

UNIVERSIDADE ESTADUAL DE CAMPINAS
SISTEMA DE BIBLIOTECAS DA UNICAMP
REPOSITÓRIO DA PRODUÇÃO CIENTÍFICA E INTELLECTUAL DA UNICAMP

Versão do arquivo anexado / Version of attached file:

Versão do Editor / Published Version

Mais informações no site da editora / Further information on publisher's website:

<https://periodicos.uninove.br/dialogia/article/view/23906>

DOI: 10.5585/44.2023.23906

Direitos autorais / Publisher's copyright statement:

©2023 by Universidade Nove de Julho. All rights reserved.

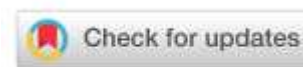
DIRETORIA DE TRATAMENTO DA INFORMAÇÃO

Cidade Universitária Zeferino Vaz Barão Geraldo

CEP 13083-970 – Campinas SP

Fone: (19) 3521-6493

<http://www.repositorio.unicamp.br>



Inteligência artificial, pós-humanismo e Educação: entre o simulacro e a assemblagem¹

Artificial intelligence, post-humanism and education: between simulacrum and assembly

 **Marcelo El Khouri Buzato**

Livre-Docente/ Doutor

Universidade Estadual de Campinas – UNICAMP.

Campinas, SP – Brasil.

mbuzato@unicamp.br

Resumo: Argumenta-se contra a personalização de agentes de inteligência artificial (IA) generativas baseadas em modelos de linguagem natural tais como ChatGPT e Bard. Propõe-se que projetar “humanidade” nesse tipo de artefato é uma forma de apego antropocêntrico improdutivo educacionalmente; sugere-se, como alternativa, o modelo da assemblagem cognitiva (novo inconsciente) de Katherine Hayles, no qual cognições técnicas (não-conscientes) produzem modos de atenção-reconhecimento que têm agência moral e política, mas não se confundem com consciências humanas. Explica-se a projeção de uma teoria da mente no artefato a partir do arcabouço da pós-fenomenologia e adverte-se que nem demonizar a inteligência artificial, nem tratar esses simulacros como inteligências de facto, “empodera” o humano, mas apenas reencena versões progressivamente mais reducionistas do humano idealizado do humanismo. Urge que se utilize de forma mais transparente as articulações entre consciência e modos não-humanos de atenção-reconhecimento para que tanto estudantes IAs não se tornem cada vez menos “inteligentes”.

Palavras chave: inteligência artificial; pós-humanismo; cognição distribuída; consciência; ChatGPT.

Abstract: Na argument is put forward against personifying generative artificial intelligence (AI) agents based on natural language models such as ChatGPT and Bard. It is proposed that projecting “humanity” onto this type of artifact is a form of educationally unproductive anthropocentric attachment; Katherine Hayles' model of cognitive assembly (new unconscious) is proposed as an alternative, in which technical (non-conscious) cognitions produce modes of attention-recognition that have moral and political agency, but are not to be confused with human consciousness. The projection of a theory of mind onto the artifact is explained through a post-phenomenological framework and it is warned that neither demonizing artificial intelligence, nor treating these simulacra as de facto intelligence “empowers” the human, but only reenacts progressively more reductionist versions of the idealized humanista human. It is urgent to use in a more transparent way the articulations between consciousness and non-human modes of awareness so that neither students nor AI become less and less “intelligent”.

Keywords: artificial intelligence; posthumanism; distributed cognition; conscience; ChatGPT.

Cite como

(ABNT NBR 6023:2018)

BUZATO, Marcelo El Khouri. Inteligência artificial, pós-humanismo e Educação: entre o simulacro e a assemblagem. *Dialogia*, São Paulo, n. 44, p. 1-20, e23906, jan./abr. 2023. Disponível em: <https://doi.org/10.5585/44.2023.23906>.

American Psychological Association (APA)

Buzato, M. E. K.. (2023, jan./abr.). Inteligência artificial, pós-humanismo e Educação: entre o simulacro e a assemblagem. *Dialogia*, São Paulo, 44, p. 1-20, e23906. <https://doi.org/10.5585/44.2023.23906>.

¹ Trabalho apresentado no evento: CIET: EnPET – ESUD: CIESUD 2022.

1 Introdução

Este trabalho é um resultado parcial do projeto “Dadificação e linguagem do ciborgue: rumo a uma Linguística Aplicada Pós-humanista” do CNPq², cujo objetivo principal é produzir subsídios teóricos e empíricos para a prática de uma Linguística Aplicada Pós-humanista focalizada no uso das TDIC no Brasil. Centrais nesse contexto de pesquisa são agentes cognitivos artificiais, chamados genericamente de IAs (inteligências artificiais), cada vez mais presentes em discursos e práticas educacionais no país.

