

# IA e Ética: a interface entre tecnologia, moralidade e sociedade

## AI and Ethics: the interface between technology, morality and society

### IA y ética: la interfaz entre tecnología, moralidad y sociedad

DOI:10.38152/bjtv8n1-026

Submitted: Feb 17<sup>th</sup>, 2025

Approved: Mar 07<sup>th</sup>, 2025

**Gabriel César Dias Lopes**

Doutor em Administração de Empresas - Especialização em Educação

Instituição: Université Américaine des Sciences et du Développement International  
(UNASDI)

Endereço: 3, Myriam II, Impasse Altidor, Port-de-Paix Haiti

E-mail: president@unilogos.edu.eu

#### RESUMO

**Introdução:** A interseção entre inteligência artificial (IA) e ética tem ganhado relevância acadêmica e social, exigindo reflexões sobre os impactos morais do desenvolvimento tecnológico. Este artigo analisa os desafios éticos da IA, destacando a necessidade de uma abordagem responsável e multidisciplinar. **Objetivo:** Investigar as implicações éticas da IA, com base em princípios como o da Responsabilidade (Jonas), e discutir a governança, a tomada de decisão automatizada e os impactos sociais. **Método:** Revisão bibliográfica sistemática de artigos científicos (2018–2023), com análise temática categorizada em: ética da responsabilidade, governança da IA, dilemas morais em sistemas autônomos e riscos sociais. Fontes incluem arXiv, IEEE Xplore e periódicos especializados. **Resultados:** Identificou-se uma lacuna entre princípios éticos teóricos e práticas industriais, além de resistência à delegação de decisões morais a máquinas (Salim *et al.*, 2019). Diretrizes como as da IEEE são pouco aplicadas, e a falta de diversidade nas equipes de desenvolvimento perpetua vieses (Kluge Corrêa *et al.*, 2022). A dualidade da IA-benefícios versus riscos-exige regulamentação robusta (Bazzan *et al.*, 2023). **Conclusões:** A ética na IA demanda frameworks multidisciplinares, supervisão humana em contextos críticos e diálogo contínuo entre academia, indústria e sociedade. Recomenda-se maior diversidade nas equipes, adoção prática de diretrizes e educação pública para mitigar riscos.

**Palavras-chave:** inteligência artificial, ética, governança, responsabilidade, tomada de decisão automatizada.

#### ABSTRACT

**Introduction:** The intersection between artificial intelligence (AI) and ethics has gained academic and social relevance, requiring reflections on the moral implications of technological development. This article examines the ethical challenges of AI, emphasizing the need for a responsible and multidisciplinary approach. **Objective:** To investigate the ethical implications of AI, based on principles such as Responsibility (Jonas), and discuss governance, automated decision-making, and societal impacts. **Method:** A systematic literature review of scientific articles (2018–2023), with thematic analysis categorized into: ethics of responsibility, AI governance, moral dilemmas in autonomous systems, and social risks. Sources include arXiv, IEEE Xplore, and specialized journals. **Results:** A gap was identified between theoretical ethical principles and industrial practices, along with resistance to delegating moral decisions to machines (Salim *et al.*, 2019). Guidelines such

as those from IEEE are poorly implemented, and a lack of diversity in development teams perpetuates biases (Kluge Corrêa *et al.*, 2022). The duality of AI-benefits versus risks-demands robust regulation (Bazzan *et al.*, 2023). Conclusions: AI ethics requires multidisciplinary frameworks, human oversight in critical contexts, and continuous dialogue among academia, industry, and society. Recommendations include greater diversity in teams, practical adoption of guidelines, and public education to mitigate risks.

