

# A INTELIGÊNCIA ARTIFICIAL DO PONTO DE VISTA FILOSÓFICO: UMA BREVE ANÁLISE

ARTIFICIAL INTELLIGENCE FROM A PHILOSOPHICAL POINT OF VIEW: A  
BRIEF ANALYSIS

Ciências Humanas • 07/06/2026

REGISTRO DOI: [10.70773/revistatopicos/780677556](https://doi.org/10.70773/revistatopicos/780677556)

Adelcio Machado dos Santos<sup>1</sup>

Cesar Henrique Wanke<sup>2</sup>

## RESUMO

Este artigo propõe uma reflexão filosófica sobre a inteligência artificial (IA), examinando suas implicações ontológicas, epistemológicas e éticas. A partir de uma abordagem interdisciplinar, discute-se como diferentes correntes filosóficas interpretam a natureza da IA e seus impactos na compreensão do ser humano. A filosofia da mente, a fenomenologia e a hermenêutica são mobilizadas para distinguir entre a inteligência computacional e a consciência humana, ressaltando a ausência de subjetividade e intencionalidade nos sistemas artificiais. Ao longo do trabalho, são explorados contributos de pensadores como John Searle, Hubert Dreyfus, Martin Heidegger e Paul Ricoeur, que argumentam sobre os limites da IA em simular aspectos existenciais e cognitivos próprios do homem. A análise aponta que, embora a IA seja capaz de realizar tarefas complexas, sua atuação permanece restrita ao plano técnico-operacional, não alcançando as esferas da experiência, da historicidade e do sentido existencial. Ao final, destaca-se a importância de uma abordagem filosófica crítica que oriente o desenvolvimento e a aplicação ética dessas tecnologias. Sugere-se, ainda, que pesquisas futuras aprofundem o diálogo entre filosofia, tecnologia e ciências cognitivas, com vistas a repensar os fundamentos do humano em tempos de transformações digitais.

**Palavras-chave:** inteligência artificial; filosofia; consciência; ética; ontologia.

## ABSTRACT

This article presents a philosophical reflection on artificial intelligence (AI), examining its ontological, epistemological, and ethical implications. Through an interdisciplinary approach, it discusses how different philosophical schools interpret the nature of AI and its impact on human understanding. Philosophy of mind,

phenomenology, and hermeneutics are employed to distinguish computational intelligence from human consciousness, highlighting the absence of subjectivity and intentionality in artificial systems. Throughout the paper, contributions from authors such as John Searle, Hubert Dreyfus, Martin Heidegger, and Paul Ricoeur are explored, emphasizing the limitations of AI in simulating existential and cognitive aspects inherent to humans. The analysis indicates that although AI can perform complex tasks, its action remains restricted to the technical-operational domain, failing to reach the spheres of experience, historicity, and existential meaning. In conclusion, the importance of a critical philosophical approach is emphasized to ethically guide the development and application of these technologies. Future research is encouraged to deepen the dialogue between philosophy, technology, and cognitive sciences, aiming to rethink the foundations of humanity in times of digital transformation.

**Keywords:** artificial intelligence; philosophy; consciousness; ethics; ontology.

## 1. INTRODUÇÃO

A ascensão da Inteligência Artificial (IA) nas últimas décadas tem suscitado debates não apenas no campo das ciências exatas e tecnológicas, mas também em áreas como a filosofia, a ética, a psicologia e o direito. A capacidade das máquinas de simular processos cognitivos humanos, como aprender, raciocinar e tomar decisões, levanta questões fundamentais sobre a natureza da mente, da consciência e da própria identidade humana. Nesse cenário, a filosofia apresenta-se como um campo fértil para a reflexão crítica sobre os limites, possibilidades e implicações da IA. Desde os primórdios da reflexão filosófica, pensadores têm buscado

compreender o que significa pensar, agir e existir, e tais questões são inevitavelmente reavivadas diante do surgimento de sistemas artificiais capazes de desempenhar funções antes atribuídas exclusivamente aos seres humanos.

Com o desenvolvimento acelerado de tecnologias baseadas em aprendizado de máquina, redes neurais e processamento de linguagem natural, tornou-se possível a criação de algoritmos que superam o desempenho humano em diversas tarefas específicas. No entanto, a performance funcional não implica, necessariamente, inteligência no sentido filosófico do termo. Seriam essas máquinas capazes de pensar? Elas possuem consciência ou apenas executam comandos baseados em padrões probabilísticos? A famosa indagação de Alan Turing — “As máquinas podem pensar?” — permanece atual e desafia os limites da compreensão humana sobre a mente e a inteligência (Turing, 1950).

