

A expansão artificial da inteligência humana

Lucia Santaella

Pontifícia Universidade Católica de São Paulo (PUC-SP)

 <https://orcid.org/0000-0002-0681-6073>

lbraga@pucsp.br

RESUMO

De poucas décadas para cá, a Inteligência Artificial viu-se transformada em um hype. Não é para menos, pois os efeitos e consequências de suas aplicações em tudo que diz respeito ao humano é indiscutível. Infelizmente, entretanto, quando um tema se torna moeda corrente de publicações nem sempre bem fundamentadas, ele fica submetido a um imbróglio de controvérsias, mal-entendidos e desinformação, justamente o que vem acontecendo com a IA. As razões para isso são muitas. Este artigo pretende colocar em discussão algumas dessas razões, com ênfase no Zeitgeist tanto filosófico quanto tecnocientífico em que a questão da inteligência se acha envolvida.

Palavras-chave: inteligência artificial; hype; Zeitgeist; divulgação científica.

The artificial expansion of human intelligence

ABSTRACT

In the last few decades, Artificial Intelligence has become a hype. No wonder, because the effects and consequences of its applications in everything that concerns humans is indisputable. Unfortunately, however, when a topic becomes the currency of publications that are not always well founded, it is subjected to an imbroglio of

controversies, misunderstanding and misinformation, precisely what has been happening with AI. The reasons for this state of affairs are many. This article intends to discuss some of these reasons, with emphasis on both the philosophical and the techno-scientific Zeitgeist in which the question of intelligence is involved.

Keywords: artificial intelligence; hype; Zeitgeist; science communication.

Submissão em: 07/09/2023 | **Aprovação em:** 30/03/2023

1. INTRODUÇÃO

A inteligência sempre foi um tema de especial interesse da psicologia e também da educação. De umas poucas décadas para cá, entretanto, a inteligência se viu transformada na grande questão do momento. Há três razões fundamentais para isso: a) no campo da filosofia, a emergência do pensamento sobre o não humano; b) na área do meio ambiente, as preocupações e dilemas trazidos pelo Antropoceno, o novo período geológico da Terra e c) o sucesso alcançado pela inteligência artificial que trouxe a inteligência para o centro dos debates.

Este artigo dará breve atenção ao contexto filosófico do não humano e, então, à problemática do Antropoceno, que se constituem em fontes de questionamento das tradicionais concepções do humano particularmente frente às inteligências não humanas. A seguir, atenção mais detalhada será dada ao debate acerca da inteligência, com ênfase na inteligência humana em seu confronto com a inteligência artificial.

2. A EMERGÊNCIA DO NÃO HUMANO

Realismo especulativo (RE), virada especulativa ou ainda ontologia orientada ao objeto (OOO) são as denominações do mais recente movimento que marcou um novo caminho para a filosofia. Pode-se determinar seu nascimento e desdobramentos a partir do *Workshop* que ocorreu no *Goldsmiths College*, na Universidade de Londres, em 27 de abril de 2007 e cujo registro consta na íntegra na Revista *Collapse* (Brassier et al., 2007, p. 307-450). Embora tenha nascido na filosofia, seu impacto se fez notar em várias áreas do conhecimento, tais como “antropologia, arqueologia, arquitetura, literatura inglesa, feminismo, artes plásticas, estudos medievais, musicologia, retórica e composição, estudos científicos e outros” (Harman, 2011, p. vii). Um dos pontos-chave desse movimento encontra-se em sua oposição e mesmo antagonismo à virada linguística que dominou e exerceu grande influência sobre muitos campos do pensamento no século XX.

Em seu lugar, o realismo especulativo propôs um outro tipo de virada, ou seja, a virada para o objeto.

São muitos os fatores que poderiam nos levar a explicar por que o RE repercutiu capilarmente em tantas áreas do conhecimento e da produção criativa. Antes de tudo, deve-se considerar o *Zeitgeist* ou sincronias temporais que propiciam que determinados fenômenos, situações ou conjunturas tomem relevo porque aparecem na hora certa e no lugar certo. Assim, sem que houvesse qualquer relação explícita entre ambos, o RE nasceu em meio a debates sobre o pós-humano que já vinham se desenvolvendo nos meios intelectuais desde finais dos anos 1980. Não custa lembrar dos efeitos provocados nos estudos culturais críticos pela publicação do *Manifesto Ciborgue* de Danna Haraway (1985). A tônica do pós-humano versava sobre as transformações do humano provocadas pelas tecnologias. Isso coincidia, sob certos aspectos, com a ruptura defendida pelo RE da relação dual entre sujeito e objeto do conhecimento, sob a primazia do sujeito, vindo daí a virada para o objeto.

O ponto nítido de convergência dos debates do pós-humano com o RE situa-se mais especificamente no deslocamento e descentramento do humano em similaridade de condições com aquilo que Bruno Latour (1994a, p. 140; 1994b) chamou de “parlamento das coisas”, no qual os novos protagonistas são os actantes, tudo que nos rodeia e que utilizamos em nossas práticas, inclusive científicas, a saber, objetos, ferramentas, instrumentos que “fazem fazer” de acordo com as operações de tradução, mediação, intermediários, mediadores etc., operações nas quais não têm mais lugar as relações dicotômicas pautadas na primazia do sujeito do conhecimento ou do fazer sobre a subserviência do objeto (ver Lemos, 2013).