**Keywords:** artificial intelligence, Ethics, governance, responsibility, automated decision-making.

## RESUMEN

**Introducción:** La intersección entre inteligencia artificial (IA) y ética ha ganado relevancia académica y social, exigiendo reflexiones sobre los impactos morales del desarrollo tecnológico. Este artículo analiza los desafíos éticos de la IA, destacando la necesidad de un enfoque responsable y multidisciplinario. **Objetivo:** Investigar las implicaciones éticas de la IA, con base en principios como el de Responsabilidad (Jonas), y discutir la gobernanza, la toma de decisiones automatizada y los impactos sociales. **Método:** Revisión bibliográfica sistemática de artículos científicos (2018–2023), con análisis temático categorizado en: ética de la responsabilidad, gobernanza de la IA, dilemas morales en sistemas autónomos y riesgos sociales. Fuentes incluyen arXiv, IEEE Xplore y revistas especializadas. **Resultados:** Se identificó una brecha entre principios éticos teóricos y prácticas industriales, además de resistencia a delegar decisiones morales a máquinas (Salim *et al.*, 2019). Directrices como las de IEEE son poco aplicadas, y la falta de diversidad en equipos de desarrollo perpetúa sesgos (Kluge Corrêa *et al.*, 2022). La dualidad de la IA-beneficios versus riesgos-exige regulación robusta (Bazzan *et al.*, 2023). **Conclusiones:** La ética en IA requiere marcos multidisciplinarios, supervisión humana en contextos críticos y diálogo continuo entre academia, industria y sociedad. Se recomienda mayor diversidad en equipos, adopción práctica de directrices y educación pública para mitigar riesgos.

**Palabras clave:** inteligencia artificial, Ética, gobernanza, responsabilidad, toma de decisiones automatizada.

## 1 INTRODUÇÃO

A interseção entre inteligência artificial (IA) e ética tem se tornado uma área de crescente relevância na academia e na sociedade em geral. A reflexão sobre as implicações éticas da IA é essencial para enfrentar os desafios que surgem com seu desenvolvimento e aplicação. Em "Inteligencia artificial y ética de la responsabilidad", (Luis Terrones Rodriguez, 2018) destaca a importância de uma abordagem ética, fundamentada no Princípio de Responsabilidade de Hans Jonas, para lidar com as transformações provocadas pela IA em várias esferas da vida. O autor enfatiza que, embora a IA represente um avanço significativo, é crucial não perder de vista os riscos associados e a responsabilidade que acompanha seu uso.

Em um contexto mais amplo, discutem a evolução das preocupações éticas na

pesquisa em IA, desde os primeiros conceitos propostos por Isaac Asimov até a contemporaneidade. O artigo revela que, apesar de um otimismo inicial, a comunidade científica enfrentou um período de pessimismo, que atualmente está sendo desafiado pelos avanços em aprendizado de máquina. A necessidade de garantir a segurança e a ética no desenvolvimento da IA se torna cada vez mais evidente, à medida que sua presença se intensifica na vida cotidiana.

(Maria Neves Rombão, 2019), em "Terão os robôs capacidade para substituir o ser humano?", explora a dualidade da IA reconhecendo tanto seu potencial para resolver problemas quanto as preocupações éticas que surgem com sua adoção. O autor menciona a necessidade de desenvolver sistemas inteligentes que sejam compatíveis com os valores humanos, levantando questões sobre a ética das máquinas e a possibilidade de direitos para sistemas de IA, especialmente quando suas ações podem causar danos.

(R. B. Salim *et al.*, 2019), no artigo "Até que Ponto Sistemas Inteligentes Podem Tomar Decisões Sem Ferir a Ética", investigam a aceitação da tomada de decisões por máquinas e os limites éticos que devem ser considerados. Os resultados revelam uma divisão de opiniões entre os entrevistados, destacando a desconfiança em relação à capacidade das máquinas de realizar escolhas éticas. Este estudo reforça a ideia de que a ética é um conceito dinâmico, que deve ser constantemente reavaliado à luz das inovações tecnológicas.

A pesquisa de (Vakkuri *et al.*, 2019) em "AI Ethics in Industry: A Research Framework" apresenta um quadro para a implementação de ética em ambientes industriais de IA. Os autores argumentam que a natureza dos sistemas de IA exige uma abordagem multidisciplinar e que há uma lacuna significativa entre a pesquisa acadêmica e a prática real em ética de IA. Essa análise é crucial para assegurar que as práticas de desenvolvimento de IA sejam éticas e responsáveis.

