

ARAÚJO, Marcos de Souza. Floridi, L. *A ética da inteligência artificial: princípios, desafios e oportunidades*. Curitiba: pucpress, 2024. 396. *Revista Intercâmbio*, v. LVII, e72899, 2025. São Paulo: LAEL/PUCSP. ISSN 2237-759X

<https://doi.org/10.23925/2237-759X2025V57e72899>

## Resenha

FLORIDI, L. *A ética da inteligência artificial: princípios, desafios e oportunidades*. Curitiba: PUCPRESS, 2024. 396 p.

Marcus de Souza Araújo  
(Universidade Federal do Pará / Universidade Federal de Minas Gerais)  
marcusaraujo@ufpa.br

*O imenso sucesso pragmático que a IA tem alcançado, junto evidentemente de sua falibilidade que não é nenhuma novidade em quaisquer áreas de conhecimento, em especial, nesse caso, as externalidades negativas ou vieses de muitos de seus resultados, tudo isso tem chamado a atenção dos críticos da ciência e da sociedade, expressa em uma profusão de textos voltados para as questões éticas e de proteção dos direitos humanos. (Santaella, 2023, p. 18).*

A sociedade contemporânea mudou e está mudando em ritmo acelerado sob a influência e o desenvolvimento da inteligência artificial generativa (IAGen), principalmente, com o lançamento do ChatGPT 3.5, em 30 de novembro de 2022. Esse agente conversacional atingiu 1 milhão de usuários em cinco dias, após ser lançado, e chegou a 100 milhões de usuários em apenas dois meses (Bentley, 2025). Nessa linha de pensamento, Laurent Alexandre (2024, p. 19), em *A guerra das inteligências na era do ChatGPT*, reflete que o ChatGPT marca uma nova etapa na produção de redes de neurônios artificiais, razão pela qual a "IA não é um detalhe da História. É o futuro da humanidade que ocorre nas suas linhas de código."

A adoção de tecnologias baseadas em IAGen para o suporte a atividades cotidianas tem se intensificado de maneira acelerada em diversas áreas do conhecimento, evidenciando seu potencial para transformar as formas de interação, solucionar problemas, fundamentar decisões e "elaborar perguntas e manter diálogos com a máquina. Portanto, quanto melhor for a habilidade de comunicação do profissional, maior tenderá a ser sua articulação no ambiente digital inteligente que emerge" (Gabriel, 2025, p. 115). Contudo, à medida que a IAGen se consolida em múltiplos contextos sociais, acadêmicos e profissionais, torna-se imprescindível uma reflexão crítica acerca dos

impactos que podem emergir dessa integração. Tal preocupação é justificada pelas evidências que apontam para os riscos relacionados ao uso inadequado desta tecnologia, seja por falhas involuntárias ou por sua aplicação com propósitos maliciosos.

Por essa razão, torna-se relevante que as pessoas estejam preparadas para saber usar a IAGen de maneira crítica, consciente, ética e responsável, como bem ressalta Santaella (2023) na epígrafe desta resenha. É preciso conhecer os princípios, as leis e as condutas que orientam o comportamento humano em uma sociedade. Nessa direção, “ser ético significa agir de maneira consistente com as virtudes que uma pessoa virtuosa demonstraria em uma situação específica.” (Córdova, 2025, p. 194). Daí, é essencial que os conteúdos criados pela IAGen operem de maneira ética respeitando os direitos dos indivíduos no século XXI.

A obra *A ética da inteligência artificial: princípios, desafios e oportunidades* apresenta e discute de maneira filosófica a relação premente da ética da inteligência artificial na era das tecnologias, abordando desde os princípios norteadores até os desafios enfrentados, e as boas práticas para seu desenvolvimento e sua implementação. O livro se destaca ao examinar os riscos da IA, com seu uso indevido, além de caracterizá-la como uma nova forma de agência para o benefício da humanidade e do meio ambiente.

A obra em tela, originalmente publicada em 2023, em língua inglesa, pela *Oxford University Press*, foi traduzida em língua portuguesa, em 2024, pela Associação Paranaense de Cultura (APC) e publicada pela editora PUCPRESS no mesmo ano. O livro é de autoria de Luciano Floridi, Diretor Fundador do *Yale Center for Digital Ethics*, professor de Ciência Cognitiva na Universidade de Yale e de Sociologia da Cultura e Comunicação na Universidade de Bolonha, Itália. Suas pesquisas estão relacionadas à filosofia da informação, à ética digital, à ética da IA e à filosofia da tecnologia.

