Revista ARACÊ
ISSN: 2358-2472

PROJETO E COGNIÇÃO: ALGUMAS LEITURAS CRÍTICAS DA IA

PROJECT AND COGNITION: SOME CRITICAL READINGS OF AI

PROYECTO Y COGNICIÓN: ALGUNAS LECTURAS CRÍTICAS DE LA IA

https://doi.org/10.56238/arev8n3-125

Data de submissão: 26/02/2026 **Data de publicação:** 26/03/2026

José Luis Menegotto
Doutor em Arquitetura e Urbanismo
Instituição: Universidade Federal do Rio de Janeiro (UFRJ)
E-mail: jlmenegotto@poli.ufrj.br
Orcid: 0000-0003-4872-2574
Lattes: http://lattes.cnpq.br/4205576885019501

RESUMO
No presente artigo são desenvolvidas algumas reflexões críticas sobre os limites da Inteligencia Artificial. Iniciando com reflexões de Heidegger, quem se questionou acerca do significado do pensamento humano, surgem questões como a distinção entre pensamento e algoritmo, os riscos éticos, a desinformação e a responsabilidade humana sobre os seus produtos tecnológicos. Destacamos que as alucinações dos sistemas de IA revelam limites estruturais. Levantamos problemas, como a ausência de sistemas abertos e a restrição imposta ao conhecimento por arquiteturas de IA fechadas. Contrastamos duas visões de hibridização divergentes: a de transumanistas como Iannis Xenakis e Vernor Vinge. Prestamos atenção às fantasias tecnológicas, como a da superinteligência, que podem revelar atitudes políticas de controle totalitário. Se defende a construção de sistemas abertos com governança democrática para calibrar esses riscos de controle, vigilância e concentração de poder. Se destacam alertas éticos da conferência de Asilomar e a necessidade de transparência dos sistemas. Utilizamos como base as recentes críticas de Landgrebe e Smith.

Palavras-chave: IA. Asilomar. Singularidade. Transumanismo. Heidegger.

ABSTRACT
In this article, we develop critical reflections on the limits of Artificial Intelligence. Beginning with Heidegger's reflections, in which he questioned the meaning of human thought, we raise issues such as the distinction between thinking and algorithm, ethical risks, disinformation, and human responsibility for technological products. We highlight that AI system hallucinations reveal structural limitations. We identify problems such as the absence of open systems and the restrictions imposed on knowledge by closed AI architectures. We contrast two divergent visions of hybridization: those of transhumanists like Iannis Xenakis and Vernor Vinge. We draw attention to technological fantasies — such as superintelligence — that may reveal political attitudes of totalitarian control. The construction of open systems with democratic governance is defended to mitigate risks of control, surveillance, and concentration of power. Ethical warnings from the Asilomar conference and the need for system transparency are emphasized. Our analysis is grounded in recent critiques by Landgrebe and Smith.

Keywords: AI. Asilomar. Singularity. Transhumanism. Heidegger.

Revista ARACÊ
ISSN: 2358-2472

RESUMEN
En el artículo se desarrollan algunas reflexiones críticas sobre los límites de la Inteligencia Artificial. Partiendo de las reflexiones de Heidegger, quien se interrogó acerca del significado del pensamiento humano, surgen cuestiones como la distinción entre pensamiento y algoritmo, los riesgos éticos, la desinformación y la responsabilidad humana sobre sus productos tecnológicos. Destacamos que las alucinaciones de los sistemas de IA revelan límites estructurales. Señalamos problemas como la ausencia de sistemas abiertos y las restricciones impuestas al conocimiento por arquitecturas de IA cerradas. Contrastamos dos visiones divergentes de hibridación: la de transhumanistas como Iannis Xenakis y Vernor Vinge. Prestamos atención a las fantasías tecnológicas, como la de la superinteligencia, que pueden revelar actitudes políticas de control totalitario. Se defiende la construcción de sistemas abiertos con gobernanza democrática para calibrar estos riesgos de control, vigilancia y concentración de poder. Se destacan las advertencias éticas de la conferencia de Asilomar y la necesidad de transparencia en los sistemas. Tomamos como base las críticas recientes de Landgrebe y Smith.

Palabras clave: IA. Asilomar. Singularidad. Transhumanismo. Heidegger.

Revista ARACÊ
ISSN: 2358-2472

1 INTRODUÇÃO

No presente artigo são desenvolvidas algumas considerações críticas a respeito da das possibilidades de uso da Inteligência Artificial. Se filósofos que analisam os limites das palavras, da linguagem e do pensamento refletiram acerca das dificuldades de compreender o que é o pensamento humano, poderá a IA, que tenta operar sobre palavras, linguagens e pensamentos superar essas barreiras? Em 1951, Alan Turing (1990) inaugurava as discussões em IA perguntando se as máquinas poderiam pensar. Nesse mesmo ano, Heidegger registrava e refletia sua visão filosófica sobre o que seria 'pensar', o que significa essa ação e, principalmente, o que nos leva imperativamente a pensar. Heidegger se perguntava:

> "[…] O que significa pensar? […] […] O pensar não precisa justificar sua existência. Não precisa prestar contas. Não precisa demonstrar nada. Apenas precisa ser pensar. Em uma época em que tudo deve ser útil, novo, emocionante, profundo, verdadeiro, o pensar parece não ter lugar. Mas justamente por isso, a pergunta 'o que significa pensar?' torna-se inevitável. É a pergunta que nos acompanha rumo a uma nova época […]" (Heidegger, 2005:216).