(Kluge Corrêa *et al.*, 2022), em "Worldwide AI Ethics: a review of 200 guidelines and recommendations for AI governance", abordam a necessidade de uma regulamentação mais forte na indústria de IA destacando a falta de diversidade e representação em discussões sobre governança da IA. O artigo sugere que uma maior atenção a esses aspectos é fundamental para moldar um futuro ético e responsável para a IA.

Finalmente, (L. C. Bazzan *et al.*, 2023), em "A Nova Eletricidade: Aplicações, Riscos e Tendências da IA Moderna", discutem os riscos e impactos da IA na sociedade contemporânea, enfatizando a necessidade de um debate informado e especializado. A crescente influência da IA em diversas áreas, incluindo o mercado de trabalho, exige uma análise crítica e fundamentada para mitigar os riscos associados à sua adoção.

Esses artigos, juntos, oferecem uma visão abrangente sobre a interface entre IA e

ética, sublinhando a importância de uma abordagem consciente e crítica no desenvolvimento e aplicação de tecnologias de IA. A análise das implicações éticas não é apenas uma consideração acadêmica, mas uma necessidade urgente em um mundo cada vez mais moldado pela tecnologia.

## 2 REFERENCIAL TEÓRICO

O artigo "Inteligencia artificial y ética de la responsabilidad" de (Luis Terrones Rodriguez, 2018) aborda a interseção crítica entre a inteligência artificial (IA) e a ética, destacando os desafios que a humanidade enfrenta com o avanço tecnológico. O autor argumenta que, embora a IA represente um significativo progresso em diversos setores, sua implementação não deve ocorrer sem uma reflexão ética aprofundada, considerando as complexas implicações que surgem em diferentes esferas da vida humana.

Terrones Rodriguez enfatiza que a rápida evolução da IA traz consigo uma série de desafios laborais, econômicos, de segurança e políticos que exigem novas abordagens éticas. O autor propõe o Princípio de Responsabilidade de Hans Jonas como um referencial ético apropriado para navegar essas questões. Essa escolha é pertinente, pois o princípio destaca a necessidade de considerar as consequências das ações humanas, especialmente em um contexto onde a tecnologia pode ter impactos profundos e, por vezes, imprevisíveis.

Um dos pontos críticos levantados no artigo é a competição entre a inteligência humana e a artificial, que se intensificará nas próximas décadas. O autor sugere que a sociedade ainda não comprehende completamente as implicações dessa competição, o que pode levar a um uso irresponsável da tecnologia. Essa falta de compreensão é alarmante, pois pode resultar em decisões que priorizam a eficiência tecnológica em detrimento de valores humanos essenciais.

Além disso, o autor alerta para o perigo do tecnocentrismo, onde a confiança excessiva na tecnologia pode obscurecer a necessidade de uma abordagem ética fundamentada. A cautela é apresentada como uma virtude necessária, enfatizando que a confiança não deve ser cega, mas informada e crítica. Essa perspectiva é crucial, pois a ética deve guiar o desenvolvimento e a aplicação da IA, garantindo que as inovações tecnológicas sirvam ao bem-estar humano e não se tornem um fim em si mesmas.

O artigo intitulado "On Quantifying and Understanding the Role of Ethics in AI Research: A Historical Account of Flagship Conferences and Journals", escrito por (Prates *et al.*, 2018), oferece uma análise abrangente sobre a intersecção entre a inteligência artificial (IA) e a ética, explorando como as preocupações éticas têm evoluído dentro da

comunidade de pesquisa em IA ao longo das décadas.

A discussão começa com uma referência ao trabalho seminal de Alan Turing, que levantou questões filosóficas sobre a possibilidade de construir máquinas com inteligência humana. O teste de Turing, que se tornou um marco na avaliação da inteligência das máquinas, também trouxe à tona preocupações éticas sobre o uso e as implicações da IA. Os autores mencionam as famosas Três Leis da Robótica de Isaac Asimov, que visavam prevenir o surgimento de máquinas maliciosas, destacando a necessidade de uma estrutura ética desde os primórdios da IA.