O parlamento das coisas brotou na consonância com a Teoria Ator-Rede (TAR) e seu ataque agudo às categorias e conceitos opositivos que dominaram o cenário do pensamento ocidental por séculos. Para a TAR, verdade/falsidade; grande/pequeno; agência/estrutura; humano/não humano; antes/depois; conhecimento/poder; contexto/conteúdo; materialidade/socialidade; atividade/passividade são visões separatistas que foram borradas no trabalho realizado pela TAR. A associação do pensamento de Latour com a TAR foi inevitável. Na sua versão, pessoas, animais, coisas,

objetos, aparelhos, dispositivos e instituições podem funcionar como actantes em rede, ou seja, agenciadores que fazem agir (Latour, 1996, 1998; Santaella e Lemos, 2010). Não obstante a reviravolta que provoca nos esquemas conservadores de interpretação do mundo, as propostas da TAR expandiram-se da sociologia da ciência e da tecnologia para a sociologia em geral, a geografia, a administração, os estudos das organizações, a antropologia e a filosofia.

Ademais, a parte fundamental do *Zeitgeist* em sintonia com a virada para o objeto do RE, conforme foi discutido em Santaella (2013), encontra-se no advento da internet das coisas (IoT) que está transformando em seres sencientes, comunicacionais, aquilo que, com o nome de meras “coisas”, costumávamos tratar como seres inertes ao nosso dispor. Essa emergência, no contexto de outras invenções que a digitalização ubíqua está trazendo, reclama pela renovação do pensamento ontológico, especialmente da ontologia do humano que herdamos do passado.

Não levou muito tempo para que tanto o RE quanto a TAR fossem absorvidos, desde o início da segunda década do nosso século, no contexto maior que ficou conhecido como a virada do não humano (*nonhuman turn*) que engloba estudos interdisciplinares das mais diversas ordens, todos eles na defesa do descentramento do humano no seio da biosfera. Assim, o não humano passou a englobar o mundo animal e seus direitos, a teoria dos afetos mobilizada pelas discussões sobre diversidade sexual, as interrogações sobre os corpos, os sistemas orgânicos e geofísicos, as novas materialidades nas teorias feministas, as neurociências, as ciências cognitivas, as vidas artificiais, as questões da inteligência e consciência. Em suma, um pool de complexidades que desafiam nossas compreensões da realidade e que exigem o nosso engajamento com o que não é humano, tais como mudanças climáticas, secas, fome, biotecnologia, genocídio, terrorismo, guerra e até mesmo o Antropoceno, o novo período geológico do planeta, fruto do peso e feridas que as ações humanas, muitas vezes insanas, imprimiram sobre a biosfera (GRUSIN, 2015, p. vii).

3. O ANTROPOCENO E SUAS CONSEQUÊNCIAS

Foi só em 2000 que a denominação Antropoceno, para marcar uma mudança no período geológico da Terra, foi proposta por Paul Z. Crutzen, e por Eugene F. Stoemer, em uma publicação do *International Geosphere-Biosphere Programme*. Sem nos determos em detalhes científicos complexos, o Holoceno, no qual estávamos inseridos, é uma época geológica pós-glacial, cujo início data de doze mil anos. Trata-se de uma época marcada, por exemplo, pelo advento da agricultura, então, pela formação de cidades, meios de comunicação e por novas formas de energia. Durante toda essa época, o regime climático estabilizou-se com climas temperados e os humanos se fizeram presentes em todos os continentes.

Entretanto, nas últimas décadas, foi se criando um consenso entre os cientistas de que a biosfera deixou de se comportar dentro dos padrões do Holoceno: a atmosfera, o clima, os oceanos e o ecossistema estão todos operando fora das normas holocênicas. A antropoesfera, a camada humana, que cresceu dentro da biosfera, deixou pegadas destrutivas profundas sobre a Terra a ponto de dar passagem a um novo período geológico, o Antropoceno (do grego *anthropos* = ser humano + *Kainos* = recente, novo).

Há muitas discussões relativas ao ponto no tempo em que esse novo período geológico teve início. Duas opções dominam nas discussões: final do século XVIII, com a revolução industrial, ou então 1945, depois da Segunda Guerra Mundial. No primeiro caso, a análise do ar mostra o início da concentração global crescente do dióxido de carbono e metano, o que coincide com a invenção da máquina a vapor, 1784, por James Watt. É a partir dessa data também que teve início a explosão demográfica que então estava na margem de um bilhão, enquanto existem prognósticos de que neste século a população chegue a dez bilhões de pessoas. As florestas tropicais desaparecem em um ritmo alarmante soltando dióxido de carbono e levando à extinção de várias espécies.

No segundo caso, diante da “grande aceleração”, ou melhor, da mudança do ritmo da própria mudança instaurada a partir do final da Segunda Guerra, 1945 é sinalizado como data inicial do Antropoceno. Foi essa época que marcou o surgimento da era nuclear e da assinatura radioativa que ela deixou sobre a superfície da Terra.

Paralelamente, a economia global cresceu dez vezes em apenas algumas décadas, provocando a perda da biodiversidade, com o ritmo de extinção das espécies da biosfera correndo cem a mil vezes mais rapidamente. Enfim, o grande problema é que, acompanhando esse ritmo, estamos consumindo exponencialmente mais recursos da natureza.