Com a naturalização e a pervasividade das TDIC nas sociedades urbanas mais abastadas, a atividade humana é continuamente saturada por essas agências cognitivas não conscientes. Assim, seria esperado que tivéssemos mais clareza de nossa integração ecológica, dos pontos de vista cognitivo e vital (BENNETT, 2010), com artefatos técnicos do tipo *smart*, numa medida mais próxima da do seu sucesso comercial. Tal sucesso é fruto, justamente, da mão-dupla em que se dá essa integração cognitiva-comportamental, isto é do trabalho cognitivo-semiótico feito pelos humanos que interagem com elas ou fornecem dados linguístico-discursivos em massa (e assim ajudam em seu desenvolvimento técnico). O que se vê, no entanto, é uma dissolução ideológica do componente não humano dessa ecologia no *background* da nossa experiência cultural, o que já transparece na naturalidade com que reimportamos dos produtores de TDIC certas metáforas reducionistas e antropomorfizantes com palavras antes emprestadas por eles, , tais como “linguagens” de computador (para códigos de automação de máquinas) e “inteligência” artificial (para processamento automatizado de modelos simbólicos ou probabilísticos adquiridos).

Meu propósito, neste trabalho, é propor uma análise crítica dos modos de relação entre humanos e IAs para uma prática educacional alinhada a um pós-humanismo (crítico) e, conseqüentemente, oposta tanto a determinismos instrumentais sobre IA quanto a sua demonização baseada em apegos humanistas ingênuos. Essa crítica é pautada pelo imperativo de resistir a propostas educacionais tecnocêntricas, de inspiração neoliberal, para a IA, pautadas pela expropriação do trabalho simbólico-cognitivo-afetivo humano por um capitalismo cognitivo (MOULIER BOUTANG, 2011) predatório e antiético (BUZATO, 2022).

Espero, com isso, contribuir para com um debate necessário e urgente sobre políticas e métodos de adoção de IA no contexto da Educação (e em geral) no Brasil sem, contudo, ceder nem à utopia da IA autônoma, nem à tecnofobia do humanismo ingênuo.

² Número de processo 312906/2020-0, chamada CNPq Nº 09/2020 - Bolsas de Produtividade em Pesquisa - PQ



2 Educação e pós-humanismo (crítico e cultural)

Pós-humanismo é um movimento intelectual e tecnocientífico ramificado entre Ciências Humanas e Exatas, nascido das contradições e idealizações que caracterizaram o humanismo histórico. Trata-se de repensar a unidade básica de estudo das humanidades (BRAIDOTTI, 2013), assim como as Humanidades enquanto campo epistemológico e praxiológico em si, a partir da suspensão analítica de certas dicotomias fundantes do pensamento moderno: humano/não humano, mente/corpo, organismo/ambiente, sujeito/objeto entre outras.

Essa suspensão analítica nos abre a consciência sobre o papel dos humanos como força da natureza, de que tratam os estudos do Antropoceno³, por exemplo, e nos conscientiza do antropocentrismo que nos faz pensar-nos "naturalmente no centro das coisas" e "inteiramente distinto(s) de animais, máquinas e outras entidades não humanas" (BADMINGTON, 2004, p. 1345, minha tradução). Temos um senso historicamente arraigado de que "o dom da razão" (e da linguagem) nos torna senhores da natureza, o que remonta aos edifícios filosóficos de Locke, Kant e de outros pilares do Humanismo Liberal. Aí se nutre o sentimento dúbio de admiração e temor que nos assalta frente a performances de IAs sofisticadas – como o afamado Chat GPT, ou então as *deep fakes* que nos tornam impotentes para utilizar a boa e velha crença humana no "ver para crer". Apavora-nos, mas também encanta, saber que o Google nos conhece(ria) melhor que nós mesmos (HARARI, 2016); ao mesmo tempo, nos reconforta nossa arrogância antropocêntrica: se chegamos ao Google (ou ao aquecimento global), podemos chegar ao anti-Google (ou à geoengenharia), porque o humano é "absolutamente conhecido e cognoscível por 'ele mesmo'" (BADMINGTON, 2004, p. 1345, minha tradução).