Ademais das questões ontológicas e epistemológicas, a Inteligência Artificial impõe desafios éticos consideráveis. A delegação de decisões a sistemas automatizados implica em uma nova configuração da responsabilidade moral. Quem é responsável por decisões tomadas por um algoritmo? É legítimo que inteligências não humanas operem em campos sensíveis como a saúde, a justiça ou a segurança pública? As possíveis respostas a essas perguntas exigem uma análise cuidadosa que ultrapasse a técnica e adentre os campos da ética e da moral, conforme propõe Jonas (2006) ao tratar da responsabilidade frente às novas tecnologias.

Do ponto de vista filosófico, pensar a IA significa, também, refletir sobre o próprio humano. A fronteira entre homem e máquina parece cada vez mais tênue, o que suscita debates sobre o pós-

humanismo e o transumanismo, movimentos que propõem uma superação dos limites biológicos por meio da tecnologia. A IA, nesse contexto, surge tanto como possibilidade emancipadora quanto como risco existencial, como advertido por filósofos contemporâneos como Bostrom (2014).

Em face de tal desse cenário, este artigo propõe uma breve análise da Inteligência Artificial sob uma perspectiva filosófica, considerando os principais questionamentos sobre mente, consciência, ética e humanidade. A abordagem adotada visa contribuir para o aprofundamento das reflexões interdisciplinares que envolvem o avanço tecnológico e seus impactos no modo de ser e de viver do ser humano contemporâneo.

## **2. FUNDAMENTOS DA INTELIGÊNCIA ARTIFICIAL**

A compreensão da Inteligência Artificial (IA) exige uma análise que vá além das aplicações tecnológicas e operacionais. Suas bases conceituais estão fortemente enraizadas na intersecção entre filosofia, matemática, lógica e ciência da computação. Desde os primeiros esforços para descrever o funcionamento da mente humana de forma lógica, a ideia de replicar ou simular processos cognitivos por meio de artefatos artificiais esteve presente. No entanto, foi apenas em meados do século XX que a IA passou a ser concebida como um campo autônomo de estudo.

O termo "Inteligência Artificial" foi oficialmente cunhado em 1956 durante a conferência de Dartmouth, organizada por John McCarthy, Marvin Minsky, Claude Shannon e Nathaniel Rochester. Os pesquisadores propuseram que "todo aspecto do aprendizado ou qualquer outra característica da inteligência pode, em princípio, ser

descrito com tal precisão que uma máquina pode ser feita para simulá-lo" (McCarthy *et al.*, 1956, apud Russell; Norvig, 2021). Desde então, a IA tem sido desenvolvida com o objetivo de criar sistemas capazes de executar tarefas associadas à inteligência humana, como reconhecimento de fala, tradução de idiomas, planejamento e tomada de decisão.

Na obra clássica de Russell e Norvig (2021), os autores dividem os sistemas de IA em dois grandes paradigmas: IA fraca (ou estreita), que é especializada em tarefas específicas, e IA forte (ou geral), que busca replicar a inteligência humana de forma abrangente e autônoma. A IA fraca é aquela que domina atualmente o cenário tecnológico, presente em assistentes virtuais, sistemas de recomendação e algoritmos preditivos. Já a IA forte permanece, por ora, no campo da especulação e da pesquisa teórica, sendo objeto de intensos debates éticos e filosóficos, especialmente no que diz respeito à possibilidade de consciência ou intencionalidade nas máquinas.

Uma das principais bases teóricas da IA encontra-se na lógica formal e na computação simbólica. Inspirados pela lógica de Frege, Boole e Russell, os primeiros pesquisadores tentaram representar o raciocínio humano por meio de regras lógicas que pudessem ser implementadas em máquinas. Esse paradigma, conhecido como IA simbólica, predominou até a década de 1980. Contudo, suas limitações logo se evidenciaram diante da complexidade e ambiguidade do mundo real, que nem sempre pode ser traduzido em regras fixas (Zuboff, 2020).

Com o tempo, emergiram abordagens conexionistas, especialmente as redes neurais artificiais, inspiradas no funcionamento do cérebro

humano. Essas redes buscam aprender a partir de grandes volumes de dados, ajustando conexões internas para reconhecer padrões e prever comportamentos. Conforme explica Goodfellow, Bengio e Courville (2016), "o aprendizado profundo é uma subárea do aprendizado de máquina que utiliza redes neurais com múltiplas camadas para representar dados em diferentes níveis de abstração". Essa abordagem, que inicialmente foi criticada por sua falta de interpretabilidade, tornou-se dominante na década de 2010 graças ao aumento do poder computacional e à disponibilidade massiva de dados digitais.