A mais de sete décadas da formulação do filósofo, estamos vivendo a nova época que ele apontava em suas aulas. Hoje em dia, em um diálogo entre Miguel Benasayag e Ariel Pennisi pode ser lida uma reflexão que, de algum modo, coincide com a perspectiva heideggeriana. Pennisi observa que:

> "[…] De alguma maneira, pensar é não saber aonde se chega; é intrinsecamente ineficiente e radicalmente vital […]" (Benasayag et al., 2023:59).

Podemos repensar o tema da IA à luz desses enunciados filosóficos? Acreditamos que sim, pois partimos do pressuposto de que ainda não sabemos com exatidão o que é a substância pensante e o que nos impulsiona a pensar. A ciência e a filosofia suspeitam que seja algo que emerge, produto de interações entre o físico, o biológico e o simbólico. Embora as máquinas simulem o pensamento — e o façam de maneira cada vez mais verossímil — ainda assim, podemos insistir que é necessário continuar diferenciando o 'pensar' do 'processamento algorítmico'. Vejamos também um pequeno alerta extraído da Constituição de Claude, publicada recentemente. Claude é um sistema de agentes artificiais criado pela empresa Anthropic.

Revista ARACÊ
ISSN: 2358-2472

> "[…] A maioria dos casos previsíveis em que os modelos de IA são inseguros ou insuficientemente benéficos pode ser atribuída a modelos que possuem valores abertamente ou sutilmente prejudiciais, conhecimento limitado sobre si mesmos, sobre o mundo ou sobre o contexto em que estão sendo implementados, ou que não possuem a sabedoria necessária para traduzir bons valores e conhecimentos em boas ações. Por essa razão, queremos que Claude tenha os valores, o conhecimento e a sabedoria necessários para se comportar de maneira segura e benéfica em todas as circunstâncias. […]" (Anthropic, 2026)

A proliferação de publicações falsas ou enganosas que circulam nas redes sociais, difundidas de forma anônima por indivíduos ou organizações humanas inescrupulosas, é uma evidência da falta de 'sabedoria' ou 'discernimento' dos sistemas de IA imperfeitos que existem hoje e que são apontadas pela constituição de Claude. Sem supervisão humana, os sistemas não conseguem filtrar comportamentos humanos negativos nem aprender valores humanos positivos que lhes permitam filtrar comportamentos odiosos. No entanto, podemos nos perguntar: se os sistemas são capazes de extrair padrões estilísticos para gerar imagens ou textos, isso significa que alcançaram algum tipo de entendimento que lhes permitiria filtrar um padrão de ódio. Portanto, a proliferação de mensagens de ódio, fraudes ou falsidades não pode ser atribuída apenas aos algoritmos, mas sim a decisões humanas que os impedem de realizar o filtro ético. De qualquer modo, no amplo teatro do mundo, o fenômeno contemporâneo da mentira — impulsionada por intenções humanas fraudulentas e processada por engenhos digitais imperfeitos — vem causando efeitos negativos sobre nossos frágeis sistemas psíquicos e nossos frágeis sistemas políticos.

2 ALUCINAÇÕES

As alucinações são falhas conhecidas dos atuais sistemas de IA. Elas afetam sua confiabilidade e não podem ser atribuídas nem às ações humanas nem à imperfeição dos sistemas. Pesquisas sobre o uso de sistemas de IA baseados em grandes modelos de linguagem (LLMs) indicam que a matemática implicada nesses sistemas não garante que os resultados sejam logicamente consistentes (Neuhaus, 2023). Esse problema de confiabilidade está relacionado aos métodos numéricos empregados para processar os dados usados para responder às solicitações e ao conhecimento limitado que os sistemas podem ter do domínio processado no momento das consultas. Embora o conhecimento distribuído na Internet possa nos parecer imenso, ainda assim é limitado como conhecimento contextual do mundo — sem contar que potencialmente pode conter erros que desviem os cálculos executados das respostas corretas, afetando a confiabilidade geral do sistema. Relatos de interações conversacionais sem sentido com um agente de IA são frequentes. Podem ser esclarecidas algumas dúvidas. Os agentes de IA conseguem fazer cálculos? Sim, isso já é comum há muito tempo. Eles podem ordenar informações? Podem, desde que recebam instruções claras e dados completos; caso contrário,

Revista ARACÊ
ISSN: 2358-2472

organizarão informações incompletas, o que pode ser motivo para gerar erros. Eles resolvem problemas? Sim, desde que esses problemas sejam bem definidos, formalizados matematicamente e tenham uma solução possível. Quando essas condições não acontecem, surgem limitações típicas da IA como a falta de memória ou um processamento infinito. LLMs conseguem interagir entre si? Conseguem, mas isso ainda é raro. Cada empresa cria seu próprio modelo e o coloca no mercado, competindo com os demais. Sistemas com vários agentes trabalhando em conjunto ainda não são comuns, embora já existam condições técnicas para fazê-lo. Diante disso, surge uma pergunta: como poderia um agente como o Claude criar consensos seguindo suas próprias regras ao interagir com outros sistemas de IA, como Copilot ou Gemini, ou com modelos especializados, como Nano Banana, Alphageometry ou com os inúmeros LLMs que aparecem todos os dias? Apesar da publicidade favorável que os sistemas de gestão automatizados por IA recebem, suas arquiteturas atuais ainda são fechadas em si mesmas, imperfeitas e rudimentares para que possam chegar a estabelecer consensos de natureza política. Em outras palavras, que levem à construção coletiva de consensos humanos sinceramente orientados para o bem comum.