Luciano Floridi é considerado uma das maiores autoridades da filosofia contemporânea e seu livro reforça que os princípios éticos estão presentes em cada setor da sociedade contemporânea, razão pela qual a compreensão desses princípios pode garantir que o desenvolvimento e uso consciente da IAGen possam beneficiar a todos de maneira justa e segura. O livro propõe um olhar crítico para que a IAGen possa beneficiar a todos, promovendo, assim, um desenvolvimento sustentável e inclusivo.

A obra em tela está estruturada em duas partes. A primeira, intitulada *Entendendo a IA*, apresenta uma introdução filosófica e concisa aos aspectos históricos, contemporâneos e prospectivos da inteligência artificial, distribuída em três capítulos. A segunda parte,

denominada *Avaliação da IA*, contempla temáticas específicas relacionadas à ética da inteligência artificial, sendo composta por dez capítulos. Nessa linha de pensamento, o conceito de ética apresentado por Floridi (2024) converge com o apresentado por Lyra e Chevitarese (2024, p. 25), a saber: “[...] como uma atividade ou disciplina filosófica que procura estabelecer nacionalmente critérios e princípios para a conduta humana, propondo-os, sempre, com pretensão de universalidade.”

A Parte 1, intitulada *Entendendo a IA*, é composta por três capítulos que delineiam a estrutura conceitual do tema. O Capítulo 1, *Passado: a ascensão da IA*, tem como objetivo apresentar, sob uma perspectiva conceitual, o processo de ascensão da inteligência artificial, destacando as transformações que contribuíram para seu desenvolvimento como uma nova forma de agência eficaz, ainda que destituída de inteligência, no sentido humano. Em outras palavras, agência no sentido, por exemplo, de “receber e usar dados do ambiente, por meio de sensores ou outras formas de entrada de dados” (Floridi, 2024, p. 49). Importa ressaltar que o foco do capítulo não reside em oferecer um panorama histórico ou tecnológico da IA, mas sim em explorar os fundamentos teóricos que sustentam sua evolução.

O Capítulo 2, *Presente: a IA como uma nova forma de agência, não de inteligência*, refere-se a uma interpretação da inteligência artificial contemporânea como um reservatório de agência, sustentado por dois fatores principais: (a) a capacidade de resolver problemas e executar tarefas orientadas a metas; e (b) a crescente necessidade de que essas ações sejam realizadas com algum grau de inteligência, em consonância com a transformação gradual dos ambientes humanos em uma infosfera favorável à atuação da IA. O autor do livro também destaca a inexistência de uma definição única e consensual para inteligência artificial, sendo esta compreendida como “um atalho útil para se referir a uma família de ciências, métodos, paradigmas, tecnologias, produtos e serviços” (Floridi, 2024, p. 55).

Concluindo a Parte 1, o Capítulo 3, *Futuro: o desenvolvimento previsível da IA*, concentra-se nas perspectivas futuras do desenvolvimento da inteligência artificial, com ênfase nos tipos de problemas que tendem a ser mais prontamente solucionados por essa tecnologia. O autor ressalta, ainda, o papel central dos modelos de linguagem em grande escala (em inglês: *Large Language Model* ou *LLM*), compreendidos como interfaces fundamentais entre diferentes agentes humanos e artificiais. Além disso, o capítulo em tela propõe uma reflexão sobre os denominados “invernos da IA”, destacando as lições já aprendidas — e aquelas ainda por aprender — com tais

períodos de estagnação, a fim de melhor explorar as potencialidades singulares dessa tecnologia.

Peter J. Bentley (2025), em *A história da inteligência artificial para quem tem pressa: o nascimento de uma nova era em 200 páginas!*, esclarece que “Invernos de IA” (em inglês, *AI winters*) são períodos na história da inteligência artificial em que houve uma forte redução no interesse, no financiamento e no desenvolvimento de pesquisas na área. O nome faz uma analogia com o inverno climático, um tempo de frio e estagnação, simbolizando momentos de desânimo e paralisação no campo da IA.