É importante notar que o funcionamento da IA atual, baseada em aprendizado de máquina e análise estatística, não representa compreensão ou cognição no sentido humano. Segundo Searle (1980), a IA pode simular comportamentos inteligentes sem, no entanto, possuir qualquer tipo de entendimento ou consciência. Em sua famosa crítica ao funcionalismo, expressa no "experimento do quarto chinês", o autor argumenta que uma máquina pode manipular símbolos com precisão sem necessariamente atribuir-lhes significado, o que distingue essencialmente a cognição artificial da consciência humana.

Ademais dos aspectos técnicos, o avanço da IA está profundamente entrelaçado a questões econômicas e políticas. A pesquisa em IA tem sido fortemente impulsionada por grandes corporações e governos, interessados no poder de processamento e decisão automatizada. Conforme argumenta Zuboff (2020), o uso da IA em plataformas digitais alimenta um novo modelo de capitalismo baseado na extração comportamental, em que dados pessoais são utilizados para prever e modificar ações humanas com fins lucrativos. Tal crítica é fundamental para entender que a IA não é

apenas uma ferramenta neutra, mas uma construção social moldada por interesses e valores específicos.

Do ponto de vista epistemológico, a IA desafia concepções tradicionais de conhecimento. Enquanto o conhecimento humano envolve aspectos contextuais, intuitivos e afetivos, os sistemas artificiais operam com base em padrões extraídos de dados. Isso levanta questões sobre a confiabilidade das decisões automatizadas, sobretudo em ambientes complexos como a medicina, o direito e a segurança pública. Conforme Dreyfus (1992), a tentativa de reduzir a inteligência a regras computacionais falha em capturar a riqueza da experiência humana, marcada por incertezas, subjetividades e práticas corporificadas.

Ademais, o desenvolvimento da IA é atravessado por uma tensão entre autonomia e controle. Por um lado, os sistemas autônomos prometem eficiência, precisão e agilidade. Por outro, sua complexidade crescente dificulta a auditoria e a responsabilização, o que pode comprometer princípios fundamentais da democracia e da justiça social (Russel; Norvig, 2021). Observa-se que isso é relevante nos casos de decisões automatizadas que afetam diretamente os direitos das pessoas, como admissões escolares, concessão de crédito ou sentenças judiciais.

Por conseguinte, os fundamentos da IA não podem ser compreendidos somente em termos técnicos ou funcionais (Zuboff, 2020). Eles exigem uma abordagem interdisciplinar, que considere seus aspectos filosóficos, sociais e éticos. O entendimento da IA como sistema lógico, conexionista ou estatístico deve ser acompanhado por uma reflexão sobre o que se entende por inteligência, conhecimento e responsabilidade. Como destacam

Floridi e Chiriatti (2020), é necessário distinguir claramente entre agentes artificiais e seres humanos, sob pena de confundirmos eficiência algorítmica com sabedoria ou julgamento moral.

### **3. A INTELIGÊNCIA ARTIFICIAL E A FILOSOFIA DA MENTE**

A relação entre IA e Filosofia da Mente representa uma das intersecções mais provocadoras do pensamento contemporâneo, pois levanta questões fundamentais sobre a natureza da consciência, da cognição e da subjetividade. Em sua essência, a Filosofia da Mente busca compreender o que é a mente, como ela se relaciona com o corpo (especialmente o cérebro) e se é possível que sistemas artificiais possam ter estados mentais semelhantes aos humanos.

A IA, por sua vez, oferece um campo empírico para testar algumas dessas hipóteses filosóficas, ao mesmo tempo que desafia suas concepções tradicionais (Tononi; Koch, 2015).

Uma das principais discussões nesse campo é se máquinas podem, de fato, “pensar” ou apenas simular o pensamento. Alan Turing, em seu artigo clássico *Computing Machinery and Intelligence*, propôs um critério funcionalista para abordar essa questão: em vez de perguntar “as máquinas podem pensar?”, ele sugeriu que deveríamos perguntar se uma máquina pode imitar com sucesso o comportamento de um ser humano a ponto de ser indistinguível dele em uma conversa textual (Turing, 1950). Esse teste, conhecido como “Teste de Turing”, influenciou profundamente o desenvolvimento da IA, mas também foi alvo de críticas e objeções, como a famosa alegoria do “quarto chinês” formulada por John Searle.

