

## Territórios do Conhecimento: o Debate Bioético da IA Responsável e Decolonial

### Territories of Knowledge: the Bioethical Debate on Responsible and Decolonial AI

\*Elen Nas<sup>1</sup> 

#### Resumo

O artigo explora os alinhamentos das perspectivas bioéticas e decoloniais com as premissas da inteligência artificial responsável (IAR). Propõe, desse modo, exames atentos sobre os conflitos implícitos nos termos ‘decolonialidade’ e ‘territórios do conhecimento’. Apresenta as aproximações dos debates acerca da biopolítica e necropolítica com a bioética, associando-os aos possíveis impactos da inteligência artificial (IA). Dado o inevitável alcance da IA em todas as esferas da sociedade, o olhar decolonial explicita o quanto a monocultura reforça ideias epistêmicas com potencial opressor sobre minorias e grupos que têm sido alvos desde as práticas coloniais até os dias atuais. Ao apresentar os princípios da IAR, o artigo ressalta o risco de absorver, sem o necessário pensamento crítico, as regras formais importadas do Norte Global como “A” solução para mitigar os possíveis impactos da IA, quando se fazem necessárias ações educativas e de comunicação dentro da cultura, que só serão possíveis partindo-se do entendimento de que a IAR é uma IA Lenta.

**Palavras-Chave:** inteligência artificial, bioética, decolonialidade, IA responsável.

#### Abstract

The article explores the alignment of bioethics and decolonial perspectives with the premises ‘responsible artificial intelligence’ (RAI). It proposes a careful examination of the implicit conflicts in terms such as ‘decoloniality’ and ‘territories of knowledge’. The article presents the similarities between the debates on biopolitics, necropolitics, and bioethics, associating them with the possible impacts of artificial intelligence (AI). Given the inevitable reach of AI in all spheres of society, the decolonial perspective explains how monoculture reinforces epistemic ideas with oppressive potential on minorities and groups that have been targeted from colonial practices to the present day. When presenting the principles of RAI, the article highlights the risk of embracing, without the necessary critical thinking, the formal rules imported from the Global North as “THE” solution to mitigate the possible impacts of AI, when educational and communication actions within the culture become necessary, and they will only be possible if RAI becomes Slow AI.

**Keywords:** artificial intelligence, bioethics, decoloniality, responsible AI.

<sup>1</sup> Universidade de São Paulo, Instituto de Estudos Avançados, Cátedra Oscar Sala (IEA/USP, São Paulo, SP, Brasil). ORCID: <https://orcid.org/0000-0002-6275-2799>.

## 1 - Introdução

Quando pensamos em territorialidade, o que nos vem à mente é o espaço geográfico. Entretanto, como salienta o filósofo Lewis Gordon (2021, p. 8) em “Freedom, Justice and Decolonization”, a euromodernidade produziu uma forma especial de alienação, reduzindo grupos inteiros a categorias tais como indígenas ou nativos, negros, colonizados e escravizados, de modo que as pessoas que tiveram sua cultura devastada passam a sofrer permanentemente de uma espécie de melancolia e luto pela separação de territórios onde a ‘casa’ e o pertencimento não estão restritos ao espaço geográfico, mas também à identidade que se compõe de elementos que vão desde a relação com o espaço e a linguagem até a dignidade do existir nas inter-relações. Territórios de conhecimento foram e são violentados, massacrados, e, no seu lugar, existe apenas um não lugar, o não ser aceito, o não fazer parte de um mundo onde os grupos vitimizados deixam de SER para se tornarem ‘problemas’.

Assim, quando o mundo inteiro se volta para os potenciais e limites da inteligência artificial (IA), temos uma oportunidade de reacender alguns debates e fazer circular outros que poderiam estar restritos a pequenos nichos, como o tópico da decolonialidade. Porque a IA é uma representação de conhecimento científico e filosófico acumulado, dentro da tradição ocidental, e suas aplicações revolucionam diversos campos da sociedade, dos mais visíveis aos invisíveis.

