

Inteligência artificial e saúde: ressituar o problema

Artificial intelligence and health: resituating the problem

Inteligencia artificial y salud: reubicar el problema

Fernanda Bruno^{1,a}

bruno.fernanda@gmail.com | <https://orcid.org/0000-0002-5950-7602>

Paula Cardoso Pereira^{2,b}

paula.cpereira@gmail.com | <https://orcid.org/0000-0001-8947-9692>

Paulo Faltay^{2,a}

pfaltay@gmail.com | <https://orcid.org/0000-0002-2467-1357>

¹ Universidade Federal do Rio de Janeiro, Instituto de Psicologia, Professora do programa de Pós-Graduação em Comunicação e Cultura. Rio de Janeiro, RJ, Brasil.

² Universidade Federal do Rio de Janeiro, Escola de Comunicação. Rio de Janeiro, RJ, Brasil.

^a Doutorado em Comunicação e Cultura pela Universidade Federal do Rio de Janeiro.

^b Mestrado em Design Comunicacional pela Universidade de Buenos Aires.

RESUMO

Diante dos avanços recentes da inteligência artificial, a presente nota de conjuntura busca recolocar questões fundamentais que emergem nesse contexto. Deslocando-se tanto das leituras salvacionistas quanto apocalípticas, argumentamos que a perda do privilégio do excepcionalismo humano pode ser uma oportunidade para repensar a inteligência a partir de uma perspectiva relacional e co-produzida entre humanos e outros-que-humanos. Tal perspectiva, no entanto, deve ser acompanhada de um olhar atento às relações de poder que em grande medida definem os destinos da IA. Sobre esse aspecto, apontamos as implicações do modelo epistêmico e de negócios hegemônico da IA, um modelo preditivo-aceleracionista dominado por grandes empresas de tecnologia. Finalmente, destacamos alguns riscos envolvidos na inclusão de máquinas inteligentes no campo da saúde, bem como os perigos da subordinação de valores e direitos públicos a interesses comerciais, o que demanda uma atenção e um cuidado coletivos e permanentes na construção dos arranjos sociotécnicos e políticos de implementação da IA nesse campo.

Palavras-chave: Inteligência artificial; Ontologia relacional; Saúde; Extrativismo de dados.

ABSTRACT

In the face of artificial intelligence recent advances, this note seeks to reassess fundamental questions that emerge in this context. Moving away from both salvationist and apocalyptic readings, we argue that the human exceptionalism privilege loss can be an opportunity to rethink intelligence from a relational and co-produced perspective between humans and other-than-humans. Such an angle, however, must be accompanied by a careful examination of the power relations that largely define the fate of AI. On this aspect, we reflect on the implications of the hegemonic epistemic and business model of AI, a predictive-

accelerationist one dominated by large technology companies. Lastly, we highlight the risks involved in the inclusion of intelligent machines in the fields of health and care, as well as the dangers of subordinating public values and rights to commercial interests, which demands attentive, collective and permanent care in the construction of sociotechnical and political arrangements for the implementation of AI in this field.

Keywords: Artificial intelligence; Relational ontology; Health; Data extractivism.

RESUMEN

Ante los recientes avances de la inteligencia artificial, el presente informe busca plantear cuestiones fundamentales que surgen en este contexto. Alejándose tanto de las lecturas salvacionistas como apocalípticas, argumentamos que la pérdida del privilegio del excepcionalismo humano puede ser una oportunidad para repensar la inteligencia desde una perspectiva relacional y co-producida entre humanos y otros-que-humanos. Sin embargo, tal perspectiva debe ir acompañada de una mirada atenta a las relaciones de poder que en gran medida definen el destino de la IA. En este aspecto, señalamos las implicaciones del modelo epistémico y de negocios hegemónico de la IA, un modelo predictivo-aceleracionista dominado por grandes empresas tecnológicas. Por último, destacamos algunos riesgos de la inclusión de máquinas inteligentes en el campo de la salud, así como los peligros de subordinar valores y derechos públicos a intereses comerciales, lo cual requiere una atención y un cuidado colectivos y permanentes en la construcción de los ensamblajes sociotécnicos y políticos de implementación de la IA en este campo.