Searle (1980), ao apresentar o experimento mental do quarto chinês, argumenta que um sistema computacional pode manipular símbolos (neste caso, ideogramas chineses) de acordo com regras preestabelecidas sem, no entanto, compreender seu significado. Para ele, isso mostra que, mesmo que um programa de computador possa se comportar externamente como se compreendesse uma linguagem, internamente ele não possui estados mentais conscientes. Essa distinção entre sintaxe e semântica é fundamental para sua crítica ao funcionalismo, que sustenta que os estados mentais podem ser reduzidos a operações formais sobre símbolos.

Em contraponto, os funcionalistas, como Fodor (1975), defendem que o que importa para caracterizar a mente não é a substância de que ela é feita, mas sim a organização causal das funções mentais. Nesse sentido, um sistema artificial que reproduzisse adequadamente as funções cognitivas humanas poderia, em princípio, ter uma mente. Para Fodor (1975), o pensamento humano pode ser compreendido como um processo computacional que opera sobre representações mentais, conceito que ele denomina de “linguagem do pensamento”.

Essa linha de pensamento encontra ecos na própria IA simbólica, que tenta representar crenças, desejos e intenções em estruturas formais manipuláveis. Entretanto, como apontado por Dreyfus (1992), essa abordagem falha ao tentar capturar aspectos não formais da experiência humana, como o saber tácito, a intuição e a corporeidade. Para ele, a inteligência humana não pode ser reduzida a regras e representações, pois ela emerge de uma interação complexa com o mundo, marcada pela historicidade e pelo contexto.

Outro campo importante da Filosofia da Mente que dialoga com a IA é o debate entre dualismo e materialismo. O dualismo cartesiano, que separa mente e corpo como substâncias distintas, foi amplamente questionado por abordagens materialistas contemporâneas, segundo as quais os estados mentais são fenômenos emergentes de processos físico-químicos no cérebro. Essa concepção naturalista da mente abre espaço para a possibilidade de que sistemas artificiais suficientemente complexos também possam gerar estados mentais (Dennett, 1991).

Dentro dessa perspectiva, o conceito de consciência artificial torna-se objeto de intensa especulação. Tononi e Koch (2015), por exemplo, propõem a Teoria da Informação Integrada (IIT), segundo a qual a consciência depende do grau de integração da informação em um sistema. Assim, seria teoricamente possível que sistemas artificiais, se suficientemente integrados, experimentassem algum tipo de consciência. No entanto, essa hipótese permanece altamente controversa, especialmente porque carecemos de critérios objetivos para identificar estados conscientes em entidades não humanas.

A Filosofia da Mente também contribui para discutir os limites da simulação cognitiva. Um sistema pode simular comportamentos empáticos ou éticos, mas isso implica que ele *sinta* empatia ou *compreenda* o valor moral de suas ações? Como ressalta Nagel (1974), há uma dimensão subjetiva na experiência consciente — o que ele chama de "qualia" — que parece escapar à descrição objetiva. Segundo ele, "há algo que é ser um morcego", ou seja, há uma experiência interna irreduzível que não pode ser captada por terceiros, tampouco por simulações externas.

Essa distinção entre simulação e realidade mental é crucial para compreender os limites da IA atual. Embora modelos de linguagem como o GPT-4 possam gerar respostas complexas e coerentes, sua produção textual não resulta de compreensão semântica, mas de cálculos probabilísticos (Rosa; Bruno, 2024). Eles não têm intencionalidade, no sentido filosófico proposto por Brentano — isto é, não dirigem seus estados mentais a objetos ou ideias com sentido.

Por outro lado, filósofos como Dennett (1991) defendem uma postura funcionalista e gradualista sobre a consciência, sugerindo que não há uma linha divisória nítida entre sistemas conscientes e não conscientes, mas sim graus de sofisticação cognitiva. Segundo ele, o que consideramos “consciência” pode ser o resultado de múltiplos processos computacionais paralelos que operam de maneira integrada. Em sua visão, a distinção entre humanos e máquinas pode se tornar cada vez mais tênue à medida que a IA evolui.

#### **4. QUESTÕES ÉTICAS E EPISTEMOLÓGICAS**

As implicações éticas da IA tornaram-se uma preocupação crescente à medida que sistemas inteligentes adquirem capacidades que antes eram consideradas exclusivamente humanas, como a tomada de decisões, o reconhecimento de padrões complexos e a interação com seres humanos em contextos sensíveis. Com isso, Rosa e Bruno (2024) observam que surgem questionamentos éticos sobre responsabilidade, autonomia, transparência, privacidade e o impacto da IA na dignidade humana. Tais questões têm mobilizado não apenas filósofos, mas também cientistas, legisladores e a sociedade civil, que buscam compreender

os riscos e limites do desenvolvimento e aplicação dessas tecnologias.