O artigo traça um panorama histórico em que, após um período inicial de otimismo, as preocupações éticas com a IA diminuíram, tornando-se mais restritas a escritores de ficção científica e filósofos. No entanto, com os avanços significativos em aprendizado de máquina a partir de 2010, essas preocupações ressurgiram, especialmente em um contexto onde a IA começou a permeiar aspectos cotidianos da vida humana. Essa nova realidade levanta questões cruciais sobre liberdade, privacidade e a confiança que depositamos nos algoritmos que nos cercam.

Os autores também abordam o impacto da automação no mercado de trabalho, referindo-se a Moshe Vardi, que sugere que a automação está eliminando empregos tradicionais sem criar novas oportunidades suficientes. A defesa de Elon Musk em favor de uma renda básica universal como uma solução para a distribuição de riqueza gerada por máquinas inteligentes é outro ponto relevante discutido, enfatizando o dilema ético da desigualdade que pode surgir com o avanço da IA.

Além disso, o artigo menciona os quatro princípios identificados pela IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems: benefício humano, responsabilidade, transparência e educação e conscientização. Esses princípios são fundamentais para garantir que a pesquisa em IA não apenas avance tecnologicamente, mas também considere as implicações éticas de suas aplicações.

O artigo "Terão os robôs capacidade para substituir o ser humano?" de (Maria Neves Rombão, 2019) apresenta uma análise crítica sobre o impacto da Inteligência Artificial (IA) nas interações humanas e nas práticas psicológicas. A autora discute como as ferramentas de IA estão sendo desenvolvidas para aprimorar a compreensão do processamento da informação humana, destacando a relevância dessas tecnologias para a psicologia, que busca entender os sentimentos e comportamentos humanos.

Um ponto central do artigo é a dualidade da IA: enquanto a autora reconhece suas potencialidades na resolução de problemas reais e na criação de soluções inteligentes, também levanta questões éticas e morais que emergem com o seu uso. A autora enfatiza que

a adoção da IA não é consensual, refletindo uma tensão entre os benefícios da tecnologia e as preocupações relacionadas à segurança e à ética. Essa ambivalência é crucial, pois sugere que, apesar dos avanços, há um risco associado à implementação de sistemas inteligentes que podem não estar alinhados com os valores humanos.

Rombão menciona que há esforços de pesquisa focados em desenvolver sistemas inteligentes que sejam compatíveis com os humanos, o que é um passo importante para mitigar os riscos associados à IA. Contudo, a autora também sugere que a ética deve ser uma consideração primordial no desenvolvimento dessas tecnologias. A falta de um consenso claro sobre as diretrizes éticas para a IA levanta questões sobre a responsabilidade dos desenvolvedores e a necessidade de um diálogo contínuo entre cientistas, psicólogos e especialistas em ética.

O artigo "Até que Ponto Sistemas Inteligentes Podem Tomar Decisões Sem Ferir a Ética", escrito por (R. B. Salim *et al.*, 2019), apresenta uma análise crítica sobre a intersecção entre a inteligência artificial (IA) e a ética, evidenciando as complexidades envolvidas na tomada de decisões por máquinas. A pesquisa revela que, apesar do crescente desenvolvimento da IA e da aceitação por parte da sociedade, existe uma resistência significativa em delegar a responsabilidade de decisões éticas a sistemas autônomos.

Os autores destacam que, entre os entrevistados que preferem salvar a vida dos pedestres, uma porcentagem considerável expressou desconforto com a ideia de que um carro autônomo poderia tomar essa decisão. Essa resistência sugere que a confiança nas máquinas ainda é limitada, especialmente em situações que envolvem dilemas morais complexos. A necessidade de uma decisão humana em contextos éticos é um ponto crucial que o artigo explora, indicando que a moralidade é uma construção social que não pode ser facilmente codificada em algoritmos.

Outro aspecto relevante abordado pelos autores é o potencial viés presente nos dados utilizados para treinar sistemas de IA. Eles argumentam que, uma vez que a IA é uma representação das visões de mundo de seus criadores, ela pode reproduzir preconceitos e erros humanos. Essa questão levanta preocupações sobre a responsabilidade ética dos desenvolvedores de IA e a necessidade de uma supervisão cuidadosa na curadoria de dados.