Palabras claves: Inteligencia artificial; Ontología relacional; Salud; Extrativismo de datos.

INFORMAÇÕES DO ARTIGO

Contribuição dos autores:

Concepção e desenho do estudo: Fernanda Bruno, Paula Cardoso Pereira, Paulo Faltay.

Aquisição, análise ou interpretação dos dados: Fernanda Bruno, Paulo Faltay, Paula Cardoso Pereira.

Redação do manuscrito: Fernanda Bruno, Paula Cardoso Pereira, Paulo Faltay.

Revisão crítica do conteúdo intelectual: Fernanda Bruno, Paulo Faltay, Paula Cardoso Pereira.

Declaração de conflito de interesses: não há.

Fontes de financiamento: não houve.

Considerações éticas: não há.

Agradecimentos/Contribuições adicionais: não há.

Histórico do artigo: submetido: 19 jun. 2023 | aceito: 19 jun. 2023 | publicado: 30 jun. 2023.

Apresentação anterior: não há.

Licença CC BY-NC atribuição não comercial. Com essa licença é permitido acessar, baixar (*download*), copiar, imprimir, compartilhar, reutilizar e distribuir os artigos, desde que para uso não comercial e com a citação da fonte, conferindo os devidos créditos de autoria e menção à Reciis. Nesses casos, nenhuma permissão é necessária por parte dos autores ou dos editores.

A história da inteligência artificial (IA) sempre foi assombrada por fantasmas de superação, destruição ou salvação do humano. Aparentemente, tais fantasmas pareciam estar adormecidos, uma vez que as implementações da inteligência artificial em diversos domínios nos últimos anos não culminaram nem nas grandes promessas nem nas grandes catástrofes antevistas. Curiosamente, eles voltam a ganhar força com os avanços recentes dos modelos de aprendizagem de máquina e aprendizagem profunda, mais especificamente com os chamados grandes modelos de linguagem e de processamento de linguagem natural, popularizados pelo loquaz lançamento do Chat GPT-3. Desde então, projeções apocalípticas e salvacionistas proliferam nos fios e *feeds* das plataformas e de diversas mídias, anunciando desde a imortalidade até a extinção da humanidade, muitas cinicamente capitalizadas pelos próprios financiadores do setor¹. Retornam também com força os debates sobre as similaridades e diferenças entre a inteligência humana e maquina, por vezes reduzidos a uma competição onde só pode haver um vencedor. Como alerta Haraway, “pagamos caro demais por viver no cronotopo das ameaças e promessas finais” (2004, p. 241, tradução nossa). Para nos deslocarmos desse cronotopo e escaparmos da (falsa) escolha entre dois lados que no fundo são faces da mesma moeda, é necessário ressituar o problema da inteligência artificial no presente. De que modo os últimos acontecimentos no campo da pesquisa, desenvolvimento e discussão em torno da Inteligência Artificial nos interpelam e nos convocam a pensar? Que questões precisam ser colocadas e o que estamos perdendo de vista? Para onde devemos dirigir nossa atenção quando sistemas de inteligência artificial passam a fazer parte do cuidado com a saúde? Tratam-se, obviamente, de questões muito amplas e cujas respostas ultrapassam os limites deste texto. No entanto, elas servem como guia para traçarmos um caminho possível de formulação dos problemas de que devemos nos ocupar.