Até recentemente, o estudo das eras, épocas e períodos geológicos parecia importar pouco para o restante dos campos de conhecimento e, certamente, importava menos para os leigos. Entretanto, se a revolução copernicana já foi capaz de transformar a visão humana sobre o ambiente em que vive, agora não há como deixar de compreender criticamente como o ser humano está implicado nos destinos do planeta. Diante de tudo isso, o tema do Antropoceno, nos últimos anos, deixou de ser privilégio dos especialistas, espalhando-se por várias áreas do conhecimento, inclusive das artes, e repercutindo cada vez mais intensamente nas mídias informativas, tais como jornais, revistas, TV e internet.

Relevante para a questão da inteligência humana e não humana, é notável que as condições do Antropoceno provocaram a emergência de movimentos ativistas e artivistas a par de fértil produção discursiva que alertam para uma necessária transformação da noção exclusivista da inteligência humana, demonstrando em pesquisas legitimadas a existência de muitas formas de inteligência não humanas que se estendem por todo o mundo vegetal, naquilo que ficou conhecido como a revolução das plantas (Mancuso, 2019) até as mais diversas modalidades de inteligência animal (Bridle, 2023). Tanto o RE e a TAR quanto o Antropoceno nas suas reivindicações, que são também de natureza política, proclamam a necessidade de expandir a noção de inteligência para além de sua reclusão estrita no reino do humano. É nesse contexto que as versões bem-sucedidas da Inteligência Artificial (IA) emergiram, trazendo, elas próprias novos questionamentos sobre a inteligência em geral, a inteligência humana, a inteligência artificial e a relação entre ambas.

4. A EMERGÊNCIA BEM-SUCEDIDA DA INTELIGÊNCIA ARTIFICIAL

Desde seu nascimento em meados do século passado, a inteligência artificial é um campo de conhecimento que pertence ao feixe das ciências cognitivas, um conjunto de ciências distintas, mas que se unem no foco comum das pesquisas sobre as habilidades cognitivas capazes de elucidar os aspectos constitutivos da inteligência. Nos seus setenta anos de existência, a IA passou por momentos de euforia, seguidos por disforias que só se dissiparam há perto de duas décadas.

De fato, atualmente a Inteligência Artificial (IA) se constitui no *hype* das pesquisas tecno científicas com repercussões em muitas esferas das atividades humanas. Entretanto, quando se trata de definir a IA as dificuldades surgem, especialmente porque a presença do substantivo “inteligência” só traz dificuldades, pois são muitas as definições de inteligência nem sempre concordantes. Quando se acrescenta o adjetivo “artificial”, as dificuldades só tendem a aumentar. Diante disso, é preciso buscar um substrato de concordância que nos leve a uma definição breve e relativamente consensual extraída daquilo que é comum em vários textos sobre o tema.

Assim, a IA significa a simulação pelos sistemas computacionais de processos de inteligência humana. Trata-se de um ramo das ciências da computação voltado para a criação de máquinas inteligentes. Isso implica o desenvolvimento maquínico de habilidades tais como são, por exemplo, a aprendizagem, o conhecimento, a aquisição de informação e as regras para usá-la, o raciocínio utilizado para alcançar conclusões definidas ou aproximadas, a autocorreção, a solução de problemas, a percepção, o reconhecimento e processamento linguístico, o planejamento, e a habilidade para manipular e mover objetos. Para realizar esses fins, o computador precisa ter acesso a objetos, categorias, propriedades e relações.

De uns anos para cá, a euforia provocada pela IA é fruto de vários fatores. Entre eles, principalmente o aumento considerável do poder computacional, o agigantamento dos dados digitais e as redes neurais artificiais. Desde que as válvulas elétricas foram substituídas pelos transistores (1947), as máquinas computacionais, munidas de

microprocessadores (1971), foram diminuindo de tamanho proporcionalmente ao aumento de seu desempenho. A escalabilidade ou capacidade de *upgrade* das máquinas acompanhou o crescimento progressivo de dados a serem processados. Há indicadores não atualizados de que as informações do mundo dobram exponencialmente a cada dois anos, sendo referidas como *Big Data*, caracterizado não apenas por seu volume, mas por uma rica mistura de tipos e formatos de dados (variedade) e pela natureza sensível ao mesmo tempo que marca um desvio do processamento em lote tradicional (velocidade). Esse enorme volume de dados permitiu o funcionamento mais eficaz das redes neurais artificiais cujo surgimento não é tão recente como se pode pensar.

Na história das ciências cognitivas, as redes neurais surgiram como reação à primeira fase da inteligência artificial regida pelo modelo computacional da mente. Conforme já desenvolvido em Santaella (2004, p. 77-78), dos anos 1950 em diante, a inteligência artificial passou a buscar as condições formais da atividade cognitiva, capazes de indicar o que é comum a todos os sistemas que exibem essa atividade, quer ela apareça em animais, máquinas ou humanos. Para isso, o modelo computacional apresentava habilidades para simular processos cognitivos, com promessas de modelar a mente. Em vista disso, esse modelo dominou por certo tempo como o paradigma clássico unificador das ciências cognitivas.