Uma das primeiras preocupações éticas está relacionada à autonomia da máquina frente à agência humana. Se um sistema de IA comete um erro ou causa um dano, quem deve ser responsabilizado? A ausência de intencionalidade e consciência nos sistemas artificiais impede que eles sejam considerados agentes morais plenos. Como observa Bostrom (2014), a IA não tem interesses ou intenções próprias no sentido humano, mas ainda assim pode causar consequências de grande escala se mal programada ou utilizada em contextos sensíveis. Assim, a responsabilidade recai inevitavelmente sobre os projetistas, programadores, usuários ou corporações que operam esses sistemas.

Um exemplo ilustrativo desse dilema ético é o uso de sistemas de IA em decisões judiciais ou na segurança pública. Em vários países, softwares de predição criminal têm sido utilizados para estimar a reincidência de acusados ou para definir prioridades policiais. No entanto, estudos como o de Angwin *et al.* (2016) mostram que esses algoritmos frequentemente reproduzem vieses raciais e sociais presentes nos dados com os quais foram treinados. Isso levanta preocupações quanto à imparcialidade e à equidade das decisões tomadas por máquinas, que podem reforçar desigualdades em vez de corrigi-las.

A par disso, há o problema da transparência algorítmica. Muitos sistemas de IA, especialmente os baseados em redes neurais profundas, operam como “caixas-pretas”, dificultando a compreensão de como determinadas decisões são tomadas. Isso

entra em conflito com princípios éticos fundamentais como o da prestação de contas e da justiça procedimental. Como afirma Floridi *et al.* (2018), uma ética da IA deve garantir que os sistemas sejam auditáveis, interpretáveis e acessíveis ao escrutínio público, principalmente quando afetam direitos fundamentais.

A privacidade também figura como uma das maiores preocupações na era da IA, especialmente com o uso massivo de dados pessoais para treinar e refinar modelos de aprendizado de máquina. Assistentes virtuais, sistemas de recomendação, vigilância automatizada e análise preditiva são exemplos de aplicações que frequentemente operam com uma coleta intensiva de informações sensíveis. Zuboff (2019) adverte que vivemos sob um novo regime de “capitalismo de vigilância”, no qual os dados pessoais são transformados em mercadoria, muitas vezes sem o consentimento explícito dos usuários. Tal contexto exige uma revisão profunda das políticas de proteção de dados e da autodeterminação informacional.

Em outro aspecto sensível diz respeito ao impacto da IA no mundo do trabalho. A automação inteligente ameaça substituir empregos em larga escala, não apenas em atividades manuais, mas também em setores altamente qualificados, como medicina, direito e jornalismo. Isso levanta dilemas sobre justiça distributiva, renda básica e o papel do trabalho na realização humana. Segundo Brynjolfsson e McAfee (2014), embora a IA traga ganhos de produtividade, seus benefícios não estão sendo distribuídos de forma equitativa, ampliando as desigualdades sociais. A ética da IA, portanto, deve considerar não apenas as ações das máquinas, mas os efeitos sistêmicos de sua implantação na sociedade.

Em um plano mais abstrato, a IA também coloca em questão os próprios fundamentos da ética. Se formos capazes de criar sistemas artificiais que simulem empatia, julgamento moral e comportamento ético, isso significa que a ética pode ser formalizada e programada? Essa possibilidade desafia a concepção tradicional de moralidade como algo intrinsecamente humano, dependente de experiência subjetiva e responsabilidade consciente. Moor (2006) propõe uma distinção entre “agentes éticos artificiais” — que seguem regras pré-programadas — e “agentes éticos autônomos”, que teriam a capacidade de deliberar moralmente. Atualmente, estamos ainda muito distantes da segunda categoria, mas a evolução da IA exige que reflitamos sobre os critérios que definem a agência moral.

Nesse sentido, diversos documentos internacionais têm proposto princípios éticos para o desenvolvimento e uso da IA. A UNESCO, em 2021, adotou a Recomendação sobre a Ética da Inteligência Artificial, que estabelece valores como o respeito à dignidade humana, à justiça, à inclusão, à diversidade e à sustentabilidade ambiental. Tais diretrizes visam orientar os países na elaboração de políticas públicas que assegurem que a IA seja usada para promover o bem comum e não a opressão ou exclusão social (UNESCO, 2021).