Nos preceitos constitucionais do modelo de IA Claude, torna-se explícita a proposta de corrigir problemas de autonomia epistêmica. Ou seja, a Constituição sugere que a autonomia do modelo de IA Claude teria mecanismos para filtrar as boas epistemologias das ameaças representadas pelas más epistemologias que, em teoria, poderiam estar contidas em algum sistema de IA sutilmente manipulador ou fraudulento. Embora a iniciativa tenha boas intenções, persiste a dúvida sobre se realmente seria possível cumpri-las, dado que o terceiro princípio dessa Constituição declara a adesão às diretrizes comerciais da empresa, o que poderia gerar vieses nas respostas do sistema. Na Constituição de Claude podem ser lidos os potenciais conflitos que surgem entre o comportamento humano e o comportamento esperado de um agente de IA sem discernimento, programado para ter um impacto positivo no mundo. O agente Claude é programado para priorizar os seus critérios de processamento seguindo quatro diretrizes na seguinte ordem: segurança, ética, aderência às diretrizes comerciais da empresa Anthropic e, por fim, ser genuinamente útil para o usuário (Anthropic, 2026), orientado pelo que define ser uma boa epistemologia.

Um LLM preparado para a gestão pode ser qualificado como superinteligente ou altamente eficiente, mas uma arquitetura construída sobre uma base de silos de dados fechados possivelmente impedirá que sejam avaliados espectros mais amplos de problemáticas ou que se possa ajudar a encontrar e construir consensos políticos benéficos para a maioria dos seres humanos. A hipótese que se formula neste artigo é que a arquitetura de qualquer sistema que pretenda ser qualificado como inteligente teria que ser necessariamente projetada para funcionar de maneira aberta, isto é, que os

Revista ARACÊ
ISSN: 2358-2472

dados gerenciados ou gerados por suas observações, medições ou inferências automáticas possam fluir livremente e ser cruzados com as soluções de outros agentes, além de serem absoluta e totalmente acessíveis para que as comunidades humanas possam usufruir e auditar. Sem mencionar que deveriam ser adaptados para lidar com realidades humanas, que são necessidades orgânicas, fisicamente estruturadas e biologicamente codificadas, que se desenvolvem em ambientes físicos e simbólicos temporais, interações que acontecem em contextos naturais, artificiais e imaginários.

3 HIBRIDAÇÃO HUMANO-MÁQUINA

Em relação aos processos de hibridização cibernética HM, podem ser contrastadas duas visões: a Singularidade de Vernor Vinge e o futuro imaginado pelo engenheiro e músico grego Iannis Xenakis. Nas décadas de 1950–1970, diante dos avanços da pesquisa genética, Xenakis especulava sobre o surgimento de uma nova espécie que poderia substituir a espécie humana. Sua visão era a de um pitagórico que imaginava uma transformação humana lenta e muito gradual, que ocorreria em um espaço temporal que ele imaginava acontecer em *"... muitas, muitas gerações..."* (Xenakis, 1992:261). Um futuro distante de nós. Embora hoje em dia já seja possível clonar, implantar ou associar diversos dispositivos eletromecânicos ou biopróteses ao nosso corpo, os processos de hibridização do animal com a máquina são possibilidades tecnológicas abertas que, com muito cuidado, vêm sendo metodicamente desenvolvidas pelas ciências médicas. Nós, continuamos sendo os mesmos seres de carne e osso. Os avanços da medicina são reservados para casos que possivelmente ninguém desejaria experimentar na própria pele, como mutilações decorrentes de desastres, acidentes, guerras, doenças ou necessidades regenerativas. Entendemos que o transumanismo não é necessariamente hibridização, nem um conjunto homogêneo de crenças. A imaginação visionária de Xenakis é muito diferente das visões propostas por Vernor Vinge. Se Vinge colocava a Singularidade acontecendo em uma janela temporal situada entre os anos de 2005 e 2030, o transumanismo humanista xenaquiano era, em realidade, uma cosmogonia poética incerta e muito distante.

Nesse sentido, acreditamos que também seja prematuro afirmar que a IA é inteligente, como é formulado pelas teorias tecnológicas mais otimistas. Parece pouco crível que possa haver inteligência em entidades fechadas em si mesmas, cujos promotores pretendem que tais engenhos demonstrem sua onipotência antecipadamente, criando expectativas desmesuradas ao fazer previsões para solucionar diversos problemas humanos, sem expressar dúvidas genuínas. Até hoje, a teórica superinteligência, tal como é prometida pelos teóricos mais otimistas da IA, parece ter os contornos das fantasias de controle totalitário. Como Varoufakis observou ao criticar o capital-nuvem, um dos mais conhecidos protagonistas e promotores dessa forma de poder econômico chegou, em certo momento, a expressar

Revista ARACÊ
ISSN: 2358-2472

o seu desejo de ter um "*App de tudo*" (Varoufakis, 2025:128). Com esta observação, não queremos expressar uma rebeldia ludita, nem minar a atividade mental dos visionários futuristas que imaginaram e ajudaram a criar as condições atuais da tecnologia digital. Se *'Imaginar é preciso'*, podemos acrescentar um 'o' a essa sentença para expressar que consideramos que *'Imaginar é precioso'*.