Um primeiro movimento para nos ressituar diante dos deslocamentos provocados pela chamada “era da inteligência artificial” diz respeito ao problema da perda do excepcionalismo humano, isto é, do lugar privilegiado e central que o homem ocupou na modernidade ocidental. Problema que agora se reapresenta de modo muito concreto com a reconfiguração dos processos de decisão, ação e criação próprios a um modelo de racionalidade no qual os algoritmos possuem lugar central (BRUNO, 2021). Esta pode ser uma ocasião para recolocarmos a questão da inteligência para além da perspectiva da competição ou da comparação. Tal lógica atravessou as últimas décadas de desenvolvimento da IA, alimentando desde confrontos públicos promovidos pelas *big techs* entre jogadores de carne e osso e de silício até os imaginários em torno da singularidade tecnológica (VINGE, 1993; KURZWEIL, 2005), essa espécie de “arrebamento cibernético” (DANOWISKI, VIVEIROS DE CASTRO, 2017, p. 69) que marcaria o ponto de inflexão a partir do qual as capacidades computacionais das máquinas ultrapassariam as competências da massa cinzenta do planeta. E se, ao invés de nutrir os fascínios e temores que a lógica competitiva suscita, aproveitarmos o “abalo existencial” desencadeado pela perda do privilégio ontológico e epistemológico² (BARAD, 2003, 2007) para começar a pensar a inteligência a partir da relacionalidade, coprodução e interpelação mútua entre agentes heterogêneos? Afinal, sabemos que grande parte dos desastres em curso foram produzidos a partir do arrogante pedestal do excepcionalismo humano (ao qual nem todos os humanos tinham acesso, uma vez que o lugar de domínio do humano frente à natureza foi acompanhado de toda sorte de hierarquias e discriminações entre humanos).

Pensar a questão da inteligência nesses termos nos permite reconhecer uma “ecologia cognitiva planetária” (HAYLES, 2017) através da qual processos cognitivos são distribuídos e postos em ação de

1 A manifestação mais contundente dessa posição fica explícita na recente “Declaração sobre os riscos da IA”, assinada por executivos da indústria da IA, pesquisadores e outras figuras proeminentes do campo, incluindo Bill Gates e o CEO da OpenAI, Sam Altman (STATEMENT..., 2023).

2 A física e teórica feminista Karen Barad propõe a noção de ontoepistemologia para descrever o estudo do entrelaçamento de práticas de saberes no ser (Barad, 2007). Para ela, “a separação entre epistemologia e ontologia é o efeito de uma metafísica que pressupõe uma separação intrínseca entre o humano e o não-humano, o sujeito e o objeto, a mente e o corpo, a matéria e o discurso.” (BARAD, 2003, p. 829; tradução nossa).

forma híbrida, interpenetrada e sistêmica por uma miríade de atores e elementos distintos, humanos e outros-que-humanos. Ao invés de atributo restrito a determinados entes (sempre demasiado humanos, masculinos, brancos, ocidentais), a inteligência pode ser entendida, de acordo com Bridle, como “algo que surge a partir de inter-relações, de pensar e agir em conjunto” (2023, p. 89). Ou como propõe Malabou (2018, tradução nossa), “a inteligência não é. Ela age. Não é de modo algum uma substância, algo que possuímos da mesma forma que possuímos órgãos, por exemplo, mas uma dinâmica. A inteligência constrói relações entre as coisas, entre os indivíduos, as ideias, as máquinas...”. Não é novidade a longa história de partilha de nosso pensamento e de nossa cognição com os objetos e aparatos técnicos. A gênese e a topologia do que se define como inteligência humana não podem ser separadas das ferramentas e tecnologias com as quais concebemos, agimos e transformamos o mundo e a nós mesmas. Temos décadas de pesquisa e conhecimento produzidos nesse terreno e as recentes máquinas computacionais nos convocam mais uma vez a recolocar o problema do pensamento e da inteligência humana e de suas conexões com outras formas de inteligência e de agência com as quais agora convivemos. Insistir na oposição, no medo e na competição com as máquinas só irá acelerar a concentração de conhecimento e o poder de transformação coletiva na mão dos poucos – e mesmos – atores que hoje governam a mais nova era da inteligência artificial, deixando a maioria dos humanos e outros-que-humanos à margem desse processo. O que nos tornamos e o que desejamos nos tornar com as máquinas computacionais que passam a coabitar nosso cotidiano no trabalho, na educação, na sociabilidade, na política, nas práticas de saúde etc.? Ou ainda: o que se torna a inteligência e a cognição em contextos de alta interpenetração e recursividade entre agentes humanos e técnicos (HAYLES, 2017)? O estranhamento que a performance das IAs generativas nos causam é, assim, uma ocasião para repensar e refazer coletivamente conexões, modos de vida e mundos com as máquinas, entendidas nem como senhoras nem como servas, mas como máquinas companheiras (BEIGUELMAN, 2023; HARAWAY, 2019).