Do ponto de vista filosófico, essas preocupações éticas nos obrigam a revisar conceitos clássicos como liberdade, responsabilidade, justiça e bem-estar. Como observa Jonas (2006), a tecnologia moderna ampliou o alcance de nossas ações a tal ponto que os critérios éticos do passado já não são suficientes. Ele defende uma “ética da responsabilidade” diante do poder transformador da técnica, que deve levar em conta as consequências futuras, muitas vezes imprevisíveis, das inovações tecnológicas. A IA, como

tecnologia emergente e disruptiva, se encaixa perfeitamente nesse novo paradigma ético.

## **5. A INTELIGÊNCIA ARTIFICIAL E A ONTOLOGIA DO SER HUMANO**

O avanço da IA impõe uma reavaliação dos fundamentos ontológicos que sustentam a concepção do ser humano. A IA, ao simular aspectos do raciocínio, da linguagem e do comportamento humano, provoca um deslocamento das fronteiras entre o natural e o artificial, o humano e o maquínico (Bostrom, 2014). Nesse sentido, pensar a IA do ponto de vista ontológico implica investigar o que significa “ser humano” em um contexto em que máquinas parecem adquirir capacidades tradicionalmente atribuídas à humanidade. Tal reflexão remete às questões centrais da filosofia, especialmente àquelas que dizem respeito à consciência, intencionalidade, identidade e existência.

A ontologia, enquanto ramo da metafísica dedicado ao estudo do ser, encontra na IA uma provocação contemporânea. A tradicional distinção entre sujeito e objeto, por exemplo, torna-se nebulosa à medida que os sistemas inteligentes passam a interagir com o mundo de forma autônoma, respondendo a estímulos, aprendendo e, em certos contextos, até “dialogando” com humanos. Para Heidegger (1995), o ser humano é o único ente que compreende o ser, o que o caracteriza como Dasein — um ser-no-mundo com abertura para o ser e com capacidade para questionar sua própria existência. A IA, apesar de sua complexidade funcional, ainda carece dessa dimensão existencial que constitui o núcleo da ontologia heideggeriana.

Embora os sistemas de IA possam realizar tarefas de linguagem, reconhecer padrões e tomar decisões, isso não implica que possuam consciência ou autocompreensão. Searle (1990), em seu famoso experimento mental da “sala chinesa”, argumenta que a manipulação sintática de símbolos por um sistema computacional não equivale à compreensão semântica. Ou seja, uma IA pode “responder” adequadamente em chinês sem, no entanto, entender o que está dizendo. Tal argumento sustenta a ideia de que a IA não possui intencionalidade — termo que, na filosofia da mente, refere-se à capacidade da consciência de estar dirigida a algo, de possuir conteúdo mental com significado.

Nesse mesmo sentido, Dreyfus (1992) critica as pretensões da IA simbólica ao tentar modelar a inteligência humana por meio de regras explícitas. Segundo ele, a cognição humana é enraizada em uma corporalidade e em uma prática situada no mundo, o que escapa a qualquer modelagem puramente lógica ou computacional. O ser humano, para Dreyfus (1992), não é apenas um processador de informações, mas um ser encarnado, imerso em um contexto cultural, histórico e existencial. A IA, portanto, ao ignorar essa dimensão do ser-no-mundo, permanece fundamentalmente distinta da ontologia humana.

Ademais disso, a questão da identidade pessoal também é colocada em xeque diante da IA. Se for possível replicar padrões de comportamento, voz, linguagem e até emoções, o que resta de singular no ser humano? As tecnologias de “deepfake”, os clones digitais e os avatares personalizados, alimentados por algoritmos de IA, desafiam a noção de individualidade e autenticidade. Ricoeur (1991) oferece uma distinção útil entre identidade-idem (mesmidade) e identidade-ipse (autoidentidade), sendo esta última

construída ao longo do tempo, por meio de narrativas e compromissos éticos. Enquanto a IA pode simular a mesmidade, ela não parece capaz de desenvolver uma identidade narrativa, fruto de uma vida vivida em primeira pessoa.

O surgimento de agentes artificiais levanta ainda um debate ontológico sobre o lugar do ser humano no universo. A partir do humanismo renascentista, consolidou-se a ideia de que o homem é a medida de todas as coisas, dotado de razão, liberdade e valor intrínseco. No entanto, o desenvolvimento da IA, capaz de superar o desempenho humano em diversas tarefas cognitivas, pode sugerir uma relativização dessa centralidade. Bostrom (2014) alerta para a possibilidade de uma “superinteligência” que ultrapasse a capacidade humana em todos os domínios relevantes, o que exigiria uma redefinição ontológica da humanidade, talvez deslocando o homem para uma posição menos privilegiada na hierarquia do ser.