O modelo partia de dois pressupostos: antes de tudo, “a relativa autonomia entre o *hardware* e o *software* das máquinas utilizadas para simular a inteligência, o que permite explicar o comportamento inteligente de qualquer sistema complexo sem pressupor o tipo físico ou biológico da inteligência de seus componentes”. Isso conduzia à “compreensão da mente como um conjunto de representações de tipo simbólico e regidas por um conjunto de regras sintáticas” (Santaella, *ibid.*). Conclui-se, a partir disso, que o pensamento seria o resultado da ordenação mecânica de uma série de representações ou símbolos, ordenação esta que não pressupõe necessariamente a existência de um cérebro. Por isso, o aparato mental concebido como dispositivo lógico pode ser descrito por meio de um conjunto de computações abstratas. Nessa perspectiva, simular a inteligência não implica a construção de máquinas com *hardware* específicos, mas sim o desenvolvimento de programas computacionais operando sobre dados ou representações. Chamado de

inteligência artificial simbólica, esse paradigma sedimentou-se a partir de final dos anos 60 tanto nos trabalhos do grupo liderado por Newell e Simon quanto nos de Marvin Minsky e Seymour Papert (Teixeira, 1998, p. 36, 43).

Não é difícil inferir que está aí implicada uma mudança no conceito de inteligência que passa a ser definida como capacidade para produzir e manipular símbolos, tendo em vista a resolução de problemas. Essa concepção da inteligência e da mente teve e continua tendo consequências filosóficas que não são desprezíveis, especialmente para as seculares questões sobre as relações entre mente e corpo e, em tempos mais recentes, diante dos avanços das neurociências, para as relações entre mente e cérebro. A teoria da mente que está por trás do paradigma computacional é o funcionalismo que exerceu domínio quase exclusivo nas ciências cognitivas até os anos 1980 quando se deu uma virada para o conexionismo. Aproveitando-se dos avanços da época das pesquisas sobre redes de neurônios formais, o conexionismo propôs a replicação da inteligência por meio da construção de redes neurais artificiais. Usando técnicas dotadas de propriedades que podem ser interpretadas em termos cognitivos, essas redes são capazes de aprender, reconhecer formas, memorizar por associações etc.

Enquanto o modelo computacional da mente, também chamado de cognitivismo, estuda os processos mentais como computações abstratas, independentemente de suas formas específicas de concretização, o conexionismo pretende simular o cérebro como meio para simular a atividade mental. Os conexionistas tomam o cérebro humano como um dispositivo computacional em paralelo que opera com milhões de unidades similares aos neurônios. Se computadores e cérebros têm como função principal processar informações, então redes neurais artificiais podem ser construídas para simular esse processo.

As redes neurais constituem uma intrincada rede de conexões entre suas unidades que são dispostas em camadas hierarquicamente organizadas. Conectadas umas com as outras, unidades estimuladas via *inputs* externos excitam ou inibem outras unidades gerando padrões de conectividade. Diferentemente dos sistemas computacionais simbólicos, os conexionistas são sistemas dinâmicos compostos por um “conjunto de processos causais através dos quais as unidades se excitam ou se inibem, sem empregar

símbolos e tampouco regras para manipulá-los”. Pretende-se que esse conjunto de neurônios artificiais modele a cognição (Teixeira, 1998, p. 84). Embora infinitamente menos complexo do que o cérebro, esse modelo da mente inspira-se na estrutura e modo de funcionamento do cérebro, chegando assim mais próximo da realidade biológica da mente. O que importa em sistemas desse tipo são os complexos padrões de atividade entre as múltiplas unidades que constituem uma rede (ibid., p. 103).

O conexionismo representou um passo importante para a compreensão do funcionamento da IA. Entretanto, o tempo não estava maduro para o seu desabrochar. O poder computacional ainda estava longe de alcançar o patamar em que se encontra desde as duas últimas décadas. Então, a quantidade dos dados era incipiente. Nos anos 1980, a WWW, as redes sociais e outras formas de captura de dados não haviam acontecido. Era preciso que tudo isso viesse à tona para que a IA emergisse com a força que hoje tem e, graças à grande estrela da aprendizagem de máquina, transformar-se no hype do momento, a ponto de asfixiar muitos outros.

5. A ÊNFASE NA INTELIGÊNCIA

A ênfase na questão da inteligência está colocada, desde meados do século passado, no âmbito das ciências cognitivas de que as ciências da computação e hoje a ciência dos dados, acrescidas das pesquisas em IA, são partes constitutivas. A par disso, para compor o *Zeitgeist*, os especialistas em tecnologias em geral e na filosofia da tecnologia em particular integraram o tema da inteligência em seus estudos, com atenção para as transformações pelas quais o conceito de inteligência vem passando no contexto de uma preocupação com a necessidade de se repensar a própria ontologia do humano.

Afinal, a espécie humana é uma espécie em evolução pelo simples fato de que o *Sapiens* é uma espécie que se constitui como inalienavelmente inacabada, sempre em busca de si mesma e de um alvo que, embora sejamos parte dele, desconhecemos. Como se não bastasse, tanto a filosofia que ilumina as mutações sociotécnicas próprias da digitalização e contemporaneamente da dataficação e performatividade algorítmica, à luz da TAR, quanto também os debates sobre o Antropoceno vêm proclamando a

urgência do deslocamento e descentramento do humano junto com o conceito de inteligência com que se viu acompanhado pela tradição.