De acordo ainda com Bentley (2025), esse “inverno” está subdividido em dois períodos, quais sejam: o primeiro inverno da inteligência artificial foi na década de 1970, cujas pesquisas iniciais sobre IA decepcionaram por não conseguir entregar resultados práticos. Nessa direção, os governos cortaram financiamentos das pesquisas. O segundo inverno ocorreu entre os anos de 1990 a 2000. Nesse período os programas que simulavam decisões humanas (sistemas especialistas) também não cumpriram as promessas e a “IA foi ficando completamente fora de moda” (Bentley, 2025, p. 97) mais uma vez.

A Parte 2 da obra, *Avaliação da IA*, é composta por dez capítulos que exploram, de maneira aprofundada, temas específicos relacionados à ética da inteligência artificial. O Capítulo 4, *Uma estrutura unificada de princípios éticos para a IA*, inaugura essa segunda parte ao apresentar uma proposta de estrutura ética unificada, fundamentada em uma análise comparativa das principais diretrizes publicadas desde 2017, ano em que os princípios éticos voltados à IA passaram a ser divulgados com maior intensidade.

Dentre os documentos de maior relevância, destacam-se os *Princípios de Asilomar sobre a IA* e a *Declaração de Montreal para um Desenvolvimento Responsável da IA*. Segundo Floridi (2024), essa estrutura ética é composta por cinco princípios fundamentais, dos quais quatro são amplamente reconhecidos no campo da bioética, a saber: beneficência, não maleficência, autonomia e justiça. O quinto princípio, denominado *explicabilidade*, é considerado essencial, uma vez que a IA representa uma nova forma de agência. O capítulo conclui com uma reflexão crítica acerca das implicações dessa estrutura para o desenvolvimento futuro de legislações, normativas técnicas, padrões e boas práticas voltadas à promoção de uma IA ética em diferentes contextos sociais e institucionais.

O Capítulo 5, *Dos princípios às práticas: os riscos de ser antiético*, discute cinco riscos que comprometem a transposição dos princípios éticos para práticas efetivas, tanto em contextos reais quanto digitais. São eles: *shopping* da ética, *bluwashing* da ética, *lobbying* da ética,

*dumping* da ética e *shirking* da ética. O autor ressalta que tais riscos não são exclusivos do campo da inteligência artificial, pois também se manifestam em outras áreas nas quais a ética é aplicada, como a ética ambiental, a bioética, a ética médica e a ética empresarial. Além disso, o capítulo destaca que uma abordagem ética no desenvolvimento e uso da IA implica, igualmente, reconhecer a natureza e a incidência desses riscos, bem como promover ações preventivas voltadas à sua mitigação.

O Capítulo 6, *Ética branda e governança de IA*, propõe uma reflexão sobre a governança ética da inteligência artificial e, de modo mais amplo, das tecnologias digitais, destacando-a como um dos novos desafios decorrentes do avanço tecnológico. Nesse contexto, é apresentado o conceito de ética branda, compreendido como a aplicação de princípios éticos em um cenário pós-conformidade, isto é, em consonância com os marcos legais vigentes. O capítulo encerra-se com a análise das inter-relações entre a ética digital, a regulamentação digital e os mecanismos de governança digital, enfatizando a importância de uma abordagem ética que vá além do cumprimento normativo.

O Capítulo 7, *Mapear a ética dos algoritmos*, examina os desafios éticos concretos suscitados pela inteligência artificial, com ênfase no debate acerca da ética dos algoritmos. A abordagem proposta busca conciliar inclusão — ao considerar todos os tipos de algoritmos, e não apenas aqueles que operam sistemas ou aplicações baseadas em IA — com especificidade analítica. O capítulo também desenvolve uma reflexão sobre as implicações éticas associadas ao uso de algoritmos, destacando questões epistêmicas e normativas. O autor enfatiza a relevância das análises éticas como estratégia para mitigar riscos, ao mesmo tempo em que se exploram os potenciais benefícios oferecidos pelas tecnologias digitais. Tais análises cumprem a função de elucidar tanto a natureza dos riscos envolvidos quanto as possibilidades construtivas derivadas do uso ético de algoritmos.