Contudo, a filosofia contemporânea também oferece caminhos alternativos para pensar essa relação. Simondon (2005), por exemplo, propõe uma ontologia das individuações, em que os seres — humanos ou técnicos — não são substâncias fixas, mas processos em constante formação. Para o autor, a técnica não deve ser vista como algo alheio ao humano, mas como uma extensão da individuação humana. Nesse sentido, a IA não ameaça o ser humano, mas o obriga a reconhecer sua própria condição técnica e relacional.

Outro ponto relevante jaz na distinção entre inteligência e sabedoria. Enquanto a IA pode exibir grande eficiência no processamento de dados e na otimização de decisões, ela não possui sabedoria, entendida como a capacidade de julgar com base

em valores, experiências e significados. Para Nussbaum (2010), a sabedoria ética e emocional é um traço essencial da humanidade, vinculado à vulnerabilidade, à empatia e à capacidade de reconhecer o outro. A IA, por mais avançada que seja, opera sem consciência moral e sem inserção em um horizonte de sentido.

Por conseguinte, a IA está longe de ameaçar diretamente a essência do ser humano, revela aspectos fundamentais da sua ontologia. Simodon (2005) argumenta que onfrontar-nos com máquinas que imitam a racionalidade, a linguagem e o comportamento, ela nos força a perguntar: o que nos torna humanos? A resposta a essa pergunta não está apenas nas capacidades cognitivas, mas também na consciência, na corporeidade, na historicidade e na alteridade. A IA, enquanto criação humana, espelha parte de nossa complexidade, mas também evidencia nossos limites e a profundidade do nosso ser.

Nesse contexto, é necessário resistir tanto ao tecnofetichismo quanto ao alarmismo. A ontologia do ser humano não pode ser reduzida aos parâmetros da eficiência técnica ou da imitação computacional. Como destaca Morin (2005), o ser humano é um ser complexo, que une razão e emoção, ordem e caos, natureza e cultura. A IA pode ampliar nossa compreensão sobre o que é pensar, mas jamais substituirá a experiência de ser — uma experiência única, encarnada e finita.

## **6. CONSIDERAÇÕES FINAIS**

A IA, enquanto fenômeno técnico-científico em constante expansão, impõe à filosofia desafios fundamentais. Ela simula processos cognitivos e comportamentais humanos e nos obriga a repensar

conceitos clássicos como consciência, intencionalidade, identidade e liberdade. Não se trata apenas de avanços tecnológicos, mas de implicações ontológicas e éticas profundas que atingem a própria noção de humanidade.

As diferentes correntes filosóficas analisadas evidenciam que, apesar da sofisticação funcional da IA, ela permanece desprovida da vivência subjetiva, do enraizamento existencial e da historicidade que caracterizam o ser humano. Muitos autores, referências na temática, reforçam a distinção essencial entre IA e existência humana, destacando a singularidade ontológica do Dasein e da consciência narrativa.

Contudo, não se trata de uma oposição simplista entre homem e máquina. A IA pode servir como espelho e extensão da condição humana, revelando tanto nossas capacidades quanto nossos limites. Ao mesmo tempo em que nos desafia, a IA também oferece oportunidades para aprofundar o autoconhecimento e repensar os rumos da vida em sociedade, desde que ancorados em princípios éticos e humanos.

Por fim, pontua-se que a filosofia cumpre papel crucial ao fornecer uma base crítica e reflexiva para o debate sobre a IA. Mais do que questionar o que a inteligência artificial pode fazer, é preciso refletir sobre o que significa ser humano diante dessas transformações. A resposta, como indicam as tradições filosóficas, não reside na técnica, mas no sentido, na ética e na existência compartilhada.

Algumas pesquisas futuras podem aprofundar a interface entre IA e filosofia, especialmente no que tange às implicações éticas e ontológicas da convivência com sistemas cada vez mais autônomos.

Investigações comparativas entre modelos filosóficos ocidentais e não ocidentais podem enriquecer o debate sobre o que significa consciência, subjetividade e agência em diferentes culturas.

Em epítome, estudos interdisciplinares que integrem filosofia, neurociência, ciência da computação e psicologia podem oferecer novas perspectivas sobre a possibilidade — ou limitação — da emergência de uma consciência artificial. Também se mostram promissoras as análises sobre o impacto da IA na formação da identidade humana em ambientes digitais e suas consequências para as relações sociais, o trabalho e a educação.