Além disso, o artigo menciona que a compreensão do que é ético pode variar significativamente entre diferentes culturas e ao longo do tempo, o que complica ainda mais a implementação de sistemas de IA em contextos globais. Essa variabilidade sugere que não existe uma resposta única ou universal para as questões éticas que surgem com a adoção da IA.

O artigo "AI Ethics in Industry: A Research Framework" de (Vakkuri *et al.*, 2019)

aborda a crescente influência dos sistemas de Inteligência Artificial (IA) na sociedade e a necessidade de incorporar a ética no desenvolvimento desses sistemas. Os autores destacam a falta de estruturas práticas para a implementação da ética em IA, o que é um aspecto crítico, considerando que, atualmente, a discussão acadêmica e pública sobre ética em IA está em ascensão, mas ainda carece de aplicação prática.

Um dos pontos centrais do artigo é a distinção entre usuários de sistemas de IA, que frequentemente são organizações, e os indivíduos, que são tratados como objetos por esses sistemas. Essa dinâmica levanta questões éticas importantes sobre a autonomia e a privacidade dos indivíduos em um contexto onde a IA é utilizada principalmente para coleta de dados e execução de ações ((Vakkuri *et al.*, 2019)). A abordagem proposta pelos autores é um framework de pesquisa que visa facilitar a implementação da ética em ambientes industriais, reconhecendo que o desenvolvimento de IA e Sistemas Autônomos (AS) é uma tarefa multidisciplinar.

Os autores também observam que, apesar do crescente interesse em ética em IA, existe uma lacuna significativa entre a pesquisa e a prática. As diretrizes do IEEE Ethically Aligned Design, por exemplo, não foram amplamente adotadas pelos profissionais do setor. Essa desconexão sugere que, embora haja um reconhecimento da importância da ética, a sua integração nos processos de desenvolvimento real ainda enfrenta barreiras significativas ((Vakkuri *et al.*, 2019)).

O framework apresentado no artigo é baseado em pesquisas conceituais existentes sobre ética em IA e busca servir como um ponto de partida para estudos empíricos nessa área. Os autores argumentam que a discussão acadêmica até o momento tem se concentrado na definição de princípios centrais, o que levou a uma resposta de governos e organizações de padronização que criaram suas próprias diretrizes. No entanto, a falta de uma aplicação prática e consistente dessas diretrizes ainda representa um desafio a ser superado.

O artigo "Worldwide AI Ethics: a review of 200 guidelines and recommendations for AI governance" de (Kluge Corrêa *et al.*, 2022) apresenta uma análise abrangente das diretrizes e recomendações existentes para a governança da inteligência artificial (IA). A pesquisa destaca a necessidade urgente de uma mudança na cultura de desenvolvimento da IA, enfatizando que a indústria atual, em grande parte desregulada, requer uma intervenção governamental mais robusta. Essa necessidade é impulsionada pela crescente conscientização dos desenvolvedores sobre os riscos e falhas associados às tecnologias de IA, que, segundo os autores, ainda são amplamente subestimados.

Os autores argumentam que a diversidade nas equipes de desenvolvimento de IA é crucial para mitigar preconceitos e falhas éticas. No entanto, eles reconhecem que a

representação de gênero e LGBTQIA+ continua deficiente, o que pode perpetuar problemas éticos na criação de sistemas de IA. Essa lacuna na diversidade é uma questão central que deve ser abordada para garantir que as tecnologias desenvolvidas sejam inclusivas e equitativas.

Além disso, o artigo sugere que a governança da IA é um campo ainda novo e em aberto, o que implica que há uma necessidade premente de mais pesquisa e exploração. Os autores oferecem ferramentas e recomendações que podem ser úteis para futuros pesquisadores e desenvolvedores que buscam navegar por esse campo complexo. A análise de 200 diretrizes fornece uma base sólida para entender as diferentes abordagens em curso e as lacunas que ainda precisam ser preenchidas.