No contexto desse *Zeitgeist* é curioso observar como o campo da IA -- especialmente o campo daqueles que rodeiam a IA e recebem seus efeitos, mas não estão dentro da especialidade das pesquisas -- tem revelado uma resistência diante da premência do enfrentamento do fato de que a IA está exigindo uma transformação na própria ideia daquilo que o passado nos fez acreditar sobre o que é inteligência, a saber, no caso, como sendo uma exclusividade humana intocável.

Mais curioso ainda é notar que a proposta de expansão da noção de inteligência para o compartilhamento da inteligência humana com aquelas dos animais e mesmo das plantas é recebida sem dificuldades, enquanto o compartilhamento com a inteligência das máquinas, que atualmente alcançou o estado da arte da IA generativa, é recebido com os mais diferentes tipos de entraves. Como explicar essa dificuldade? A tradição nos mostra que a busca de separação entre o orgânico, particularmente o organismo humano, e as máquinas vem de longe e se encontra bem representada tanto na filosofia cartesiana como em versão mais renovada na filosofia de Heidegger.

Sugestões de resposta para as oposições entre humano e máquinas podem ser encontradas no fato de que o humano pode experimentar muitos afetos, inclusive compaixão e piedade; máquinas não sentem nada. Elas funcionam dentro de limites estreitamente definidos, e quanto mais padronizados, mais esses limites se tornam rígidos. Os organismos, ao contrário, são mais abertos às potencialidades do acaso, das surpresas, das renovações e mesmo das mutações. A vida é experiência, o que significa improvisação, a vida é tentativa e busca em todos os respeitos.

Na concepção obsoleta de máquinas que ainda abrigamos, máquinas são compostas de partes. Elas são armadas e desarmadas. Elas estão abertas a modificações ou reconstruções. Para o ser humano, a ideia de ganhar novas partes é apenas um pouco menos horrível do que perder partes. A máquina tanto solta quanto adquire novas partes com a maior facilidade. A perspectiva humana parece insistir sobre a integridade orgânica como a única norma possível (Wilson 1996, p. 246). Tudo isso indica que “somos vítimas de uma esperança nostálgica de retorno à plenitude de um corpo

uno-primordial. Uma esperança teimosamente renovada, mas fadada ao malogro, pois tal plenitude só existe na saudade sonhada da carícia materna que tremeluz nas neblinas da imaginação” (Santaella, 2003, p. 228).

Isso explica por que, dentre todas as linguagens, técnicas e tecnologias, só a fala e o gesto são sentidos como naturais enquanto as outras técnicas são tidas como artificiais, estranhas, irreconhecíveis. Isso ocorre porque, não obstante sejam também técnicas, fala e gesto ainda estão inseridos no próprio corpo, são ainda partes integrantes do corpo e, portanto, não são percebidas como técnicas. O estranhamento e as resistências manifestam-se quando as técnicas começam a se lançar para fora do corpo como extrojeções de habilidades do cérebro, a extrojeção da memória, por exemplo, na escrita, ou extrojeções do gesto, na máquina fotográfica como amplificação do sentido da visão, ou nos aparelhos eletroacústicos como amplificação do sentido da audição. Desde essas extensões sensoriais até a emergência da cultura mediada pelo computador, as máquinas deixaram de funcionar como meras ferramentas, para se transformarem em tecnologias da inteligência em evolução acelerada que, no ponto em que estamos, alcançaram a simulação das faculdades discursivas e imagéticas humanas. É o estranhamento e mesmo preconceito diante das tecnologias inteligentes, que a IA leva a um certo limite, que se constitui na celeuma em torno do tema da inteligência que hoje reina no campo da IA, mesmo antes do abalo sísmico que a IA generativa vem provocando.

6. O IMBRÓGLIO DA INTELIGÊNCIA

No seu estágio atual a IA é movida por aprendizado de máquina (AM) e por sua vertente mais bem sucedida, a aprendizagem profunda (AP). De acordo com a literatura especializada, existem pelo menos cinco grupos diferenciados de desenvolvimento da IA: os simbolistas, as connexionistas, os evolucionários, os bayesianos e os analogistas. Tudo indica que, entre eles, o grupo que tem alcançado maior sucesso, pela performatividade de suas aplicações, é o connexionismo.

Conforme já desenvolvido em Santaella (2023), na sua versão da IA preditiva, antes da IA generativa, o connexionismo opera com redes neurais artificiais mais

profundas do que em qualquer outra época e elas definem novos padrões para a visão, o reconhecimento da voz, a descoberta de medicamentos e outras áreas. Para os conexionistas, aprendizado é a função do cérebro que aprende ao ajustar as forças das conexões entre neurônios, e o problema crucial é descobrir quais conexões são responsáveis por quais erros e alterá-las de acordo. Mimeticamente ao funcionamento das redes sinápticas do cérebro, o algoritmo mestre dos conexionistas é a *backpropagation* (propagação retrógrada, retropropagação ou propagação reversa), que compara a saída de um sistema com a saída desejada e altera sucessivamente as conexões em cada camada de neurônios para aproximar mais a saída daquilo que ela deveria ser, ou melhor, que se espera que ela seja (Domingos, 2017, p. 77-78).