Recolocar o problema a partir dessa perspectiva é uma aposta em uma ontologia relacional (BARAD, 2007, HARAWAY, 2023). Ou seja, não se trata somente de compreender a inteligência como processo relacional, mas também de entender que a relacionalidade modifica os próprios entes nela envolvidos. A questão não é tanto se o humano se tornará (ou não) obsoleto nesse novo arranjo, mas sobretudo como essa relação maquínica-humana, que tende a ser cada vez mais íntima e complexa, transforma o que somos e o que historicamente se entende como o humano.

Aproveitar as perturbações e incômodos desencadeados pelos últimos desenvolvimentos da IA para recolocar a questão da inteligência fora do excepcionalismo humano e a partir de uma perspectiva coproduzida não pode significar, no entanto, abrir mão de um olhar atento às relações de poder e forças econômicas em jogo nessas novas ecologias cognitivas planetárias que, em grande parte, definem seus arranjos possíveis, sob o risco de cairmos numa espécie de pós-humanismo ingênuo.

Grande parte dos arranjos assumidos pela IA em diversos contextos são definidos por modelos epistêmicos e de negócios hoje hegemônicos. Desse modo, um outro movimento para recolocar o problema deve levar em conta as implicações desses modelos. Uma delas é o “realismo da IA”, isto é, do senso de inevitabilidade que vem acompanhado das retóricas mais carregadas em torno dessas tecnologias (MCQUILLAN, 2022) e invisibilizam suas contingências históricas. Apesar da história de a inteligência artificial ser marcada por uma série de disputas, ramificações e composições, atualmente ela tem sido dominada por apenas uma matriz, materializada em grandes modelos estatísticos capazes de descobrir padrões e correlações em vastos conjuntos de dados com a finalidade de gerar previsões, cuja propriedade se concentra na mão de poucas corporações.

Esse modelo epistêmico e de negócios dominante depende do extrativismo de dados em larga escala (RICAURTE, 2022), da expropriação privada de saberes coletivamente produzidos (BENDER *et al.* 2021),

da mão de obra precarizada em países do Sul global (PERRIGO, 2023) e do alto consumo de energia (STRUBELL *et al.* 2020). Pautado por interesses estritamente comerciais e concentrado em um pequeno número de corporações de tecnologia que vêm redefinindo praticamente sozinhas o destino da IA, esse modelo hegemônico também se manifesta num tipo de performatividade voltada para a aceleração. É um modelo preditivo-aceleracionista, regido por uma temporalidade que aposta na otimização ininterrupta como sinal de eficácia, alheio, por consequência, a outras temporalidades mais estendidas e fundamentais para o debate público e para as deliberações políticas coletivas.