Se a bioética surge a partir de avanços técnico-científicos que impactam a sociedade e o meio ambiente, o cuidado da saúde, assim como novos modos de comunicar e educar, é porque os valores humanistas que emergem no mundo moderno são confrontados não apenas nas suas inconsistências e contradições, mas também quanto a onde haveria consenso sobre o que é ou não aceitável.

Desse modo, o objetivo deste texto é trazer reflexões bioéticas sobre os desafios da Terra e de os todos que nela habitam frente à corrida desenfreada dos desenvolvimentos de inteligência artificial (IA) aplicados em todo o mundo de maneira ubíqua, pervasiva, com impactos visíveis e invisíveis. A bioética surge como uma ponte entre as ciências e humanidades, quando os projetos tecnocientíficos se expandem com ameaças diretas à vida, esta compreendida como um bem inegociável.

Buscar entender tais ameaças a partir de relações de poder tem conduzido a bioética a uma viagem exploratória que se inicia com a atenção à ética aplicada às ciências da vida, desde a filosofia clássica ocidental até os temas interseccionais das ciências humanas contemporâneas. Assim, um primeiro passo de ruptura com as epistemologias mais tradicionais do Ocidente é entender as relações de poder em suas expressões de controle – biopoderes e biopolíticas – embutidas no conhecimento e suas linguagens. Conforme se argumenta na obra de Michel Foucault (1988, 2001, 2008), novas políticas da era moderna geraram hábitos que se inserem e interferem na vida de maneira mais incisiva, a partir dos modos por meio dos quais o conhecimento é organizado e aplicado. Em outras palavras, Foucault chamou nossa atenção para o fato de que ‘a verdade’ evocada pelas instituições, sejam do conhecimento científico ou das estruturas político-administrativas, não existe fora do poder ou sem o poder. É uma verdade deste mundo, produzida graças a múltiplas coerções (Foucault, 1993, p. 12). Sua influência sobre os comportamentos demonstra uma espécie de ‘adestramento’ capaz de influir no agenciamento sobre o próprio corpo. Foucault argumenta que “o Século XVIII encontrou um regime por assim dizer sináptico de poder, de seu exercício **no** corpo social e não **sobre** o corpo social” (Foucault, 1993, p. 131).

A partir de mudanças estruturais nos modos de vida, as tecnologias, aliadas aos meios de produção, tornam-se mais incisivas sobre os corpos, permeando desde a relação com o trabalho e controle do tempo, até o policiamento e influência sobre a imaginação. Este contexto levou as histórias de opressão para um outro nível, o da repressão, em que há a passagem de regimes fundamentalmente baseados na punição para um outro, em que sobressai a vigilância como instrumento de autoridade (Foucault, 1993, p. 130).

O entendimento sobre os mecanismos de controle como exercício de autoridade é fundamental para compreender como o conhecimento encontra barreiras territoriais de modo a dificultar a livre expansão do pensamento decolonial para todas as esferas institucionais. Pensar a decolonialidade na IA vai além da denúncia sobre os modos de exploração nos usos indevidos e não consentidos de dados, ou da maneira como as classificações e orientações algorítmicas desses dados perpetuam desigualdades e preconceitos históricos, pois na base destes problemas existem territórios de conhecimento intocados, em suas formas de validação e compreensão. E 'descolonizar' como um campo de estudos em que apenas alguns especialistas terão autoridade para pronunciar enunciados científicos não gerará a transformação necessária evocada pelo significado do termo. As vozes de alguns teóricos relacionados à filosofia africana e aos conhecimentos indígenas não serão capazes, sozinhas, de transformar o *status quo*. É preciso construir pontes entre os conhecimentos que regem toda a nossa estrutura educacional em todos os níveis e outros saberes que fazem parte das heranças culturais de nosso território. E o ato de construir é o ato de inovar, transformar e requer pensamento criativo em metodologias não lineares: exige processos em movimentos espirais e um reconhecimento sobre os saberes que se constroem para além das fronteiras da escola e da universidade. E onde e como estes conhecimentos aparecem ou estão invisíveis? Nos dias de hoje, nas infraestruturas computacionais. Por este motivo, pensar a decolonialidade para a IA requer atenção às infraestruturas tecnológicas, que são compostas de conhecimentos, práticas sociais: os dados por elas produzidos refletem acesso ou exclusão, violência ou proteção institucional, super-representação ou sub-representação de imaginários; do mesmo modo, as definições de prioridades sobre os desenvolvimentos tecnológicos são determinadas a partir dos interesses de quem os domina.