4 CRÍTICAS À IA DESDE A PERSPECTIVA POLÍTICA

Se os sistemas políticos evoluíram para diversas formas de democracias representativas, abandonando sistemas de governo autoritários e absolutistas, que motivo teríamos para dar forma a sistemas políticos controlados por engenhos e sistemas de IA concentradores de poder, que pretendam avançar seus controles sobre todas as ordens de nossas vidas? O desafio que é imaginado daqui em diante é trabalhar em sistemas que equilibrem o que Benasayag e Pennisi conceituaram como delegação de potências entre o humano e o maquínico (Benasayag *et al.,* 2023:99). Uma possível delegação de potência poderia relacionar-se, entre outras coisas, à intenção de reduzir cargas cognitivas desnecessárias no trabalho diário do ser humano. Quando tratamos do tema da programação de interfaces HM controladas por voz para ambientes de projeto CAD-BIM, a carga cognitiva à qual nos referíamos era aquela relacionada à necessidade que um projetista tem de lembrar as inúmeras posições e relações entre ícones, menus e procedimentos das interfaces gráficas (*Graphic User Interface, GUI*) dos programas de projeto. Essas interfaces costumam mudar de uma versão para outra do programa e pouco ou nada acrescentam como conhecimento para o objetivo principal das tarefas, que é desenvolver o projetos de arquitetura e engenharia. A hipótese era que uma interface de voz (VUI) eliminaria essa carga cognitiva inútil durante o esforço mental projetual, além de ajudar a diminuir a manipulação por toque dos complexos modelos BIM. Ao implementar a interface de voz, pensávamos nela como uma metáfora da folha em branco no início de um projeto. Testando opções, comprovamos a tensão que sentimos entre nosso corpo sensorial e a ação puramente intelectual. A interface de voz tensionou nosso sentido de motricidade gestual, até hoje necessário para o exercício gráfico (corporal), com o processo intelectual (mental) do nosso pensamento projetual (Menegotto, 2015). O fenômeno manifestou-se como se sempre tivéssemos que resolver uma dupla carga intencional: a do nosso corpo e a da nossa mente, atuando em velocidades diferentes. Como se a intenção gestual e a intenção intelectual não fossem parceiras e lutassem entre si para prevalecer. Quando interagimos com agentes inteligentes como o Copilot, torna-se cada vez mais difícil negar que o modo de comunicação do sistema parece aproximar-se muito do pensamento inteligente. Parece que as redes neurais multicamadas estão muito perto de cumprir a suspeita que Dreyfus observava nos anos iniciais da IA. Dreyfus observava que as técnicas de redes neurais desafiavam o pressuposto

Revista ARACÊ
ISSN: 2358-2472

filosófico segundo o qual o pensamento e a inteligência precisam contar com uma teoria unificada e prévia do mundo (Dreyfus, 1990:394). Nesse sentido, continuam sempre atuais e pertinentes os alertas filosóficos e sociológicos sobre as diversas distopias que podem ser imaginadas ao analisar a intromissão e o papel manipulador que os sistemas de IA podem exercer sobre nossas vidas. Talvez seja verdade que mecanismos sofisticados de controle biopolítico ou necropolítico possam ser utilizados para colonizar comportamentos humanos, até o limite do condicionamento conductista. Os temores de que a IA tenha efeitos contraproducentes para o processo civilizatório dos sistemas humanos são reais e atuais. Não se trata de teorias conspiratórias, pois são os próprios criadores desses sistemas que alertam sobre isso, como podemos ler na recente publicação da Constituição de Claude. A União Europeia vem tomando providências, colocando sinais de alerta na Consideração nº 7 do Regulamento Europeu sobre Inteligência Artificial (Diário Oficial da União Europeia 2024/1689). Outro alerta significativo provém da Conferência *Beneficial AI*, realizada em 2017. A conferência ocorreu em Asilomar, uma localidade situada em Pacific Grove, na costa oeste dos Estados Unidos. Em 1975, Asilomar sediou outra conferência que tratou de princípios éticos para pesquisas em biotecnologia. A conferência de 2017 colocou como pauta de debate as implicações éticas do uso da IA. Ao final do evento, foi redigido o documento *Princípios de Asilomar sobre a IA* (FLI, 2017), no qual consta uma lista de 23 recomendações sobre o direcionamento que os sistemas de IA deveriam seguir para alinhar-se em prol de uma IA benéfica para o ser humano. O documento considera princípios éticos para a prosperidade do desenvolvimento humano; o estabelecimento de bases de responsabilidade; a necessidade de evitar uma corrida armamentista letal com IA; e, diversas recomendações sobre formas de delegação HM.