A ampla repercussão que os modelos generativos dessa matriz hegemônica ganhou recentemente, envolta em exercícios de futurologia voltados para perigos futuros em nada plausíveis, desvia o debate e escrutínio públicos sobre os danos passados e presentes desse modelo, como vários pesquisadores da chamada IA Crítica vêm alertando (GEBRU, 2022; BELLI; DA HORA, 2023; WONG, 2023). À diferença das futurologias apocalípticas, esses danos já são reais, produzindo efeitos que atingem principalmente populações historicamente marginalizadas: pretas, periféricas e pobres. Focar nossa lente analítica no presente revela que as aplicações de maior impacto da IA tendem a ser aquelas que permeiam instituições e práticas de maneira muito menos visível e cujos efeitos incidem na vida das pessoas de maneira muito mais corriqueira, seja no uso de sistemas biométricos em espaços públicos e privados ou na automatização de decisões em setores como o trabalho, crédito financeiro, segurança pública, habitação, justiça, educação etc. Trata-se, assim de um momento oportuno para dirigirmos nossa atenção para o uso de IA em campos sensíveis como o da saúde e do cuidado.

Sabemos que as noções de saúde, doença e cuidado, bem como as práticas terapêuticas e técnicas de diagnóstico são constantemente transformadas por tecnologias e processos socioculturais. Da mesma maneira, o modo como se produz e se coleta dados sobre indivíduos e populações incide fortemente no modo como as pessoas, instituições e grupos são compreendidos e compreendem a si mesmos e se comportam (FOUCAULT, 1977; ROSE, 2001).

As aplicações da IA na área da saúde são vastas e em muitos países, sobretudo no Norte global, já vêm sendo utilizadas para aumentar a velocidade e precisão de diagnósticos e a triagem de doenças; apoiar decisões clínicas; prever e acompanhar epidemias e surtos; informar decisões para formulação de políticas públicas e/ou alocação de recursos nos sistemas de saúde; auxiliar o desenvolvimento de pesquisas e medicamentos; fornecer ferramentas de autocuidado e automonitoramento para pacientes (ORGANIZAÇÃO MUNDIAL DE SAÚDE, 2021; DONIA; SHAW, 2021, BRUNO *et al.*, 2021). Tais aplicações tendem a produzir reconfigurações gigantescas em diversos domínios da saúde e do cuidado, seja numa redistribuição de saberes entre humanos e máquinas, nas relações entre médico e paciente, nos problemas que são privilegiados por uma “saúde orientada por dados”, no modelo sujeito-paciente subjacente a essas tecnologias e práticas ou nos modos de produção, uso e gestão dos dados na área da saúde.

Sem deixar de reconhecer os potenciais da IA para a saúde pública e a medicina, um relatório recente da OMS (2021) adverte para os perigos de superestimar seus benefícios e subestimar seus riscos, relacionados majoritariamente a três questões: discriminações reproduzidas e amplificadas por essas tecnologias, potenciais perigos à segurança dos pacientes e do meio ambiente e coleta e uso antiético de dados sensíveis e de saúde.

Casos que materializam esses riscos se avolumam à medida que a IA se dissemina na área de saúde. Um grave viés racial foi relatado por Obermeyer *et al.* (2019), ao analisarem o algoritmo usado para orientar decisões de gestão e administrativas no sistema de saúde dos Estados Unidos. Orientado a otimizar custos, o sistema automatizado presumia incorretamente que pacientes brancos precisavam de maior cuidado que pacientes negros com o mesmo nível de necessidade de atendimento, alocando assim menos da metade dos recursos para pacientes negros do que para brancos, segundo estimativa dos pesquisadores. Outro estudo

concluiu que ferramentas de IA treinadas para a detecção de câncer de pele têm dificuldade de diagnosticar pessoas não brancas (LASHBROOK, 2018).

Investigações apontam ainda para manejos indevidos dos dados e sistemas que produzem diagnósticos não solicitados, arbitrários e potencialmente discriminatórios, com enorme capacidade de reforçar condições prévias de vulnerabilidade (VENTURINI, 2019). São exemplos os casos de modelos usados para prever episódios psicóticos por meio da análise de padrões de fala (BEDI *et al.*, 2015); os casos da ferramenta testada para identificar indivíduos que desconheciam estar com tuberculose e torná-los candidatos ao protocolo de profilaxia pré-exposição ao HIV por supostamente terem maior risco de infecção (MARCUS *et al.*, 2019); o acordo entre a província de Salta, na Argentina, e a Microsoft para o desenvolvimento de um sistema automatizado capaz de avaliar a probabilidade de gravidez em adolescentes em populações vulneráveis (PEÑA; VARON, 2019); e a controversa parceria entre a Google DeepMind e o sistema nacional de saúde do Reino Unido, que utilizou dados de 1.6 milhão de pacientes para tratamento de doenças renais sem o conhecimento e a anuência dos titulares.