Aos conceitos de *biopolítica* apresentado por Michel Foucault (2008) e *necropolítica* de Achille Mbembe (2016), acrescenta-se – a respeito dos modos com os quais as estruturas de conhecimento da racionalidade moderna aprisionam, controlam e vigiam os corpos, limitando sua autonomia – o fato de que também nestas estruturas do conhecimento existem conflitos inerentes às dinâmicas de poder que reforçam os ditames sobre as vidas que mais importam e as que menos importam. Apesar de princípios universais de direitos humanos (ONU, 1948), a tradição do pensamento ocidental segue influenciada por uma epistemologia que se organiza em distinções binárias como o princípio hilemórfico aristotélico e o entendimento sobre *bios* e *zoe*, a vida cidadã e a vida nua. Assim, a dupla categoria fundamental da política ocidental não é aquela em que ou se é amigo ou se é inimigo. Ela se expressa muito mais nas tensões entre *zoe*, a 'vida nua' ou sem existência política, e a *bios*, ou seja, a vida salva da exclusão (Agamben, 2002, p. 16). *Bios* é a vida incluída nas equações com maior riqueza de detalhes, que vai além da vida reduzida a blocos e categorias de informação direcionada a alimentar estatísticas.

Portanto, enquanto as tecnologias emergentes possuem largo alcance, estando presentes na vida diária de mais da metade da população global, podemos nos perguntar a que conhecimentos e territórios elas mais correspondem, quais políticas e ideais de mundo estarão nelas embutidas, e, quando elas falham, quem tenderão a prejudicar. A ‘vida nua’ é a matéria de preenchimento dessas tecnologias para os propósitos de melhor atender à ‘vida cidadã’; e esta cidadania, conforme verificamos, não é, ainda no século 21, acessível a todas as pessoas como um direito universal.

Tomando como exemplo alguns usos da IA na área da saúde, pensar em decolonialidade é um convite para observar as formas de conhecimento que geram dados, organizam, classificam e produzem resultados. Isso é o que chamei de ‘bioética da não presença’ como um exercício contínuo de análise de fatos para além de ‘evidências’ (Nas, 2021).

Ruha Benjamin (2019), por exemplo, argumenta que o viés algorítmico se expressa também em um racismo invisível advindo de suposições sobre riscos à saúde e avaliação de acesso ao tratamento de acordo com características da população negra cuja situação socioeconômica é, em geral, precária. Este fato gerador de contraste nas Américas é uma das consequências do colonialismo. A autora aponta que, no ambiente da saúde, em muitos casos se pressupõe que as pessoas negras são mais fortes e não precisam de anestesia; pela mesma razão, estes sujeitos sociais recebem atendimentos mais apressados do que pessoas com características brancas. Desse modo, podemos perceber que há preconceitos implícitos sobre as vidas que têm maior ou menor valor dentro do imaginário social; tais práticas são interpretadas como dados que geram estatísticas e, quando são encapsuladas dentro de uma IA, elas se tornam verdades opressoras. Uma vez que a opressão atinge de maneira recorrente os grupos que mais sofreram com o colonialismo da Era Moderna, vivemos o risco da IA inaugurar um novo colonialismo que se impõe em camadas distintas, do nível mais objetivo e superficial ao mais profundo e subjetivo.

O colonialismo impactou as relações do corpo com o território, impactando sobre o que é considerado conhecimento ou não-saber, como mover-se, ter ou não direito de ir e vir, ter ou não soberania sobre o próprio corpo. Soma-se a este fator o quanto as formas de ver e conhecer o mundo estão inscritas no corpo como formas de se portar, vestir, organizar a comunidade e educar. O patriarcado e o colonialismo euromoderno impuseram o que entendem como modos superiores no agir, reivindicando inclusive autoridade sobre o conhecimento do que deveria ser a ética, e sobre o que deve ser considerado mais eficaz, avançado e adequado.