Em 1950, Turing formulava objeções àqueles que criticavam suas teses sobre a máquina pensante. Chama a atenção e até provoca certa perplexidade, recordar como reagiu a comunidade de especialistas em IA quando surgiu a versão pública do ChatGPT, no final de 2022. Num paralelismo que faz lembrar da brincadeira com a qual Turing ironizava de seus críticos, os especialistas pediram aos laboratórios de IA que suspendessem por seis meses o desenvolvimento de agentes inteligentes. Fizeram isso na carta aberta *Pause Giant AI Experiments: An Open Letter*, publicada no site do Future of Life Institute (FLI, 2023). Em um trecho da carta, solicita-se aos laboratórios de pesquisa de IA que:

> "[…] A pesquisa e o desenvolvimento em IA deveriam ser reorientados para tornar os sistemas atuais, potentes e de ponta, mais precisos, seguros, interpretáveis, transparentes, robustos, alinhados, confiáveis e leais […]" (FLI, 2023)

Revista ARACÊ
ISSN: 2358-2472

Foi como se, de repente a comunidade de especialistas tivessem adotado a posição do avestruz que Turing, com humor, atribuía a seus críticos. É verdade que a IA deu um salto evolutivo surpreendente e abrupto; ainda assim, é difícil entender como aqueles que vinham trabalhando nesse campo de estudos há tanto tempo puderam ser surpreendidos a ponto de solicitar uma pausa de seis meses no desenvolvimento de uma tecnologia que acumulava não menos de 60 anos de pesquisas e trabalho. Vernor Vinge não apenas imaginava essas possibilidades na década de 1990, como também havia estabelecido uma data para que isso aconteça (entre 2005 e 2030).

O que se impõe agora é reconhecer a seriedade que o tema merece e concentrar esforços na criação de mecanismos sólidos de governança democrática, distribuída, transparente e humana para os renovados engenhos da IA. Mas construir esses engenhos não é tarefa simples, pois são os próprios criadores que, na carta-manifesto, chegam a reconhecer que não conseguem compreender totalmente os métodos de processamento do tipo "caixa-preta", sugerindo que os projetos de IA deveriam ser concebidos como sistemas de "caixa-transparente". Chega também a ser paradoxal, e até contraditório, criar uma tecnologia e não poder conhecer como ela funciona. Landgrebe e Smith são céticos quanto à solução do problema da explicabilidade das Redes Neurais Profundas (dNNs), problema que deu origem a uma linha de estudos em IA que tenta entender causalmente como um modelo estocástico dNN chega aos seus resultados (Landgrebe *et al.,* 2023:165). Podemos suspeitar que, no pedido de suspensão de seis meses formulado na carta, esteja oculto um problema que não é de natureza técnica, mas política. Algo assim como uma trégua estratégica entre aqueles que ambicionam acumular e concentrar poder e aqueles que defendem a ideia de redes democráticas e descentralizadas de IA, capazes de distribuir equitativamente o poder que o conhecimento confere. A célebre fórmula da filosofia política — "saber é poder" — parece ser o epicentro do problema colocado pela carta-manifesto do *Future of Life Institute*. A IA é uma tecnologia digital que, de algum modo, amplifica a sensação de onipresença panóptica imaginada por Jeremy Bentham. No século XVIII, Bentham idealizou um tipo de edifício penal em que os guardas estariam estrategicamente posicionados em um centro óptico, a partir do qual poderiam ver todos os presos. O conceito foi analisado por Michel Foucault em *Vigiar e Punir*. Como os detentos sentem a presença constante da vigilância panóptica, o panóptico torna-se um método de controle exercido por pressão psicológica, uma coerção que conduz à autodisciplina. De certo modo, as tecnologias da informação possuem esse potencial panóptico (e panacústico), hoje potencializado pela IA que pretende saber tudo. Tais sistemas não apenas poderiam exercer um controle posicional, como já exercem certo controle preditivo sobre nossas intenções. Esse potencial pode ser um incentivo para a concentração de poder por novas classes dominantes, configurando novas alianças produtivas e ambientes vitais que Varoufakis (2025)

Revista ARACÊ
ISSN: 2358-2472

caracterizou como *tecno-feudos*, dos quais muitas populações e sistemas produtivos poderiam ser excluídos. Varoufakis sugere que o poder e a vigilância hoje podem ser exercidos pela simples exclusão de um link nos motores de busca. Como ele observa, o exercício de uma forma de *"terror higienizado é a base do tecno-feudo"* (Varoufakis, 2025:124). Podemos levar em conta os alertas de outros observadores da realidade, que já vinham notando há bastante tempo — muito antes do surgimento dos atuais agentes de IA — que ao controle sobre os corpos das populações deveria somar-se o controle exercido de modo mais sutil e sofisticado pela manipulação mental. A biopolítica e a necropolítica encontram-se no epicentro dessas discussões. Peter Sloterdijk, por exemplo, ao referir-se aos objetivos de sua obra *Esferas*, expressa com certa ironia que a escreveu como

> "[…] uma tentativa de relativizar essa insuportável e crescente cisão entre os mundos do saber e criar algo como um esoterismo democrático. […] Para ele, […] o 'espaço psíquico' do ser humano atravessa uma época de confusão e esquecimento em uma situação única de entorpecimento […]" (Sloterdijk et al., 2007:127).