Esses problemas revelam também os impactos da crescente presença das *big techs* em áreas fundamentais para o bem-estar social, especialmente quando isso ocorre às custas de investimentos essenciais para a implementação de políticas públicas universais e de garantia a direitos fundamentais. Tradicionalmente pouco interessadas no campo, nos últimos anos essas empresas passaram a investir em pesquisas na saúde e no setor biomédico, movimento nomeado por Tamar Sharon (2018) de “Googlização da pesquisa em saúde”. Mais do que más práticas no uso de dados sensíveis ou mau funcionamento dos sistemas, os casos manifestam os perigos de delegar soluções da esfera comum a entes privados, subordinando os valores e direitos envolvidos em uma concepção pública e universal de saúde – como inclusão, equidade, transparência e interesse público – aos valores e interesses comerciais de grandes empresas.

A crescente inclusão de máquinas inteligentes nas práticas de saúde, de cuidado ou em outros campos, como observamos anteriormente, modifica não só tais práticas, mas também aquilo que podemos nos tornar com tais máquinas e vice-versa. Por isso, a construção dos arranjos sociotécnicos de convívio com as novas formas de inteligência artificial precisa estar atenta às implicações na experiência e na qualidade do cuidado com a saúde. Além disso, é preciso que as inovações nesse campo estejam a serviço da redução das desigualdades sociais da saúde (Fassin, 2020) e não percam de vista as assimetrias de poder envolvidas, o que passa pela revisão contínua dos impactos de tais artefatos em indivíduos e populações e pela criação mecanismos de responsabilização e controle voltados para o bem coletivo (COSTANZA-CHOCK, 2020). Passa também pelo combate à dependência dos mercados e das *big techs* e pelo fortalecimento de práticas colaborativas baseadas em contextos locais, assim como nas diferentes necessidades das comunidades e em seus contextos históricos (ESCOBAR, 2018). A produção da saúde coletiva é marcada pela atenção continuamente renovada ao tecido heterogêneo de processos sociais, políticos, econômicos e subjetivos que constituem a experiência da saúde, jamais reduzida a uma definição biológica ou médica. Essa atenção é desafiada a incluir as recentes formas de inteligência maquina na teia de coprodução de políticas de saúde, escapando tanto do fascínio pelo solucionismo tecnológico quanto do conforto meramente instrumental.

REFERÊNCIAS

BARAD, Karen. **Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning**. Durham: Duke University Press, 2007.

BARAD, Karen. Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter. **Signs: Journal of Women in Culture and Society**, Chicago, v. 28, n. 3, p. 801-831, 2003.

BEDI, Gillinder; CARRILO, Facundo; CECCHI, Guillermo A.; SLEZAK, Diego Fernández; SIGMAN, Mariano; MOTA, Natália B. *et al.* Automated analysis of free speech predicts psychosis onset in high-risk youths. **NPJ Schizophrenia**, London, v. 26, n. 1, p. 15030, 2015. DOI: <https://doi.org/10.1038/npschz.2015.30>.

BEIGUELMAN, Giselle. Máquinas companheiras. **Morel**, Santo André, n. 7, p. 76-86, 2023.

BELLI, Luca; DA HORA, Nina. ChatGPT: O que anima e o que assusta na nova inteligência artificial. **Folha de S.Paulo**, 20 jan. 2023. Disponível em: <https://www1.folha.uol.com.br/tec/2023/01/chatgpt-o-que-anima-e-o-que-assusta-na-nova-inteligencia-artificial.shtm>. Acesso em: 14 jun. 2023.