Assim, ainda que a comunicação através deste texto venha a requerer palavras amplamente adotadas em nosso território como parte da herança do modelo europeu ocidental nas nossas universidades, e que tenhamos benefícios em muitas áreas das ciências e humanidades, a leitura decolonial visa chamar a atenção para os modos por meio dos quais os padrões de violência, exploração, subalternização e usos de pessoas, corpos, grupos e territórios também estão presentes nestes conhecimentos de modos invisíveis e, com a IA, poderão se tornar ainda mais pulverizados e de difícil identificação.

Portanto, quando o mundo se volta aos esforços de regulamentação da IA dentro da perspectiva de implementar a IA Responsável (IAR) como uma estrutura capaz de condensar os princípios de justiça, transparência, explicabilidade e prestação de contas, devemos estar atentos para que as proposições de legislação não se tornem um novo contrato social iluminado por especialistas, que na prática apenas mantém

desigualdades e desequilíbrios expressos em entendimentos enviesados sobre a lei, conferindo os privilégios de acesso e aplicação da justiça aos mesmos atores hegemônicos hoje, sob as mesmas óticas da cultura colonialista<sup>2</sup>. E, uma vez que os sistemas de informação operam em caixas fechadas através da IA, devemos nos perguntar em que medida eles estão reforçando a cultura de violência colonial, agora, por meio de instrumentos que tornam ainda mais difícil a sua identificação. Estes instrumentos são também os modos pelos quais os modelos de conhecimento hegemônicos exercem sua autoridade, desqualificando narrativas em territórios de conhecimento que não seguem os modelos lineares e cercados de fronteiras.

Mesmo na tradição do conhecimento ocidental que tem no topo de suas hierarquias os homens e pessoas brancas, há um grande número de autores e obras produzidas nos últimos 50 anos em que se entende a necessidade de resgatar o diálogo interdisciplinar para a melhor compreensão de uma cultura que se movimenta a partir da ‘hibridação’ nas relações humano-máquina (Nas, 2020). E o que se constata é uma ‘retração’ contínua das buscas que desafiam os ideais da *formalidade*.

De acordo com Gilbert Simondon (1995, p. 49), o conceito de ‘forma’ é uma mostra da influência da vida política sobre a teoria. Assim, a separação entre ‘forma’ e ‘matéria’ reflete a estrutura do pensamento formado na antiguidade clássica em que a ‘forma’ provém de ideias da vida qualificada, dos cidadãos, e a ‘matéria’ é o que preenche a forma, serve a propósitos, em que se entende a passividade ou ausência de liberdade de escolha (Nascimento, 2017, p. 135).

Por conseguinte, sugere-se perguntarmos o que é conhecimento, quais os seus propósitos ou como se pode entender a sua relação com a vida, para finalmente situarmos onde a ‘forma’ contribui para novas descobertas, e onde ela apenas poderá reproduzir o que se chama de *conhecimento como ‘produtos’ de autoridade*. Fundamentalmente, se o conhecimento disforme (que foge ou destoa da forma) ainda é conhecimento, ou se, como um produto ‘defeituoso’ (um tijolo com rachadura ou em proporção distinta do formato retilíneo), deve ser descartado<sup>3</sup>.

## 2 - IA Responsável

O advento da internet concretiza o modo supraterritorial das novas políticas industriais e financeiras que, em última instância, regem as políticas de governo em todas as suas áreas. A globalização inaugura o deslocamento do espaço físico para a

2 Podemos citar muitos exemplos cotidianos da vida política brasileira, como o caso de Rafael Braga, que foi preso como ativista no Rio de Janeiro em 2013, e, enquanto ativistas brancos de classe média foram liberados, ele ficou retido, mesmo que o advogado tenha se apresentado para liberação de todos os envolvidos nas manifestações. Ou as declarações de políticos de extrema-direita minimizando a gravidade do assassinato da vereadora da Cidade do Rio de Janeiro Marielle Franco em 2018 pelo fato de ser mulher e negra. E assim poderíamos seguir com a lista, resumindo que, segundo o Atlas da Violência (IPEA, 2023), quase 80% das vítimas de homicídios no Brasil são pretos e pardos entre os homens e quase 70% entre as mulheres (também pretas e pardas).