O Capítulo 8, *Práticas ruins: o uso indevido da IA para o mal social*, examina os impactos negativos e as ameaças previsíveis associadas à inteligência artificial. Conforme exposto no capítulo, uma consequência não intencional do recente aumento considerável das pesquisas em IA é a reorientação potencial dessas tecnologias para a facilitação de atos criminosos, denominados *crime de IA* (doravante CIA). Segundo Floridi (2024), o CIA constitui uma área relativamente nova e, por sua natureza interdisciplinar, pode abranger tanto estudos sociojurídicos quanto ciências formais; nesse sentido, permanece certa incerteza quanto ao desenvolvimento futuro desse campo. O capítulo também oferece uma síntese dos problemas contemporâneos relacionados ao CIA e delinea possibilidades iniciais para a construção de um espaço de solução.

O Capítulo 9, *Boas práticas: o uso adequado da IA para o bem social*, dedica-se à análise da inteligência artificial voltada ao bem social (do inglês: *Artificial Intelligence for Social Good* – AI4SG), um campo que vem ganhando destaque tanto nas sociedades da informação quanto na comunidade científica especializada em IA, de acordo com a pesquisa do autor do livro em tela. Essa abordagem apresenta potencial significativo para enfrentar problemas sociais por meio da criação de soluções baseadas em tecnologias de inteligência artificial. Em outras palavras, um projeto de AI4SG “ajuda a reduzir, diminuir ou erradicar um determinado problema social ou ambiental sem introduzir novos danos, ou ampliar os já existentes” (Floridi, 2024, p. 228).

O capítulo também evidencia uma lacuna no entendimento sobre o tema, iniciando com a apresentação de uma definição de IA4SG e, em seguida, identificando sete fatores éticos fundamentais para orientar futuras iniciativas desse tipo. São eles: (1) falseabilidade e implementação incremental; (2) mecanismos de proteção contra a manipulação de preditores; (3) intervenção contextualizada pelo receptor; (4) explicação contextualizada pelo receptor e objetivos transparentes; (5) garantia de privacidade e consentimento do titular dos dados; (6) equidade situacional; e (7) semantização acessível ao ser humano. Com base nesses fatores, o capítulo propõe um conjunto de boas práticas voltadas à condução ética e eficaz de projetos de IA para o bem social.

O Capítulo 10, *Como criar uma sociedade de IA boa: algumas recomendações*, apresenta um conjunto de diretrizes construtivas e concretas voltadas à avaliação, ao desenvolvimento, ao incentivo e ao respaldo de uma inteligência artificial orientada para o bem. Em determinados contextos, tais recomendações podem ser implementadas diretamente por formuladores de políticas públicas em nível nacional ou supranacional. Em outros casos, as transformações propostas podem ser conduzidas por diversos atores sociais, incluindo organizações da sociedade civil, agentes do setor privado e entidades setoriais. O autor do livro ressalta que, uma vez adotadas, essas recomendações podem contribuir significativamente para a consolidação de uma base sólida que favoreça a construção de uma Sociedade de Inteligência Artificial Boa.

O Capítulo 11, *O gambito: impacto da IA nas mudanças climáticas*, aborda o princípio da beneficência no contexto da sustentabilidade planetária, examinando os impactos ambientais, tanto positivos quanto negativos, decorrentes do uso da inteligência artificial. O principal objetivo do capítulo consiste em discutir recomendações de políticas públicas que promovam um desenvolvimento da IA mais sustentável e ambientalmente responsável. Conforme argumenta o

autor, este é um momento decisivo, dado que a IA já está sendo empregada na modelagem de fenômenos climáticos e na contribuição para estratégias de mitigação do aquecimento global. Nesse sentido, a tecnologia pode representar uma força significativa em prol de uma sociedade mais equitativa e sustentável.

O Capítulo 12, *IA e os Objetivos de Desenvolvimento Sustentável da ONU*, examina os esforços existentes para compreender e fomentar iniciativas de Inteligência Artificial para o Bem Social (AI4SG), destacando, contudo, as limitações impostas pela escassez de análises normativas e de evidências empíricas. Nesse sentido, o capítulo argumenta em favor da utilização dos Objetivos de Desenvolvimento Sustentável (ODS) da ONU como parâmetro para monitorar o alcance e a disseminação das ações relacionadas à abordagem AI4SG. Ademais, o capítulo apresenta, com detalhes, um banco de dados contendo projetos vinculados a essa abordagem, oferecendo, assim, uma visão abrangente sobre sua implementação prática. Por fim, o capítulo enfatiza a necessidade de identificação de problemas críticos que, se negligenciados, podem comprometer a eficácia das iniciativas voltadas ao uso socialmente benéfico da inteligência artificial.