Não apenas seria ingênuo, mas também uma imprudência não levar em conta todas essas observações ao propor sistemas de trabalho com os novos meios digitais. Apoiados nas críticas de Landgrebe e Barry Smith (2023), podemos duvidar dos anúncios proféticos feitos pelos teóricos mais otimistas da IA contemporânea. Mas é evidente que a IA evoluiu até consolidar-se como um conjunto de engenhos diferenciados e potentes, cujo controle configura um novo campo de disputas.

5 CRÍTICAS À IA DESDE A PERSPECTIVA TÉCNICA

As respostas que obtemos dos novos sistemas que utilizam grandes modelos de linguagem (*Large Language Models*) são, em geral, suspeitosamente afirmativas e positivas. Um sistema de IA que alucina implica que em suas respostas podem estar misturadas afirmações que parecem verdadeiras (e talvez o sejam) com afirmações que reconhecemos como falsas. Mas o limite entre o que é certo e o que é errado, falso ou ficcional torna-se cada dia mais difuso e difícil de distinguir. Os problemas processados por sistemas LLM são sempre transformados em problemas decidíveis e tratáveis. Qualquer solicitação feita a um agente LLM é reduzida a uma matriz numérica que é respondida por outra matriz numérica, após passar por um processamento de cálculo no qual se combinam métodos numéricos que usam equações diferenciais (determinísticas) e equações probabilísticas (estocásticas). Anand Nikhil explica como os modelos de *Machine Learning* aplicam o cálculo de derivadas para encontrar gradientes e atribuir pesos relativos aos conjuntos de dados processados, com o objetivo matemático de melhorar as previsões feitas ao longo do tempo, ajustando os erros estatísticos (Nikhil, 2025).

Revista ARACÊ
ISSN: 2358-2472

6 O LIMITE DOS SISTEMAS NÃO-ERGÓDICOS PARA A IA

No oitavo capítulo do livro *Why Machines Will Never Rule the World*, Jobst Landgrebe e Barry Smith (2023) apresentam um compilado das diversas técnicas computacionais usadas para implementar modelos de IA. Eles destacam as dificuldades enfrentadas pelos modelos matemáticos para realizar previsões em situações nas quais participam sistemas complexos não ergódicos. A palavra ergódico vem do grego *ergon*, que significa 'trabalho', e *hodos*, que significa 'caminho'. A teoria ergódica surgiu no campo da matemática para estudar o comportamento estatístico dos sistemas dinâmicos, ou seja, aqueles sistemas cujo comportamento pode ser descrito em um espaço de fases, configurado como planos bidimensionais, espaços tridimensionais ou espaços abstratos de dimensões arbitrárias, que evoluem no tempo. George David Birkhoff foi o matemático que contribuiu com desenvolvimentos fundamentais nesse campo. Suas pesquisas matemáticas também influenciaram o campo da pesquisa estética. Birkhoff se concentrou em determinar se, nos sistemas complexos, existe uma média estatística única e coincidente entre o evento temporal observado e sua configuração espacial. Colocando a pergunta: podem ser encontrados padrões estatísticos estáveis na evolução de um sistema complexo? Nos sistemas ergódicos, sim. Mas Landgrebe e Smith pensam que muitos sistemas vitais complexos podem ser sistemas não ergódicos, com processos cujos eventos dificilmente apresentam um padrão de evolução homogêneo ou estável e, portanto, seriam matematicamente não modeláveis (Landgrebe *et al.,* 2023:149).

Por exemplo, um fenômeno raro e cuja natureza talvez seja não ergódica poderia ser a floração gregária de algumas plantas. Se tomarmos como exemplo a palmeira Talipot (também chamada palmeira do Ceilão), sabe-se que sua floração é gregária e monocárpica, ou seja, acontece apenas uma vez na vida, antes de morrer. Acredita-se que esse momento seja causado por algum mecanismo temporal interno à planta, que pode ocorrer quando ela tem entre 60 e 80 anos de vida. Depois de florescer, a palmeira lança suas sementes no ambiente e morre ao longo do ano seguinte. Em novembro de 2025, no Parque Urbano do Flamengo, no Rio de Janeiro, pôde-se observar esse raro e vistoso fenômeno de floração de indivíduos de palmeiras Talipot, plantadas pelo paisagista Roberto Burle Marx na década de 1960. Em casos assim, os sistemas de aprendizado de máquina não poderiam fazer previsões precisas do evento. Se o fizessem, seria possivelmente por meio de um cálculo estatístico ingênuo e simplificado, que poderia ocultar ou eliminar as incertezas dos processos vitais. Por isso é tão importante que os processamentos sejam transparentes ou combinados com métodos paralelos que permitam adaptar o tipo de cálculo realizado, informando sempre ao ser humano o método aplicado. Outros fatores que ocasionam imprecisões no aprendizado com modelos estocásticos incluem o fato de que os sistemas não conseguem lidar bem com espaços de dados dispersos ou com dados novos

Revista ARACÊ
ISSN: 2358-2472

que não estavam presentes nos cálculos prévios (Landgrebe *et al.,* 2023:169). Poderíamos pensar que o florescimento recente da IA é um prenúncio de sua hipotética morte ou de que um novo inverno se aproxima? Algo como a morte das palmeiras Talipot? De tempos em tempos, a IA entra em seu inverno e congela por um período, depois de passar anos acumulando energia. Mas, ao morrer, espalha novas sementes que a farão renascer. Poderiam as necessidades energéticas apontadas por Landgrebe e Smith causar um novo inverno para a IA? Imaginem um engenho gigantesco que, para ser superinteligente, precisasse ser alimentado com toda a energia do mundo. Seria a força bruta em sua máxima expressão. O processamento por força bruta é um fato conhecido e temido pela IA, que a acompanha desde o seu nascimento e que ela parece não conseguir superar. Hoje em dia, a força bruta do processamento não se manifesta apenas na computação realizada dentro dos microprocessadores dos nossos computadores pessoais. Ela saiu da CPU para manifestar-se no novo mundo dos grandes centros logísticos de dados da IA, uma força bruta gigantesca que precisa ser alimentada com quantidades consideráveis de energia.