Há diversos pós-humanismos, em verdade (FERRANDO, 2013). Há um que pretende construir uma crítica abrangente do humanismo como base ideológica da opressão (racial, de gênero, de classe, da natureza etc.) (HERBRECHTER, 2018). Há, em oposição exata, um pós-humanismo transumanista (BOSTROM, 2003), na verdade, um ultra-humanismo, que visa a maximizar a supremacia do humano (branco ocidental) por meio da intervenção biotecnológica na espécie e da tradução radical de corpos humanos em padrões informacionais "imortais", futuros discípulos de uma IA consciente e benévola na "Era da Singularidade" (KURZWEIL, 2005).

Há, finalmente, um pós-humanismo derivado de estudos interdisciplinares sobre as figurações e ansiedades do humano na cultura digital, que tematiza a subjetividade humana no trânsito entre informação e materialidade. Essa vertente, à qual me filio neste trabalho, se baseia em uma profunda pesquisa sobre os fundamentos e implicações da epistemologia cibernética e das

³ Acerca da relação entre Antropoceno e Pós-humano, ver (BUZATO, 2020).

ciências da computação sobre vida e cognição humanas (HAYLES, 1999; 2007). Aqui o pós-humano é uma condição existencial presente e cotidiana pela qual somos capturados pela dialética corpo/matéria *versus* informação/virtualidade. É esse, particularmente, o pós-humanismo que nos dá a chave para formas de relação entre humanos e IAs que interessam a uma Educação reformadora contemporânea, que nos convida a “ir além das formas humanistas tradicionais de pensar sobre o (ser humano como) agente individual autônomo e obstinado, a fim de tratar o próprio humano como uma assemblagem, coevoluindo com outras formas de vida, enredado com o ambiente e com a tecnologia” (NAYAR, 2014, p. 4, minha tradução). É daí que devemos extrair novas concepções de cognição, alteridade e ética, desapegadas do humanismo liberal, que celebrem nossa história coevolutiva com as máquinas, mas entendam que o sentido da existência humana está ligado à finitude e à dignidade de todo vivente (HAYLES, 1999).

3 Humanos e outros cognizadores na assemblagem pós-humana

Boas IAs gerativas, como os notórios ChatGPT e DALL-E⁴, são fascinantes, do ponto de vista de uma crítica pós-humanista, menos pelas suas proezas, do que pelas manchetes que anunciam suas competências em veículos de mídia interessados no que pode resultar dos bilhões de dólares injetados em tais projetos. Por exemplo, em: “ChatGPT ‘passa’ em prova de MBA, ‘OAB’ dos EUA e de capacitação médica”⁵, vê-se projetar a IA como um “eu (cogito cartesiano)” dotado de um predicado humano que até então se reservaria, no viés humanista, a um estudante consciente, sujeito autônomo da razão. Em: “Inteligência artificial DALL-E mostra como seria uma cidade em Marte”⁶, celebra-se o bom o velho hábito colonialista do ocidente liberal de imaginar (e invadir) algum “novo mundo” em que possa projetar sua cosmovisão e hierarquia de valores sobre o (até então) exótico e o selvagem. Já em: “Perguntamos ao ChatGPT: Como a inteligência artificial muda o ensino?”⁷, revela-se o determinismo tecnológico incrustado autodescrições da modernidade humanista, que atribui às suas ferramentas tecnológicas a força causal da história, presente e indicativa – como no modo verbal empregado.

É no mínimo irônico que os humanos cedam suas imagens mentais traduzidas em palavras e seus textos produzidos arduamente para processos seletivos, normalmente sem sabê-lo e sem

4 ChatGPT - Chat Generative Pre-trained Transformer e DALL-E são programas baseados em deep learning, criados pela empresa Open IA, baseada em San Francisco, EUA, em 2022 e 2021, respectivamente.

5 REDAÇÃO. ChatGPT ‘passa’ em prova de MBA, ‘OAB’ dos EUA e de capacitação médica - Estadão. Estadão (online), 24 jan. 2023. Disponível em: <<https://www.estadao.com.br/link/empresas/chatgpt-passa-em-prova-de-mba-oab-dos-eua-e-de-capacitacao-medica/>>. Acesso em: 2 fev. 2023.

6 BRAGADO, Louise. Inteligência artificial DALL-E mostra como seria uma cidade em Marte; veja as imagens. Época NEGÓCIOS, Rio de Janeiro, 25 out. 2022. Tecnologia. Disponível em: <<https://epocanegocios.globo.com/tecnologia/noticia/2022/10/inteligencia-artificial-dall-e-mostra-como-seria-uma-cidade-em-marte-veja-as-imagens.ghtml>>. Acesso em: 2 fev. 2023.