Em termos de avaliação crítica, o trabalho de Corrêa *et al.* é valioso por sua abrangência e pela clareza na identificação dos desafios atuais da governança da IA. No entanto, poderia ter se aprofundado mais nas implicações práticas das diretrizes estudadas e como elas poderiam ser implementadas de forma eficaz. Além disso, uma discussão mais detalhada sobre as consequências da falta de diversidade nas equipes de desenvolvimento poderia enriquecer ainda mais a análise.

O artigo "A Nova Eletricidade: Aplicações, Riscos e Tendências da IA Moderna" de (L. C. Bazzan *et al.*, 2023) aborda a crescente importância da Inteligência Artificial (IA) na indústria de jogos e discute os impactos e riscos associados à sua aplicação na economia e na sociedade. A análise crítica apresentada pelos autores revela a complexidade do debate em torno da IA, especialmente após o surgimento de modelos de linguagem como o ChatGPT, que catalisaram uma discussão mais ampla sobre as implicações éticas e sociais da tecnologia.

Os autores destacam que a IA se tornou uma ferramenta indispensável, mas também levantam preocupações significativas sobre os riscos que ela representa. Um ponto central do artigo é a influência da IA na formação da opinião pública, que pode ser distorcida por informações provenientes de indivíduos sem conhecimento técnico adequado. Essa dinâmica pode resultar em desinformação e, consequentemente, em uma percepção errônea dos riscos e benefícios da IA ((L. C. Bazzan *et al.*, 2023)). Essa crítica é especialmente relevante em um contexto onde as decisões políticas e sociais são frequentemente moldadas por opiniões populares, que podem não estar fundamentadas em dados ou em uma compreensão profunda da tecnologia.

Além disso, o artigo aponta que governos de nações líderes em IA estão formando grupos de trabalho para investigar e mitigar esses riscos. Essa iniciativa é um passo positivo, pois sugere uma abordagem proativa na regulamentação e na governança da IA. No entanto,

a necessidade de consultar especialistas na área é enfatizada, o que levanta a questão sobre a acessibilidade do conhecimento técnico e a importância de uma educação adequada para o público em geral ((L. C. Bazzan *et al.*, 2023)).

Os autores também mencionam que a discussão sobre IA deve ser inclusiva, engajando tanto especialistas quanto leigos, mas com um foco na educação e na informação precisa. A crítica à falta de experiência de muitos comentaristas sobre o desenvolvimento de tecnologias de IA é um chamado para que a sociedade busque um entendimento mais profundo sobre as implicações dessa tecnologia.

### 3 METODOLOGIA

Este artigo adotou uma abordagem qualitativa e exploratória, baseada em revisão bibliográfica sistemática e análise crítica de artigos científicos, livros e documentos relevantes sobre inteligência artificial (IA) e ética. A metodologia consistiu em três etapas principais: a) Seleção e Coleta de Fontes onde foram selecionados trabalhos acadêmicos publicados entre 2018 e 2023, obtidos de bases de dados como arXiv, IEEE Xplore, periódicos científicos e repositórios institucionais. B) Análise Temática, sendo os materiais categorizados em eixos temáticos, como: ética da responsabilidade, governança da IA, tomada de decisão automatizada e impactos sociais. C) Síntese e Interpretação com os resultados interpretados à luz dos objetivos do estudo, relacionando as discussões teóricas com desafios práticos no desenvolvimento e aplicação da IA.

### 4 RESULTADOS E DISCUSSÕES

Adiante, tem-se a discussão acerca dos achados do estudo, interpretando os dados obtidos à luz dos objetivos e das hipóteses de pesquisa, e relacionando-os com a literatura existente.

### 5 CONCLUSÃO

A análise da interseção entre inteligência artificial (IA) e ética revela a necessidade urgente de uma abordagem crítica e informada no desenvolvimento e aplicação dessas tecnologias. O artigo de (Luis Terrones Rodriguez, 2018) destaca o Princípio de Responsabilidade de Hans Jonas como um referencial ético essencial para lidar com os desafios emergentes da IA, enfatizando que a implementação dessa tecnologia deve ser

acompanhada de uma reflexão ética profunda para evitar riscos significativos. Essa perspectiva é corroborada por , que traça a evolução das preocupações éticas na pesquisa em IA, ressaltando a necessidade de considerar as implicações da automação e a ética nas decisões que envolvem a vida humana.