O teorema ergódico de Birkhoff, pode ser relacionado a outra de suas especulações e teorias matemáticas, concebida no início da década de 1930: a medida estética de Birkhoff. O matemático imaginou um coeficiente calculado em função de dois parâmetros que sempre estão presentes em obras de arte ou, de modo geral, em objetos produzidos pelos seres humanos: Ordem e Complexidade. Simplificando, a medida estética é calculada como uma razão entre a Ordem e a Complexidade de uma obra (Birkhoff, 2013)

$$M = f(O/C)) \qquad (1)$$

Três décadas depois, essa teoria influenciaria pesquisas no campo da arte. Os pesquisadores se dedicaram a estudar as produções artísticas com esses métodos de orientação matemática. O objetivo era investigar se seria possível mensurar conceitos como o da beleza, isto é, buscavam respostas para questões que tensionam o objetivo e o subjetivo, o físico e o espiritual, o que é mensurável e tudo aquilo que parece escapar à medição concreta. A tese matemática de Birkhoff foi seguida por Max Bense, que, na década de 1960, buscou racionalizar a estética, sistematizando a teoria da medição estética. Bense dedicou-se a examinar composições de obras artísticas (literatura, pintura, música, arquitetura etc.), imaginando uma autonomia do mundo estético em relação ao mundo físico. Ele conceituava a atividade artística como um processo orientado a alcançar um estado de entropia mínima (Bense, 2003:231). A entropia é um conceito utilizado pela física para descrever o grau de dispersão da energia contida em sistemas físicos (como partículas e gases). Bense transportou a entropia para o

Revista ARACÊ
ISSN: 2358-2472

campo da teoria da arte, imaginando que uma composição artística persegue alcançar um estado de entropia mínima, com o qual chegaria ao seu ideal, algo assim como uma ordem absoluta intrínseca à obra, a forma da perfeição estrutural desejada pelo artista. A estratégia de buscar patrões de estilo é um método de geração compositiva usado pelos modelos de IA generativa. Esses modelos numéricos são devedores das técnicas matemáticas e analíticas herdadas dos estudos de Birkhoff, Bense e outros pesquisadores do campo da arte eletrônica ou da estética informacional. Pode ser formulada a pergunta: a geração artística é um processo ergódico, formalizável e previsível? Do ponto de vista da filosofia da arte, há dúvidas de que assim seja. Embora reconheça méritos nas teorias da medição estética, a filósofa da arte Susanne Langer faz ressalvas às propostas que tentam caracterizar as produções artísticas a partir de leis matemáticas (Langer, 2006:112). A arte pode usar a matemática para sua criação, mas isso não significa que ela seja o resultado de um cálculo, nem que a arte precise de engenhos de previsibilidade. O humanista e pitagórico Xenakis, criador da música estocástica, que para compor suas obras musicais usava sons calculados em turbulências e fluxos sonoros imaginados, também fez críticas taxativas às especulações entrópicas e analíticas de Bense (Xenakis, 1992:77). Com isso, podemos retornar às dúvidas de Landgrebe e Smith, que sustentam que a IAG não poderá superar os limites impostos pelos sistemas complexos não ergódicos, que não são apenas físicos. A eles também poderíamos acrescentar os sistemas complexos não físicos, como os sistemas da metafísica estética, que carregam consigo cargas simbólicas contidas nas produções artísticas. Ao definir a arte da morfologia geral e a música estocástica, por exemplo, Xenakis falava sobre a *metamúsica*, no que consideramos ser um evidente desejo de escapar das limitações do mundo físico por meio dos pórticos imateriais da arte. A palavra grega *stoa* significa pórtico, uma estrutura composta por três elementos: duas colunas e uma viga, sistema estrutural usado pelos gregos em seus templos e edifícios cívicos. Portanto, sejamos cautelosos, pois os resultados obtidos com sistemas de aprendizado de máquina podem transitar por inferências e narrativas que atravessam repetidas vezes as fronteiras entre o real e o ficcional, ou entre a criação artística autêntica e a criação de pastiches clonados de segunda ordem.