7 LEON FERRARI. Perguntamos ao ChatGPT: Como a inteligência artificial muda o ensino? Tilt - UOL, São Paulo, 1 fev. 2023. Inteligência artificial. Disponível em: <<https://www.uol.com.br/tilt/ultimas-noticias/estado/2023/02/01/perguntao-ao-chatgpt-como-a-inteligencia-artificial-muda-o-ensino.htm>>. Acesso em: 2 fev. 2023.



receber pagamento por isso, a sistemas distribuídos de modelização estatística e processamento simbólico-formal só para, em seguida, poderem atribuir o resultado desse empréstimo a alguma suposta consciência, quase humana, do artefato. Mais do que irônica, essa visão é míope, porque também nos impede de perceber o quanto somos, em alguma medida, seres alcançáveis pela ontologia imaterial e epistemologia pragmática da cibernética.

Hayles (2017) propõe um alargamento do conceito de “cognição” nas Ciências Humanas como forma de superar esse tipo de projeção humanista/antropocêntrica alienante. Trata-se de pensar em cognição como efeito de uma assemblagem de processos materiais, orgânicos e não orgânicos, fluxos informacionais, signos culturais e outras produções coletivas de toda uma assemblagem de interagentes humanos e não humanos conectados, em que se processam, contínua e extensivamente no espaço e no tempo, os mais diversos modos de atenção-reconhecimento (*modes of awareness*)⁸, entre os quais está, não exclusivamente, a consciência, restrita a certos organismos biológicos superiores.

Com base na biologia cognitiva (KOVÁČ, 2000)⁹, na teoria da cognição distribuída (HUTCHINS, 2000)¹⁰ e na cibernética de segunda ordem (BATESON, 1979)¹¹, entre outros aportes, Hayles (2016) propõe que se caracterize a cognição pós-humana, para fins analíticos, como uma pirâmide de três níveis: na base, processos materiais não cognitivos; acima, em menor volume, cognições não conscientes (técnicas e biológicas); e, no topo, em pequeno volume, modos de atenção-reconhecimento emergentes da atividade nos níveis inferiores, entre os quais as cognições dos autômatos que se adaptam a seus ambientes físicos (como fazem robôs com reconhecimento visual) ou lógico-simbólicos (como fazem IAs que respondem a perguntas), além de emoções e ideias humanas e de alguns mamíferos superiores, basicamente.

Tal formulação alia o conceito de *assemblage* (agenciamento) de Deleuze e Guattari (1995) à definição sistêmica de cognição como “um processo de interpretação de informações em contextos que o conectam com o significado¹²”. Assim, busca expandir “a visão tradicional de cognição como

⁸ O termo *modes of awareness* é cunhado por Hayles (2017) justamente para contrapor-se à posição tradicional humanista de que a cognição seria uma função da consciência (humana), o que excluiria cognizadores não-humanos em geral. Traduzo *awareness* como um composto em português para explicitar os dois sentidos do termo conjugados em inglês, mais claramente atribuíveis, por exemplo, a um sistema de reconhecimento de faces, um girassol ou uma porta automática de elevador.

⁹ A biologia cognitiva explica que todo indivíduo vivo depende do conhecimento incorporado mínimo relevante para estabelecer alguma correspondência (informacional-semiótica) com seu ambiente, ainda seja apenas a correspondência entre as moléculas do ambiente e as de seu corpo unicelular. Assim, o organismo é um sistema interpretante que consegue “resolver problemas”, como capturar alimento ou se afastar de uma toxina, sem necessidade de consciência, exatamente como numa IA.

¹⁰ Essa abordagem das Ciências Cognitivas explica a inteligência percebida em indivíduos por referência a um nível sistêmico indivíduo-artefatos-ambiente. As funções cognitivas dentro de uma cultura (e de um comportamento inteligente qualquer nela) são distribuídas e propagadas pelo contato entre seus membros, e desses com seus artefatos e ambiente, conforme possibilidades agenciais e determinações estruturais de cada agente.

¹¹ Na cibernética de Bateson, “qualquer conjunto contínuo de eventos e objetos que tenha a complexidade apropriada de circuitos causais e as relações de energia apropriadas certamente mostrará características mentais”, mas “nenhuma parte desse sistema interativo internamente pode ter controle unilateral sobre o restante ou sobre qualquer outra parte” (BATESON, 1971, p. 4, minha tradução).