Além disso, a discussão sobre a dualidade da IA, abordada por (Maria Neves Rombão, 2019), revela que, embora a IA tenha potencial para resolver problemas, também levanta questões éticas que não podem ser ignoradas. A resistência em delegar decisões éticas a máquinas, como observado por (R. B. Salim *et al.*, 2019), indica uma desconfiança generalizada em relação à capacidade das máquinas de realizar escolhas morais, o que reforça a necessidade de uma supervisão humana em contextos críticos.

A pesquisa de (Vakkuri *et al.*, 2019) apresenta um quadro para a implementação de ética em ambientes industriais, destacando a lacuna entre a teoria e a prática na aplicação das diretrizes éticas em IA. A análise de (Kluge Corrêa *et al.*, 2022) sobre a governança da IA também sublinha a necessidade de uma regulamentação mais robusta e inclusiva, enfatizando que a diversidade nas equipes de desenvolvimento é crucial para mitigar preconceitos e falhas éticas.

Por fim, (L. C. Bazzan *et al.*, 2023) discute os riscos associados à IA na sociedade contemporânea, ressaltando a importância de um debate informado e especializado, que envolva tanto especialistas quanto o público em geral. A crescente influência da IA em diversas áreas, especialmente na formação da opinião pública, exige uma análise crítica e fundamentada para mitigar os riscos associados à sua adoção.

Em conclusão, a interface entre IA e ética demanda uma abordagem multidisciplinar que considere as implicações éticas em todas as etapas do desenvolvimento tecnológico. A responsabilidade dos desenvolvedores e a necessidade de um diálogo contínuo entre diferentes áreas são fundamentais para garantir que a IA sirva ao bem-estar humano e não se torne um fim em si mesma.

## REFERÊNCIAS

Terrones Rodríguez, A. L. (2018). Inteligencia artificial y ética de la responsabilidad. *Cuestiones de Filosofía*, 4(22), 141–170.  
<https://doi.org/10.19053/01235095.v4.n22.2018.8311>

Rombão, M.M.N. *Terão os robôs capacidade para substituir o ser humano?..* [Dissertação de Mestrado, Laureate International Universities].  
<https://core.ac.uk/download/223223799.pdf>

R. B. Salim, S., B. Lacerda, V., C. G. Maia, L., & F. Villela, H. (2019). Até que Ponto Sistemas Inteligentes Podem Tomar Decisões Sem Ferir a Ética.  
<https://revista.fumec.br/index.php/computacaoesociedade/article/view/7311>

Vakkuri, V., Kemell, K. K., & Abrahamsson, P. (2019). AI Ethics in Industry: A Research Framework. <https://arxiv.org/pdf/1910.12695>

Kluge Corrêa, N., Galvão, C., William Santos, J., Del Pino, C., Pontes Pinto, E., Barbosa, C., Massmann, D., Mambrini, R., Galvão, L., Terem, E., & de Oliveira, N. (2022). Worldwide AI Ethics: a review of 200 guidelines and recommendations for AI governance. <https://www.cell.com/action/showPdf?pii=S2666-3899%2823%2900241-6>

L. C. Bazzan, A., R. Tavares, A., G. Pereira, A., R. Jung, C., Scharcanski, J., Luis Carbonera, J., C. Lamb, L., Recamonde-Mendoza, M., L. T. da Silveira, T., & Moreira, V. (2023). A Nova Eletricidade: Aplicac{c}~oes, Riscos e Tend^encias da IA Moderna - The New Electricity: Applications, Risks, and Trends in Current AI.  
<https://arxiv.org/pdf/2310.18324>

Prates, M., Avelar, P., & C. Lamb, L. (2018). On Quantifying and Understanding the Role of Ethics in AI Research: A Historical Account of Flagship Conferences and Journals.  
<https://arxiv.org/abs/1809.08328>