BENDER, Emily M.; GEBRU, Timnit; MCMILLAN-MAJOR, Angelina; SHMITCHELL, Shmargaret. On the dangers of stochastic parrots: Can language models be too big?. **Proceedings of the ACM Conference on Fairness, Accountability, and Transparency**, p. 610-623, March 2021. DOI: <https://doi.org/10.1145/3442188.3445922>.

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

BRUNO, Fernanda. Tecnopolítica, racionalidade algorítmica e mundo como laboratório. In: GROHMANN, Rafael. **Os laboratórios do trabalho digital**: entrevistas. São Paulo: Boitempo, 2021.

BRUNO, Fernanda Glória; PEREIRA, Paula Cardoso; BENTES, Anna Carolina Franco; FALTAY, Paulo; ANTOUN, Mariana; COSTA, Debora Dantas Pio da; STRECKER, Helena; ROCHA, Natássia Salgueiro. "Tudo por conta própria": autonomia individual e mediação técnica em aplicativos de autocuidado psicológico. **RECIIS - Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**, v. 15, n. 1, p. 33-54, 2021. DOI: <https://doi.org/10.29397/reciis.v15i1.2205>. Disponível em: <https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/2205>. Acesso em: 21 jun. 2023.

COSTANZA-CHOCK, Sasha. **Design Justice**: Community-led Practices to Build the Worlds We Need. Cambridge, Massachusetts: The MIT Press, 2020.

DANOWSKI, Déborah; VIVEIROS DE CASTRO, Eduardo. **Há mundo por vir? Ensaio sobre os medos e os fins**. 2. ed. (Cultura e Barbárie). Desterro [Florianópolis]: Instituto Socioambiental, 2017.

DONIA, Joseph; SHAW, James A. Co-design and ethical artificial intelligence for health: An agenda for critical research and practice. **Big Data & Society**, v. 8, n. 2, 2021.

ESCOBAR, Arturo. **Designs for the Pluriverse**: Radical Interdependence, Autonomy, and the Making of Worlds. Durham, NC: Duke University Press, 2018.

FOUCAULT, Michel. **O nascimento da clínica**. Rio de Janeiro: Forense Universitária, 1977.

GEBRU, Timnit. Effective Altruism Is Pushing a Dangerous Brand of 'AI Safety'. **Wired**. 30 nov. 2022. Disponível em: <https://www.wired.com/story/effective-altruism-artificial-intelligence-sam-bankman-fried/>. Acesso em: 10 jun. 2023.

HARAWAY, Donna. **Ficar com o problema**: fazer parentes no chthluceno. Trad.: Ana Luíza Braga. São Paulo: n-1 edições, 2023.

HARAWAY, Donna. Manifesto ciborgue: ciência, tecnologia e feminismo-socialista no final do século XX. In: HOLLANDA, Heloisa Buarque de (org.). **Pensamento feminista**: conceitos fundamentais. Rio de Janeiro: Bazar do Tempo, 2019.

HARAWAY, Donna. The Promise of Monsters: A Regenerative Politics for Inappropriate/d Others. In: HARAWAY, Donna. **The Haraway Reader**. New York, London: Routledge, 2004.

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

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

LASHBROOK A. AI-driven dermatology could leave dark-skinned patients behind. **The Atlantic**. 16 ago. 2018. Disponível em: <https://www.theatlantic.com/health/archive/2018/08/machine-learning-dermatology-skin-color/567619/>. Acesso em: 12 jun. 2023.

MALABOU, C. L'intelligence n'est pas. Elle agit. **Le Point**. 13 abr. 2018. Disponível em: https://www.lepoint.fr/sciences-nature/catherine-malabou-l-intelligence-n-est-pas-elle-agit-13-04-2018-2210517_1924.php. Acessado em: 14 jun. 2023.