¹² Em cibernética, o sentido da informação é dado pelo processo (material, biológico ou não) que a interpreta e/ou no entrecruzamento desses processos (BATESON, 1987).

pensamento humano para a de processos que ocorrem em vários níveis e locais dentro da vida biológica e sistemas técnicos” (HAYLES, 2016, p. 32, minha tradução). O “normal” da cognição, portanto, não seria mais relação sujeito consciente-objeto discreto), mas os processos interpretativos que se dão entre partes biológicas e não-biológicas de um sistema híbrido distribuído.

Perceba-se como essa definição se assenta muito mais convincentemente do que as manchetes anteriormente mencionadas ao verdadeiro funcionamento de IAs conexionistas: elas modelam e mimetizam esses comportamentos sem dispor de uma consciência. Quem interpreta o que uma IA faz como tendo sentidos humanos somos nós, seus interagentes conscientes, porque isso nos convém (BUZATO, 2010).

Mais interessante ainda é entender que esse modelo é escalável e monádico, isto é, aplica-se a um estudante interagindo com Alexa, mas também a *fake news* que se alastram em “bolhas” de rede social, ou a *smart cities* em que integram comportamentos cognitivos de populações humanas e de sistemas de controle de trânsito, energia e poluição ambiental, perfazendo um subsistema que troca informação-entropia com o sistema climático planetário. Vai por água abaixo, nesse modelo, a fantasia das IAs como sujeitos humanos atomizados artificiais, inspirada no humanismo, e abrem-se novas perspectivas para pensarmos, de forma menos preconceituosa, em questões fundamentais para o ensino: ética, trabalho, autoria, justiça, cidadania, veredicação etc.

4 Modos de relação humano-tecnologia-mundo e a recepção social das IAs

Se queremos evitar tanto o determinismo tecnológico alienante quanto as projeções antropocêntricas sobre as IAs, que só servem para obscurecer a compreensão pública sobre essas tecnologias, isso sem jamais negar o grande mérito que existe em empreendimentos tecnocientíficos que visam a articular processos cognitivos humanos e não humanos na assemblagem de modo a facilitar e dinamizar tarefas intelectuais úteis às comunidades viventes, precisamos considerar as práticas e fenomenologia dessa relação humano-não humano. É importante atentar para como a experiência humana estrutura as relações com IAs porque, ao mesmo tempo em que a antropomorfização das IAs deriva do *hype* e das metáforas promovidas pelos seus produtores comerciais, o modo como interagimos cumulativamente com nossas tecnologias influencia, cumulativamente, as competências e estruturas físicas e cognitivas humanas, até mesmo no sentido evolucionário, o que Mitchell, (2002) e Hayles (2017) chamam de *technogenesis*.

De certo modo, IAs generativas de grande capacidade baseadas em processamento de linguagem natural (ou de signos culturais humanos de qualquer tipo), como ChatGPT e DALL-E,

tendem a deslocar a relação hermenêutica (sujeito-objeto) que tradicionalmente outorgamos a programas de computador para o território da alteridade (eu-outro). Isso é importante, em Educação, porque a construção da identidade-alteridade humana (no humanismo) é mediada por práticas educacionais, sobretudo letramentos, ligadas a mídias específicas, o que Sloterdijk (2016) chama de antropotécnica.

Encontramos na pós-fenomenologia (ROSENBERGER; VERBEEK, 2015) um caminho interessante de pesquisa e análise dos modos de relação entre humanos e IAs, sendo que parte de nosso grupo de pesquisa se dedica, no momento, a articular esses modos de relação com concepções de IA educacional e políticas de letramento de IA no mundo e no Brasil (GONSALES, a sair).

O conceito de modos de relação Humano –Tecnologia –Mundo em pós-fenomenologia foi desenvolvido a partir da distinção, por Heidegger (2007), entre artefato “presente-à-mão” (*Vorhandenheit*), isto é, o objeto tecnológico em si, passível de descrição teórica, e artefato “pronto-na-mão” (*Zuhandenheit*), isto é, um objeto que se anuncia como um campo de ação, o que torna a tecnologia uma estrutura utilizador-artefato-mundo. A pós-fenomenologia estipula diversos modos, acessíveis pela observação empírica, pelos quais se estrutura a relação “Eu–Tecnologia–Mundo”. Na relação de incorporação, o artefato presente-à-mão recede ao *background* da experiência, passando a contar como parte despercebida do corpo (ou da mente) do utilizador – é o caso de um par de óculos bem limpos, confortáveis e adequados a nossa patologia ocular específica. Também pode ser uma relação hermenêutica, pela qual tecnologia e mundo se fundem – por exemplo, um médico a “ver uma doença” numa imagem de ressonância magnética, ou alguém tomando o que ChatGPT diz sobre o mundo como critério do que está de fato no mundo.