## REFERÊNCIAS BIBLIOGRÁFICAS

ANGWIN, Julia et al. *Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks.* ProPublica, 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 20 abr. 2025.

BOSTROM, Nick. *Superinteligência: caminhos, perigos, estratégias.* Trad. Rafael Mantovani. São Paulo: Editora Valentina, 2014.

BRYNJOLFSSON, Erik; MCAFEE, Andrew. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies.* New York: W. W. Norton & Company, 2014.

DENNETT, Daniel C. *Consciousness Explained.* Boston: Little, Brown and Co., 1991.

DREYFUS, Hubert L. *What Computers Still Can't Do: A Critique of Artificial Reason.* Cambridge, MA: MIT Press, 1992.

FLORIDI, Luciano et al. AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, v. 28, n. 4, p. 689–707, 2018.

FLORIDI, Luciano; CHIRIATTI, Michele. GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds and Machines*, v. 30, p. 681–694, 2020.

FODOR, Jerry A. *The Language of Thought*. Cambridge, MA: Harvard University Press, 1975.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. *Deep Learning*. Cambridge, MA: MIT Press, 2016.

HEIDEGGER, Martin. *Ser e Tempo*. Tradução de Fausto Castilho. Campinas: Papirus, 1995.

JONAS, Hans. *O princípio responsabilidade: ensaio de uma ética para a civilização tecnológica*. Trad. Marijane Lisboa. 3. ed. Rio de Janeiro: Paz e Terra, 2006.

MOOR, James H. The Nature, Importance, and Difficulty of Machine Ethics. *IEEE Intelligent Systems*, v. 21, n. 4, p. 18-21, 2006.

MORIN, Edgar. *Introdução ao pensamento complexo*. 4. ed. Porto Alegre: Sulina, 2005.

NAGEL, Thomas. What Is It Like to Be a Bat? *The Philosophical Review*, v. 83, n. 4, p. 435-450, 1974.

NUSSBAUM, Martha C. *Not for Profit: Why Democracy Needs the Humanities*. Princeton: Princeton University Press, 2010.

RICOEUR, Paul. *Soi-même comme un autre*. Paris: Seuil, 1991.

ROSA, J. F.; BRUNO, D. R. Evolução e o desenvolvimento do chatbot GPT-3 e GPT-4. *Revista Interface Tecnológica*, Taquaritinga, SP, v. 21, n. 1, p. 40–51, 2025. DOI: 10.31510/infa.v21i1.1849. Disponível em: <https://revista.fatectq.edu.br/interfacetecnologica/article/view/1849>. Acesso em: 21 abr. 2025.

RUSSELL, Stuart; NORVIG, Peter. *Inteligência Artificial*. 4. ed. São Paulo: Pearson, 2021.

SEARLE, John R. Minds, Brains, and Programs. *Behavioral and Brain Sciences*, v. 3, n. 3, p. 417–457, 1980.

SEARLE, John R. *Minds, Brains and Science*. Cambridge: Harvard University Press, 1990.

SIMONDON, Gilbert. *L'individuation à la lumière des notions de forme et d'information*. Grenoble: Jérôme Millon, 2005.

TONONI, Giulio; KOCH, Christof. Consciousness: Here, There and Everywhere? *Philosophical Transactions of the Royal Society B*, v. 370, n. 1668, 2015.

TURING, Alan Mathison. Computing Machinery and Intelligence. *Mind*, v. 59, n. 236, p. 433-460, 1950.

UNESCO. *Recomendação sobre a Ética da Inteligência Artificial*. Paris: UNESCO, 2021.

ZUBOFF, Shoshana. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York:

PublicAffairs, 2019.

ZUBOFF, Shoshana. *A era do capitalismo de vigilância: a luta por um futuro humano na nova fronteira do poder*. Trad. George Schlesinger. Rio de Janeiro: Intrínseca, 2020.

---

<sup>1</sup> Doutor em Engenharia e Gestão do Conhecimento (UFSC). Pós-Doutor em Gestão do Conhecimento e Administração. Docente e Pesquisador da Universidade Alto Vale Rio do Peixe (Uniarp). <https://orcid.org.0000-0003-3916-972X>. E-mail: [acesse o artigo original para visualizar o e-mail](#).

<sup>2</sup> Doutor em Ciência dos Materiais pela Universidade Federal do Rio Grande do Sul (UFRGS). Universidade Alto Vale do Rio do Peixe (UNIARP) <https://orcid.org/0000-0002-9054-8920>. E-mail: [acesse o artigo original para visualizar o e-mail](#).