MARCUS, Julia L.; HURLEY, Leo B.; KRAKOWER, Douglas S.; ALEXEEFF, Stacey; SILVERBERG, Michael J.; VOLK, Jonathan E. Use of electronic health record data and machine learning to identify candidates for HIV pre-exposure prophylaxis: a modelling study. **Lancet HIV**, v. 6, n. 10, e-688-e695, Jul. 2019. DOI: [https://doi.org/10.1016/S2352-3018\(19\)30137-7](https://doi.org/10.1016/S2352-3018(19)30137-7).

MCQUILLAN, Dan. **Resisting AI: An Anti-fascist Approach to Artificial Intelligence**, Bristol, UK: Bristol University Press, 2022.

OBERMEYER, Ziad; POWERS, Brian; VOGELI, Christine; MULLAINATHAN, Sendhil. Dissecting racial bias in an algorithm used to manage the health of populations. **Science**, v. 366, n. 6464, p. 447-453, Oct. 2019. DOI: <https://doi.org/10.1126/science.aax2342>.

ORGANIZAÇÃO MUNDIAL DA SAÚDE – OMS. **Ethics and governance of artificial intelligence for health: WHO guidance**. Geneva: World Health Organization; 2021. Disponível em: <https://www.who.int/publications/item/9789240029200>. Acesso em: 08 jun. 2023.

PEÑA, Paz; VARON, Joana. Decolonising AI: A transfeminist approach to data and social justice. In: **GLOBAL INFORMATION SOCIETY WATCH 2019**. Artificial intelligence: Human rights, social justice and development. [S.l.]: Association for Progressive Communications, 2019. Disponível em: https://giswatch.org/sites/default/files/gisw2019_web_th4.pdf. Acesso em: 12 jun. 2023.

PERRIGO, Billy. OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic. **Time**. 18 jan. 2023. Disponível em: <https://time.com/6247678/openai-chatgpt-kenya-workers/>. Acesso em: 09 jun. 2023.

RICARTE, Paola. Ethics for the majority world: AI and the question of violence at scale. **Media, Culture & Society**, v. 44, n. 4, p. 726-745, 2022.

ROSE, Nikolas. The politics of life itself. **Theory, culture & society**, London, v. 18, n. 6, p. 1-30, 2001.

STATEMENT ON AI RISK. AI experts and public figures express their concern about AI risk. **Center for AI Safety**, San Francisco. Disponível em: <https://www.safe.ai/statement-on-ai-risk#open-letter>. Acesso em: 21 jun. 2023.

SHARON, Tamar. When digital health meets digital capitalism, how many common goods are at stake?. **Big Data and Society**, v. 5, n. 2, 2018.

STRUBELL, Emma; GANESH, Ananya; MCCALLUM, Andrew. Energy and Policy Considerations for Deep Learning in NLP. In: ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 57 th. **Proceedings** [...], Florence, Italy, Association for Computational Linguistics, p. 3645-3650, 2020.

VENTURINI, Jamila. Vigilância, controle social e desigualdade: a tecnologia reforça vulnerabilidades estruturais na América Latina. **Derechos Digitales**, 15 out. 2019. Disponível em: <https://www.derechosdigitales.org/13921/vigilancia-control-social-e-desigualdade-a-tecnologia-reforca-vulnerabilidades-estructurais-na-america-latina/>. Acesso em: 14 jun. 2023.

VINGE, Vernor. **Technological Singularity**, 1993. Disponível em: <https://frc.ri.cmu.edu/~hpm/book98/com.ch1/vinge.singularity.html>. Acesso em: 15 jun. 2023.

WONG, Matteo. AI doomerism is a decoy. **The Atlantic**. 02 jun. 2023. Disponível em: <https://www.theatlantic.com/technology/archive/2023/06/ai-regulation-sam-altman-bill-gates/674278/>. Acesso em: 09 jun. 2023.