Esses rudimentos de explicação técnica parecem suficientes para colocar em discussão aquilo que importa neste artigo, ou seja, a inteligência humana e a IA, colocando em relevo o imbróglio em que esse tema está hoje envolvido. Como ponto de partida, é preciso considerar que a IA é uma área de pesquisa ultra especializada, envolvendo complexidades matemáticas, estatísticas e probabilidades que não estão ao alcance dos leigos. Todavia, quando um *hype* toma conta da cultura, da política e da economia, todos se julgam autorizados a se pronunciar sobre o tema, mesmo quando não tiveram a menor curiosidade para enfrentar as informações especializadas, tanto internas quanto externas sobre a questão. Esse é o péssimo legado que as redes sociais nos deixaram e que se alia ao costumeiro sensacionalismo das mídias noticiosas que nadam a braçadas naquilo que pode atrair seu público. A par disso, não obstante a indiscutível relevância da divulgação científica, ela pode produzir um desserviço quando se deixa contaminar pelo sensacionalismo ou pela autocomplacência.

O primeiro grande foco do imbróglio, composto por controvérsias, equívocos e desinformação vem da confusão entre consciência e inteligência. Felizmente os pesquisadores estão longe de cair nessa cilada, antes de tudo porque sabem que seu *metier* diz respeito à inteligência e não à consciência, esta última uma questão que é melhor ser deixada para os filósofos e psicólogos e, daqui a algum tempo, provavelmente para a neurociência.

De fato, há um certo consenso entre os pesquisadores de IA de que consciência não se confunde com inteligência, embora, no animal humano, ambas estejam indiscutivelmente entrelaçadas. Infelizmente na filosofia cognitiva, a consciência tem sido reduzida às discussões sobre o *qualia* quer dizer, “o sentido subjetivo de uma experiência difícil de capturar em quaisquer tipos de sistemas formais”: por exemplo, a vermelhidão do vermelho como é sentida por cada um de nós. “Tomamos como garantido que a outra pessoa está vendo a mesma cor, mas não sabemos se ela tem a mesma experiência subjetiva que eu tenho.” O segundo aspecto refere-se ao sentido do self. Experimentamos o mundo “em uma certa espécie de modo unitário e experimentamos a nós mesmos como estando nele.” Não o experimentamos em termos de dezenas de milhões de neurônios acesos (Tenenbaum, 2018, p. 482-483).

Outros pontos de convergência dos especialistas acerca da consciência dizem respeito à persistência no tempo da autoconsciência (Bostron (2018, p. 11), a saber, à constatação de que a consciência consiste em nossas experiência subjetivas, ou seja, consiste em nosso mundo interior intransferível e intraduzível. Diante disso, Koller (2018, p. 395) afirma que chegar à IA Geral ou superinteligência dispensa a consciência, “pois é possível ter um sistema incrivelmente inteligente, que não tem nada a ver com uma consciência interior.” Turing já havia afirmado que a consciência não é conhecível. Tanto isso é verdade que o surgimento da IA generativa tem sido acompanhada pelos prognósticos de uma possível inteligência geral para o futuro, nos quais o tema em foco é a inteligência em si e não a consciência.

De todo modo, a partir dessas convergências, tendo a concordar que a consciência é um fenômeno interior, e por ser interior, é difícil encontrar formas de compartilhamento entre consciências, tanto quanto é difícil saber se outros animais também a têm. Por outro lado, por ser simultaneamente interior e exterior, a inteligência humana, por seu lado, é compartilhável por meio de linguagens e sistemas de codificação externos à mente, tendo, por isso, uma natureza coletiva (para mais detalhes, ver Santaella, 2022 e 2023, onde essa questão é tratada detalhadamente).

O outro foco do imbróglio em que a questão da inteligência, em especial a computacional está envolvida, provém dos negacionistas da inteligência artificial, quer

dizer, da negação peremptória de que haja qualquer forma de inteligência na inteligência artificial, o que leva à consideração de que o substantivo “inteligência” não lhe cabe. Essa negação se constitui em uma verdadeira ladainha de vozes.

Na maior parte das vezes, o que está por trás dessa ladainha é o *parti-pris* de que inteligência é tão só e apenas humana. Quaisquer outras formas de inteligência são, portanto, corruptelas de inteligência ou nem mesmo inteligência, um problema que é criticamente comentado por Rus (2018, p. 263). Para essa autora, “há um grande equívoco na imprensa popular sobre o que é e o que não é IA. Quando dizem IA referem-se a aprendizado de máquina e, mais que isso, aprendizagem profunda dentro da aprendizagem de máquina. As pessoas tendem a antropomorfizar o que esses termos significam. Os não especialistas associam com inteligência e tomam esta como inteligência humana.

Quanto ao termo “artificial”, ele é utilizado para se referir à inteligência das máquinas e não à expressão mais ampla de inteligência não humana que inclui a inteligência dos animais para aqueles que não se limitam a uma visão antropocêntrica da inteligência. Para Russel (2018, p. 63) uma entidade pode ser considerada inteligente na medida em que faz a coisa certa, significando que se espera que suas ações atinjam seus objetivos. Esse princípio aplica-se tanto aos humanos quanto às máquinas. “Fazer a coisa certa é a chave do princípio unificador da IA. Para isso, a IA precisa de algumas habilidades-chave, percepção, visão, reconhecimento de fala, e ação. Especificando, trata-se “da habilidade de tomar decisões, planejar, e resolver problemas, também da habilidade de comunicar, portanto, o entendimento da linguagem natural se torna muito importante para a AI” (ibid.).

