina Leonelli, professora na Universidade de Exeter (GOUVEIA, 2022, p. 241–266), são explorados diversos aspectos da relação entre dados e IA. Ela enfatiza que o *Big Data*, apesar de sua vastidão, frequentemente é incompleto, tendencioso e limitado, desafiando a noção geral de compreensibilidade dos dados usados pela IA (GOUVEIA, 2022, p. 242 – 244). Leonelli discute as limitações do uso da IA na saúde e na lei, observando que a contextualização de informações requer julgamento humano com base em seu ambiente, algo que sistemas de *Machine Learning* não conseguem realizar (GOUVEIA, 2022, p. 248 – 249).

CONSIDERAÇÕES FINAIS

A ética da IA é um campo em crescimento, ainda relativamente novo nos currículos acadêmicos e no discurso social. Apesar disso, as tecnologias de IA



permeiam vários aspectos da vida moderna, oferecendo tanto benefícios significativos quanto potenciais perigos. *Thinking the New World* destaca os desafios éticos que surgem em conjunto com esse avanço.

Enquanto a IA pode melhorar a eficiência e lidar com problemas complexos como mudanças climáticas, também apresenta riscos como discriminação, amplificação de viés, manipulação e consequências letais (por exemplo, armas autônomas). A compreensão profunda e urgente dessas questões, por filósofos, eticistas e especialistas, torna a leitura dessa obra indispensável para a apreensão de muitos dos problemas contemporâneos e futuros.



REFERÊNCIAS BIBLIOGRÁFICAS

GOUVEIA, Steven S. *Thinking the New World: Conversations on Artificial Intelligence*. Braga, Portugal: Independently published, 2022.

KELLEHER, J. D. **Deep learning**. Cambridge, Massachusetts: The MIT Press, 2019.

LEONELLI, SABINA. Scientific Research and Big Data. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Summer 2020 ed. [s.l.] Metaphysics Research Lab, Stanford University, 2020.

SEARLE, J. R. **The Mystery of Consciousness**. First Edition ed. New York: The New York Review of Books, 1997.