O Capítulo 13, *Conclusão: o verde e o azul*, apresenta considerações finais acerca da natureza da inteligência artificial e propõe uma síntese prospectiva sobre os rumos futuros dessa tecnologia. O autor sugere uma transição conceitual da ética centrada na agência da IA para uma abordagem orientada por políticas de ação social, com vistas à construção de um projeto humano fundamentado na integração entre o “verde” — representando os ambientes e *habitats* naturais — e o “azul” — símbolo das tecnologias digitais, em especial a IA.

Essa convergência é concebida como um caminho promissor para o desenvolvimento de uma sociedade mais equitativa e de uma biosfera mais saudável. O capítulo também enfatiza que toda forma de agência, inclusive a nova modalidade representada pela IA, requer mecanismos de governança. Ressalta-se, ainda, que a governança é inerentemente uma prática sociopolítica e que os meios digitais têm reconfigurado, significativamente, os modos de exercício da agência sociopolítica.

Nesse contexto, o livro resenhado destaca-se por sua abordagem crítica e propositiva, ao reconhecer as limitações estruturais das iniciativas atuais, sem desconsiderar o potencial transformador da IA quando orientada por princípios éticos e compromissos globais. A articulação entre tecnologia e desenvolvimento sustentável apresentada no texto reforça a urgência de uma governança responsável da IA, baseada em dados empíricos, marcos regulatórios e no diálogo entre múltiplos atores sociais. Trata-se, portanto, de uma contribuição

ARAÚJO, Marcos de Souza. Floridi, I. *A ética da inteligência artificial: princípios, desafios e oportunidades*. Curitiba: pucpress, 2024. 396. *Revista Intercâmbio*, v. LVII, e72899, 2025. São Paulo: LAEL/PUCSP. ISSN 2237-759X

relevante para o campo da ética da IA, ao propor caminhos concretos para integrar inovação e responsabilidade social.

Assim sendo, Floridi (2024) propõe uma integração simbiótica entre IA e natureza, baseada em valores éticos e em estruturas sociopolíticas responsáveis. O livro resenhado cumpre o papel de ampliar o horizonte do debate, ao destacar que o futuro da IA não deve ser guiado apenas por avanços técnicos, mas por decisões coletivas informadas, que considerem o bem-estar da humanidade e do planeta como um todo. Nessa direção, a obra em questão apresenta uma análise aprofundada sobre os impactos, os desafios e as possibilidades da inteligência artificial no contexto contemporâneo, propondo uma reflexão ética, política e ambiental sobre os caminhos que essa tecnologia pode seguir.

## Referências bibliográficas

ALEXANDRE, L. *A guerra das inteligências na era do ChatGPT*. Tradução de Idalina Lopes. Santana de Parnaíba, SP: Amarylis, 2024.

BENTLEY, P. J. *A história da inteligência artificial para quem tem pressa: o nascimento de uma nova era em 200 páginas!* Rio de Janeiro: Valentina, 2025.

CÓRDOVA, P. R. *Inteligência artificial: entre o fascínio e o medo*. São Paulo: Contexto, 2025.

GABRIEL, M. *A (r)evolução das habilidades para o futuro: do trabalho na era da inteligência artificial*. São Paulo: Editora SENAC São Paulo, 2025.

LYRA, E.; CHEVITARESE, L. *Ética: conceitos, fundamentos e aplicações contemporâneas*. Rio de Janeiro: Editora Vozes; Editora PUC-Rio, 2024.

SANTAELLA, L. *A inteligência artificial é inteligente?* São Paulo: Edições 70, 2023.

Recebido em: 09/08/2025

Aprovado em: 11/11/2025



Esta obra está licenciado com uma Licença [Creative Commons Atribuição 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/) que permite o uso irrestrito, distribuição e reprodução em qualquer meio, desde que a obra original seja devidamente citada