Um argumento que abre caminhos contra os discursos negacionistas baseia-se na questão do aprendizado. A inteligência humana é viva e metabólica, implicando o aprendizado contínuo, base para a consideração da plasticidade do cérebro. Mas, se o aprendizado se encontra no coração das habilidades de que a inteligência é feita, esse parece ser o nó-górdio que implica considerar equivocado negar a inteligência às máquinas. Sim, os algoritmos de IA são inteligentes e, assim o são, primordialmente porque aprendem, auto-organizam-se, corrigem seus erros e chegam aos seus alvos.

Tanto é assim que a comparação entre os processos humanos de aprendizagem e aqueles dos algoritmos tornou-se um tema de investigação dos especialistas em IA, isto porque, certamente, para construir máquinas inteligentes, eles não podem escapar da comparação com a complexidade dos humanos.

De fato, os defensores da inteligência computacional consideram a IA inteligente porque o computador adquiriu o potencial de aprender e tomar decisões com base nas informações que recebe. Isso não desconsidera o fato de que esse tipo de inteligência é decididamente diferente da nossa. Por suas habilidades de imitação, a IA tem a qualidade de identificar padrões informacionais que otimizam tendências relevantes para os resultados a que chegam. Além disso, ao contrário dos humanos, a IA não possui um corpo metabólico regulado por hormônios, um corpo que sente fadiga, tem sono, um corpo inseparável da mente que é muitas vezes atormentada por complicações psicológicas que, inclusive, podem ser somatizadas. Por ser corporificada a inteligência humana está submetida às flutuações químicas do nosso corpo, flutuações, de resto, que têm propósitos. A nossa máquina química interfere e, por vezes, supera a máquina raciocinante diante de urgências, pois as emoções são uma máquina química de ajustamento de nossas prioridades existenciais.

Para compensar as exigências corporais, as habilidades humanas ganham em expansividade, pois, enquanto a IA só responde aos dados disponíveis, os humanos têm a capacidade de imaginar, antecipar, sentir e julgar situações de mudança. Nós humanos somos providos de sentimentos e afetos, inclusive tristes, que nem nós mesmos capazes de compreender. Nas sábias palavras de Fernando Pessoa, “amamos sempre no que temos o que não temos quando amamos”, fonte primeira de nossas contradições. Somos dotados de senso comum, pois vivemos em um mundo de experiências factuais que muitas vezes nos dão força e outras nos fragilizam.

Em função de tudo isso, não fazem sentido as comparações apressadas entre a inteligência humana e a IA, nem mesmo no ponto em que nos encontramos da IA generativa que aprendeu a conversar conosco. Mas igualmente não faz sentido afirmar que a IA não tem inteligência pelo simples fato de que suas modalidades de inteligência são distintas da nossa. Embora distintas, a IA opera eficazmente com o raciocínio

dedutivo, uma forma de inteligência que também é processada por mentes humanas, com a diferença de que a IA é capaz de lidar com um grande número de dados com os quais as redes neurais artificiais aprenderam a lidar.

Na verdade, do confronto entre IA e inteligência humana, origina-se um paradoxo: o que é difícil para o humano, a IA faz, o que é difícil para a IA, o humano faz. Na realidade são dois tipos distintos de inteligência e, por serem distintos, por mais que a inteligência humana possa ou não nos orgulhar, isso não permite que se negue a potencialidade cognitiva e a expressão da inteligência na IA sem negar que há características e propriedades que funcionam como entraves para que ambas possam se igualar.

O que se pode concluir com Lecun (2018, p. 133) é que os sistemas de IA irão amplificar a inteligência humana do mesmo modo que as máquinas mecânicas foram uma amplificação da força física. Não serão uma substituição. Em suma, devemos prever um futuro em que sistemas computacionais terão habilidades complementares às humanas, funcionando muito mais como uma inteligência aumentada na união de dois tipos de inteligência distintos.

7. O RECRUDESCIMENTO DOS PRINCÍPIOS ILUMINISTAS

Desafortunadamente, pouco efeito podem ter os argumentos a favor do tipo de inteligência de que a IA está dotada frente à exacerbação anacrônica a que estamos assistindo dos princípios iluministas como capa de proteção da exclusividade da inteligência humana contra quaisquer modalidades de IA.

Tenho repetido com certa insistência que o advento da IA trouxe consigo um infeliz revigoramento de velhos e cínicos valores iluministas (ver Sloterdijk, ...) , agora anacronicamente tidos como o supassumo do humano. A filosofia do século XX demoliu esses valores. Essa demolição teve continuidade no pós-humano crítico e na teoria do ator-rede entre outras, assim como na necessidade imperiosa de reconsideração da ontologia do humano frente à sua solidariedade com outras formas de inteligência da

biosfera. De onde vem tanto temor e protecionismo contra formas de inteligência não-orgânicas, processadas no interior de tecnologias inteligentes?

Se continuarmos a pensar a relação entre humano e IA, de modo opositivo, a partir de valores herdados do Iluminismo (sujeito, individualismo, intencionalidade, livre-arbítrio etc.), ou seja, valores que engordam o egocentrismo e que foram demolidos pela filosofia da desconfiança (Foucault, Deleuze, Derrida, Lyotard, que infelizmente tão poucos leram), nossa compreensão da IA ficará presa ao imbróglio sensacionalista (para o bem da mídia) em que nos encontramos.