Um terceiro tipo de relação é a da alteridade. Nesse caso, a estrutura da experiência coloca o mundo no *background*: utilizador e artefato passam a se comportar como parceiros sociais. É esse o caso quando tratamos IAs como “alguém que conversa com a gente”, que nos compreende e se “faz compreender” por nós. O problema, neste caso, é que, para que haja eficácia de uso, o usuário é obrigado a projetar no artefato uma teoria da mente (WITTKOWER, 2022), isto é, algum simulacro de consciência. O próprio design do *software* tenta “azeitar” essa relação, usando vozes humanas sintetizadas (normalmente dóceis e femininas), encenando pequenos rituais sociais tecnicamente desnecessários (cumprimentar, agradecer, despedir-se), utilizando modalizadores ou frases feitas que “disfarçam” sua falta de compreensão (BUZATO, 2010) ou mesmo tentando demonstrar empatia etc.¹³.

¹³ No caso específico da projeção de uma relação de alteridade facilitada pela demonstração de afeto e empatia, nosso grupo vem estudando o funcionamento da Chat Bot *Replika*, afamada por seu uso terapêutico no tratamento da depressão, e frequentemente considerada uma parceira romântica por seus (KOGA, 2022).

Em outras palavras, a assemblagem que permite o funcionamento da IA (IA que, de fato, não pode ser considerada alguém que pensa, sequer pelos parâmetros do notório Teste de Turing) é capturada para funcionar segundo os modos de identidade e alteridade do humano do humanismo, disfarçando, simultaneamente, a inconsciência de sua cognição e o fato de que quem a faz inteligente, na experiência, é nossa capacidade humana de fabular e de adaptar nosso comportamento semiótico a interlocutores menos inteligentes (como crianças pequenas e animais de estimação) que desejamos “hominizar” (SLOTTERDIJK, 2016) (ou domesticar).

5 Um argumento contra a IAs humanistas na Educação do capitalismo cognitivo

À guisa de conclusão, eis meu argumento explícito: creio que a Educação não pode ser vítima do preconceito antropocêntrico que reduz o todo da cognição a apenas um de seus modos de atenção-reconhecimento, redirecionando esse mesmo preconceito, ironicamente, aos próprios artefatos. Que isso acontece está patente no excerto de imprensa a seguir:

O Estadão resolveu questionar o que o próprio ChatGPT "pensa" sobre seus impactos na educação. Ele reconhece que sua introdução pode aumentar casos de plágio, mas destaca que "a prevenção da prática depende da educação dos estudantes sobre a importância da integridade acadêmica e da autoria original", e frisa que embora prime pela objetividade, suas respostas podem refletir "vieses" e "desigualdades" das fontes do banco de dados dele (FERRARI, 2023, s.p.)

Tendo sido programada pelos produtores comerciais (o que é provável), ou apenas modelada estatisticamente a partir de enunciados humanos do senso comum neutralista tecnológico¹⁴ no meio educacional, a resposta é tratada, ainda que com aspas de ressalva, como vinda de uma mente autorreferente. Esse tipo de “enquadramento inicial” tem influenciado sobremaneira a percepção pública sobre o impacto social das IAs e facilitado a fantasia da IA autônoma como produto comercial. A pergunta realmente pertinente talvez fosse: “Como se usa a IA para não mudar o antropocentrismo humanista em Educação em benefício da venda de software antropomorfizado?”.

O senso comum educacional acredita que o benefício da IA esteja em inseri-la na categoria de “ferramenta (hermenêutica)”, em que já estão, grosso modo, as demais tecnologias educacionais (GONSALES, a sair) – em casos mais raros, em escolas “de ponta”, IAs têm sido, ou serão, usadas como soluções de modelagem e geração projetadas por professores e alunos (como já se faz com robótica). Mas o mais provável, sabemos, sobretudo na escola pública, é que IAs continuem sendo, cada vez mais, alimentadas pelos próprios estudantes e professores com seus dados de experiência

¹⁴ Refere-se à posição de que as tecnologias, em si, são neutras, de modo que a beneficência ou maleficência de seus efeitos sociais depende exclusivamente das intenções e capacidades humanas (LATOUR, 1994).

vivida e elaborações não remuneradas da função “consciência humana”, alienadas de seu poder na assemblagem (GONSALES; AMIEL, 2020).