7 CONCLUSÕES

Embora possam obter altas taxas de acerto, gerar imagens vistosas, composições musicais ou contos, os sistemas de IA parecem incapazes de adotar a dúvida metódica cartesiana sem enfrentar sérios problemas de processamento, nem de incorporar as energias simbólicas intrínsecas às obras humanas. Ainda assim, geram em nós o sentimento da presença de algo que parece ser inteligente, um espelho imperfeito daquilo que acreditamos ser e fazer como seres humanos. É pouco provável que

Revista ARACÊ
ISSN: 2358-2472

existam respostas para todas as perguntas. Os modelos de IA parecem carecer da inteligência do silêncio. Eles transformam todos os problemas em problemas decidíveis e tratáveis; por isso, suspeitamos que ainda não seja possível qualificar os sistemas de IA como inteligentes, nem confiar cegamente em suas previsões, nem qualificar suas composições artísticas como produtos originais. Nós, seres humanos, sabemos na própria carne que um pensamento que valha a pena ser retido pode demorar a surgir e, ainda assim, somos conscientes de que podem existir variáveis que não consideramos durante nossos processos de pensamento, raciocínio ou intuição. A eliminação da dúvida talvez seja uma boa razão para pensar a Inteligência Artificial contemporânea como um gigantesco engenho que pode nos enganar e que, apesar de seu desempenho, ainda precisa conquistar a confiabilidade geral da inteligência do silêncio e da lentidão.

Revista ARACÊ
ISSN: 2358-2472

REFERÊNCIAS

ALPHAGEOMETRY. https://github.com/google-deepmind/alphageometry . 2024.

ANTHROPIC. *Claude Constitution*. www.anthropic.com/constitution , 2026.

BENASAYAG, M., PENNISI, *A. La inteligencia artificial no piensa. (El cerebro tampoco)*. Ed. Prometeo. Ciudad Autónoma de Buenos Aires. 2023.

BENSE, M, *Pequena estética*. São Paulo. Perspectiva. 2003.

BIRKHOFF, G, D. *Aesthetic measure*. Harvard University Press. 2013.

BODEN, M. A. Filosofía de la Inteligencia Artificial. Londres: Oxford University Press. 1990.

_____ Dimensões da criatividade. Porto Alegre: Ed. Artes Médicas Sul. 1999.

CHURCHLAND, P. M. Matéria e consciência. Uma introdução contemporânea à filosofia da mente. São Paulo. Fundação Ed. As UNESP. 1998.

DREYFUS, H.; DREYFUS, S. La Construcción de una mente versus el modelaje del cérebro: La inteligencia artificial regresa a un punto de ramificación. Artificial Intelligence Vol. 117 n°1. 1988.

FLI. FUTURE LIFE INSTITUTE. *Asilomar AI Principles. Beneficial AI 2017.* https://futureoflife.org/open-letter/ai-principles , 2017.

FLI. FUTURE LIFE INSTITUTE. *Pause Giant AI Experiments: An Open Letter,* 2023. https://futureoflife.org/open-letter/pause-giant-ai-experiments/ , 2023.

HEIDEGGER, M. *¿Qué significa pensar?* Traducción de Raúl Gabás. Ed. Trotta. Madrid, 2005.

LANDGREBE, J., SMITH, B. Why Machines Will Never Rule the World. Artificial Intelligence without Fear. New York. Routledge. 2023.

LANGER, S. K. Sentimento e forma. São Paulo: Perspectiva. 2006.

MENEGOTTO, J.L. A Framework for Speech-Oriented CAD and BIM Systems. en: Gabriela Celani; David Moreno Sperling; Juarez Moara Santos Franco. (Org.). Communications in Computer and Information Science. 1°Ed. Berlin: Springer Berlin Heidelberg, 2015, v.1, pp. 329-347. 2015.

NEUHAUS, F. Ontologies in the era of large language models - a perspective. Applied Ontology. v.18. n.4. pp.399-407. Dezembro 2023. DOI: 10.3233/AO-230072. 2023.

NIKHIL, A. I wasted years confused about how LLMs learn before learning this. I wish someone had told me this when I started. Medium. https://medium.com/me/lists, 2024.

NON-ERGODICSYSTEMS.https://energy.sustainability-directory.com/term/non-ergodic-systems , 2025.

SLOTERDIJK, P. Crítica de la Razón Cínica. Madrid. Ed. Siruela. 2003.

Revista ARACÊ
ISSN: 2358-2472

_____ Regras para o parque humano. São Paulo: Estação Liberdade. 2000.

SLOTERDIJK, P, HEINRICHS H. J. O sol e a morte. Investigações dialógicas. Lisboa: Relógio D' Água Editores. 2007.

SMITH, B. New desiderata for biomedical terminologies. in K. Munn and B. Smith (eds.), Applied Ontology: An Introduction, Frankfurt/Lancaster: Ontos, p. 83-109. 2008.

TURING, A. La maquinaria de la computación y la inteligencia. En: Boden Margaret Ann. Filosofía de la Inteligencia Artificial. México, D.F.: Ed. Fondo de Cultura Económica. 1990.

UE. Regulamento da União Europeia. 2024/1689 del Parlamento Europeu e do Conselho da União Europeia. https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L_202401689, 2024.

VAROUFAKIS, Y. Tecno feudalismo. O que matou o capitalismo. Tradução: Érika Nogueira Vieira. São Paulo. Ed. Planeta do Brasil. 2025.

VINGE, V. The Coming Technological Singularity: How to Survive in the Post Human Era. En: Vision-21: Interdisciplinary Science and Engineering in the Era of Cyberspace, G. A. Landis, Ed., NASA Publication CP-10129, pp.11-22. 1993.

XENAKIS, I. *Formalized Music: Thought and Mathematics in Composition.* New York: Pendragon Press. 1992.