Levar em consideração as modalidades de inteligência operadas pela IA não significa ocultar ou negligenciar os imensos desafios que a IA está trazendo para tudo o que diz respeito ao humano. O caráter fundamental das tecnologias, desde sempre, encontra-se na ambivalência. Se, de um lado, a IA e suas aplicações trazem, de fato, contribuições inestimáveis para a gestão da vida humana, de outro lado, não devem ser minimizadas as contradições e dilemas que ela traz e que têm sido objeto de discussões éticas levadas a efeito no campo das regulamentações cuja urgência não pode ser procrastinada. Contudo, contradições e dilemas devem ser enfrentados cara a cara. Disfarces e dissimulações proferidos em nome de um protecionismo baseado em concepções obsoletas do humano só colaboram para o afastamento dos verdadeiros problemas.

REFERÊNCIAS

BOSTROM, Nick. Entrevista concedida a Martin Ford. In Ford, Martin (ed.). **Architects of intelligence**. The truth about AI from the people building it. Birmingham: Packt Publishing, p. 97-116, 2018.

BRIDLE, James. **Maneiras de ser**. A busca por uma inteligência planetária. Trad. Daniel Galera. São Paulo: Todavia, 2023.

BRYANT, Levi; SRNICEK, Nick; HARMAN, Graham. Toward a Speculative Philosophy. In **The Speculative Turn: Continental Materialism and Realism**, Levi Bryant, Nick Srnicek, and Graham Harman (eds.). Melbourne: re.press, 2011.

CRUTZEN, Paul; STOERMER, Eugene F. The “anthropocene”. **IGBP Newsletter**, n. 41, p. 17-18, May 2000. Disponível em:

<http://www.igbp.net/download/18.316f18321323470177580001401/NL41.pdf>. Acesso: 10 ab, 2022.

GRAHAM, Harman. **Quentin Meillassoux**. Philosophy in the making. Edinburgh University Press. 2011.

GRUSIN, Richard. Introduction. In Grusin, R. (ed.). **The nonhuman turn**. Minneapolis: The University of Minnesota Press, p. vii-xxix, 2015.

HARAWAY, Donna. Manifesto ciborgue: ciência, tecnologia e feminismo-socialista no final do século XX. In Silva, Tomaz Tadeu (org, trad.). **Antropologia do ciborgue**. Belo Horizonte: Autêntica, p. 37-130, (1985) 2000.

KOLLER, Daphne. Entrevista concedida a Martin Ford. In Ford, Martin (ed.) **Architects of intelligence**. The truth about AI from the people building it. Birmingham: Packt Publishing, p. 387-403, 2018.

LATOURE, Bruno. **Jamais fomos modernos**. Ensaio de antropologia simétrica. Trad. de Carlos Irineu da Costa. Rio de Janeiro: Ed. 34, 1994a.

_____. On technical mediation: Philosophy, sociology, genealogy. **Common Knowledge**, 3 (2), p. 29-64, 1994b.

_____. On actor-network theory: a few clarifications. *Soziale Welt*, p. 369-381, 1996.

_____. On recalling ANT. In Law, J; Hassard, J. (eds.) *Actor-network theory and after*. Oxford: Blackwell Publishers, p. 15-26, 1998.

LECUN, Yann. Entrevista concedida a Martin Ford. In Ford, Martin (ed.). **Architects of intelligence**. The truth about AI from the people building it. Birmingham: Packt Publishing, p. 119-142, 2018.

LEMOIS, André. **A comunicação das coisas**. Teoria ator-rede e cibercultura. São Paulo: Annablume, 2013.

MANCUSO, Stefano. **A revolução das plantas**. São Paulo: Ubu, 2019.

RUS, Daniela. Entrevista concedida a Martin Ford. In Ford, Martin (ed.). **Architects of intelligence**. The truth about AI from the people building it. Birmingham: Packt Publishing, p. 253-268, 2018.

RUSSEL, Stuart J. Entrevista concedida a Martin Ford. In Ford, Martin (ed.) **Architects of intelligence**. The truth about AI from the people building it. Birmingham: Packt Publishing, p. 39-68, 2018.

SANTAELLA, Lucia. Pressupostos teóricos da cognição. In Santaella, Lucia. **Navegar no ciberespaço**. O perfil cognitivo do leitor imersivo. São Paulo: Paulus, p. 73-92, 2004.

_____. O pós-humano na convivência com a internet das coisas. In Santaella, Lucia. **Comunicação ubíqua**. Repercussões na cultura e na educação. São Paulo: Paulus, p. 23-38, 2013.

_____. **Neo humano**. A sétima revolução cognitiva do Sapiens. São Paulo: Paulus, 2022.

_____. **A inteligência artificial é inteligente?** São Paulo: Almedina. Edições 70, 2023.

SANTAELLA, Lucia e Lemos, Renata. **Redes sociais digitais**. A cognição conectiva do Twitter. São Paulo: Paulus, 2010.

TEIXEIRA, João de Fernandes. **Mentes e máquinas**. Uma introdução à ciência cognitiva. Porto Alegre: Artes Médicas, 1998.

TENENBAUM, Joshua. Entrevista concedida a Martin Ford. In Ford, Martin (ed.). **Architects of intelligence**. The truth about AI from the people building it. Birmingham: Packt Publishing, p. 473-491, 2018.