A mesma relação de alteridade com um simulacro de consciência que encanta o jornalista possivelmente apavora professores e pais arraigados à mentalidade humanista radical, ou por temerem perder seus empregos, ou por tornar muito difícil saber o que foi ensinando/aprendido e o que foi gerado por modelagem não-consciente. Do outro lado da moeda, estudantes já começam a supor que é mais *smart* pedir que o ChatGPT faça a tarefa de casa do que “perder tempo” escrevendo ou desenhando sobre o que pode ser requisitado da IA a qualquer momento.

Esses problemas são óbvios, e, nem por isso, menos graves. Contudo, pensar na assemblagem nos indica que é mais *smart* nos ocuparmos menos de encontrar IAs detectoras de plágios de IA para redações de vestibular e mais em do que uma redação (“algorítmica”) de vestibular ainda diz, de fato, sobre “inteligência”; ou investir menos em IAs educacionais fantasiosas e mais em tempo e incentivo de pesquisa, autodesenvolvimento e de autonomia criativa para professores que possam delegar tarefas de “educação bancária” a IAs “honestas” em troca de mais interação humano-humano significativa. Sobretudo, é preciso impedir gestores e pais de, burramente, imporem performances de IAs como *benchmarks* para o trabalho criativo e consciente do professor e do aluno¹⁵, como já se faz com trabalhadores no mercado.

A noção de tecnogênese (MITCHELL, 2002) nos dá certa clareza de que o custo de tratar simulacros como seres humanos conscientes é o da adaptação biológica e sócio-histórica de nossa cognição incorporada aos mecanismos do simulacro que gerencia o ambiente. Como nós humanos também somos capazes de agir assim – afinal quem, numa sociedade humana, não age, ocasionalmente, de modo a simular que entende o que não entende, para não romper o fluxo social? (BUZATO, 2010) –, é fácil prever que tratar IAs como humanos nos “empodera” menos do que nos impele a uma remodelagem contínua de uma versão reducionista de nós mesmos (na medida do reducionismo do modelo computacional); logo, quanto mais “burramente” nos comportamos para sustentar um simulacro de consciência, a longo prazo, mais arriscamos que as próprias IAs fiquem mais “burras” (e, consequentemente, mais vulneráveis ao racismo, à misoginia, ao etnocentrismo, às teorias conspiratórias da moda etc.).

Que não se confunda esta crítica justamente com seu oposto. Nada neste trabalho significa demonizar, nem mesmo desmerecer, a Inteligência Artificial como campo de investigação tecnocientífica e artefato técnico útil às comunidades, tampouco reforçar ou defender a dicotomia humano/não humano, cuja suspensão é justamente a base de qualquer crítica pós-humanista.

¹⁵ Me refiro, aqui, ao relato anedótico de uma pesquisadora de nosso grupo, que atual como professora de redação na rede privada de ensino paulista, sobre escores de correções automatizadas de redações de Enem feitas por IA estarem sendo usadas por pais e gestores como modelo para as correções a serem feitas pelas professoras na escola.

Trata-se, justamente, de argumentar por uma Educação menos antropocêntrica e mais protetora da dignidade, num sentido radicalmente ecológico, de tudo que tenha sensibilidade e imaginação.

Referências

- BADMINGTON, N. Mapping Posthumanism. *Environment and Planning A: Economy and Space*, v. 36, n. 8, p. 1344–1351, ago. 2004. DOI: <https://doi.org/10.1068/a3712>
- BATESON, G. The Cybernetics of “Self”: A Theory of Alcoholism. *Psychiatry*, v. 34, n. 1, p. 1–18, fev. 1971.
- BATESON, G. *Mind and nature: a necessary unity*. 1st ed ed. New York: Dutton, 1979.
- BATESON, G. *Steps to an ecology of mind*. Northvale, N.J: Aronson, 1987.
- BENNETT, J. *Vibrant Matter: A Political Ecology of Things*. Durham: Duke University Press, 2010.
- BOSTROM, N. The transhumanist FAQ: a general introduction the World Transhumanist Association, 2003. Disponível em: <<https://www.nickbostrom.com/views/transhumanist.pdf>>. Acesso em: 17 nov. 2020.
- BRAIDOTTI, R. *The posthuman*. Cambridge, UK ; Malden, MA, USA: Polity Press, 2013.
- BUZATO, M. E. K. Can reading a robot derobotize a reader? *Trabalhos em Linguística Aplicada*, v. 49, n. 2, p. 359–372, dez. 2010. DOI: <http://dx.doi.org/10.1590/S0103-18132010000200004>
- BUZATO, M. E. K. Entre o padrão e a presença: buscando uma heurística para a biopolítica do pós-humano. Em: BUTTURI JÚNIOR, A.; SCARSO, D.; LEME, J. L. C. (Ed.). *Antropoceno, biopolítica e pós-humano*. 1. ed. Campinas, SP: Pontes, 2020. p. 328.
- BUZATO, M. E. K. (How) Can Critical Posthumanism Help to Decolonize Tertiary Education in the South in the Age of Cognitive Capitalism? Em: MAKONI, S. et al. (Ed.). *The languaging of higher education in the global South*. New York: Routledge, 2022. p. 211–227.
- DELEUZE, G.; GUATTARI, F. *Mil platôs: capitalismo e esquizofrenia*. Tradução Célia Pinto Costa. 1. ed. Rio de Janeiro: Editora 34, 1995. v. 1
- FERRANDO, F. Posthumanism, Transhumanism, Antihumanism, Metahumanism, and New Materialisms: Differences and Relations. *Existenz*, v. 8, n. 2, p. 26–32, 2013. Disponível em <<https://existenz.us/volumes/Vol.8-2Ferrando.html>> Acesso em 31 Dez, 2021.
- GONSALES, P. *AI literacies: letramentos digitais de inteligência artificial e a educação na perspectiva pós-humanista*. a sair. Universidade Estadual de Campinas, Campinas, SP, a sair.
- GONSALES, P.; AMIEL, T. Inteligência Artificial, Educação e Infância Educação na contemporaneidade: entre dados e direitos. *Panorama Setorial Cetic* v. 12, n. 3, p. 1–21, 2020. Disponível em <https://cetic.br/media/docs/publicacoes/6/20201110120042/panorama_setorial_ano-xii_n_3_inteligencia_artificial_educacao_infancia.pdf> Acesso em 31 Dez, 2021.
- HARARI, Y. N. *Homo Deus: a brief history of tomorrow*. London: Harvill Secker, 2016.

HAYLES, N. K. Cognitive Assemblages: Technical Agency and Human Interactions. *Critical Inquiry*, v. 43, n. 1, p. 32–55, set. 2016. DOI: <https://doi.org/10.1086/688293>

HAYLES, N. K. *Unthought: the power of the cognitive nonconscious*. Chicago: The University of Chicago Press, 2017.

HEIDEGGER, M. *Ser e tempo*. Trad. Marcia Schuback. Petrópolis: Vozes [u.a.], 2007.

HERBRECHTER, S. Critical posthumanism. Em: BRAIDOTTI, R.; HLAVAJOVA, M. (Ed.). *Posthuman glossary*. Theory. London: Bloomsbury Academic, 2018. p. 94–96.

HUTCHINS, E. *Cognition in the wild*. Nachdr. Cambridge: MIT Press, 2000.

KOGA, J. *Euforia ou Fantasia: Agentes Conversacionais e sua relação com a linguagem*. Campinas, SP, 2022. .

KOVÁČ, L. Fundamental Principles of Cognitive Biology. *Evolution and Cognition*, v. 6, n. 1, p. 51–69, 2000. Disponível em < https://dai.fmph.uniba.sk/courses/CSCTR/materials/CSCTR_03sem_Kovac_2000.pdf> Acesso em 01 Mar., 2019

KURZWEIL, Ray. *The singularity is near: when humans transcend biology*. New York: Viking, 2005.

LATOUR, B. On technical mediation: Philosophy, Sociology, Genealogy. *Common Knowledge*, v. 3, n. 2, p. 29–64, 1994. Disponível em < <http://www.bruno-latour.fr/sites/default/files/54-TECHNIQUES-GB.pdf>> Acesso em 12 Set, 2009

MITCHELL, S. *Technogenesis*. New York: Roc, 2002.

MOULIER BOUTANG, Y. *Cognitive capitalism*. Cambridge: Polity Press, 2011.

NAYAR, P. K. *Posthumanism*. Cambridge: Polity Press, 2014.

ROSENBERGER, R.; VERBEEK, P.-P. (ed.). *Postphenomenological investigations: essays on human-technology relations*. Lanham: Lexington Books, 2015.

SLOTERDIJK, P. Rules for the human park: a Response to Heidegger’s “Letter on ‘Humanism’”. Em: *Not saved: essays after Heidegger*. Malden: Polity, 2016. p. 193–216.

WITTKOWER, D. E. What Is It Like to Be a Bot? Em: WITTKOWER, D. E. (Ed.). *The Oxford Handbook of Philosophy of Technology*. Oxford: OUP, 2022. p. 357–373.