3 Esse debate me é caro desde que pausei minha vida acadêmica (iniciada nas Ciências Sociais com trabalho premiado em Ciência Política) por mais de 20 anos, quando ganhei dois festivais de música e fui para o exterior. Mas em épocas de tecnologia analógica e concentração de recursos nos agentes de mercado, o conhecimento incapaz de ser produto é insignificante e uma tecnocracia dita que o ‘ruim’ pode ser ‘bom’ desde que em formas apropriadas, aceitas e vendáveis, e que o ‘bom’ não ‘serve’ para nada se não está formatado adequadamente pelas tecnologias em voga e, principalmente, se não consegue ser ‘produto’. Esse comentário não foge ao que é aqui apresentado: na sociedade industrial tudo deve ‘funcionar’ como uma fábrica, inclusive – e infelizmente – a produção de conhecimento na esfera acadêmica.

nuvem, regida pela supremacia dos cálculos e algoritmos. A política neoliberal passa a se desresponsabilizar tanto quanto possível dos impactos dos negócios nos territórios e nas vidas (Bauman, 1999)<sup>4</sup>.

Desse modo, enquanto nos novos desenvolvimentos tecnológicos as informações chegam por vias imateriais e tornam-se “modelos” abarcados em softwares, aplicativos e dispositivos, ainda existem lacunas sobre como mensurar os seus impactos nos territórios e vidas, individuais e coletivas, humanas e não humanas.

Enquanto o debate sobre o colonialismo digital aponta para a ampla extração de dados em poder de grandes empresas de tecnologia, tal debate se adensa quando identificamos onde tais empresas estão sediadas. São novas formas de imperialismo e colonialismo que expandem o modelo inicial de exploração do território, extração de minérios, para novas explorações de vidas através de extração e mineração de dados. Quando estes dados se tornam parte de uma IA que se apresenta como um bem para toda a humanidade, torna-se necessário observar como os vícios inerentes ao sistema poderão se amplificar através da IA e causar danos.

Considerando que o Brasil não desenvolve as tecnologias fundamentais e os principais produtos e ferramentas globais de inteligência artificial, nos perguntamos de que maneira a sociedade poderá obter informações de modo a construir uma análise crítica sobre como o país pode adquirir e desenvolver tecnologias de IA capazes de promover inovações radicais em favor do benefício comum, em vez de apenas realizar ajustes superficiais em produtos, serviços e sistemas já existentes. Enquanto tal questão está hoje na pauta de vários países, o fato de que as economias intensivas em IA irão se distanciar de forma crescente das demais aponta não só para uma corrida pela produtividade do futuro, mas também para a perpetuação de territórios de conhecimento onde há maior representatividade masculina, branca e do Norte Global. A única maneira de modificar este cenário seria através de uma maior participação popular indicando os modos de se apropriar das tecnologias, assim como as decisões sobre implementá-las ou não de acordo com os setores, definindo prioridades e limites.

Muitas instituições internacionais têm formulado diferentes conjuntos de princípios instrumentais para a implementação da inteligência artificial responsável (ACM, UNESCO e OECD)<sup>5</sup>. Esses princípios observam como minimizar riscos específicos do uso dessas tecnologias, como equidade, viés, autonomia, privacidade, proporcionalidade, responsabilidade e prestação de contas (OECD, 2019; UNESCO, 2023). Buscar minimizar riscos através de normas e princípios não é o mesmo que garantir a plena observância de princípios éticos e de justiça para os desenvolvimentos de IA desde as infraestruturas até as interações humano-algorítmicas (IHA), interações humano-computador (IHC) e interações humano-robô (IHR). Ou ainda, se pensarmos no principialismo evocado na bioética, de respeito à autonomia, o princípio da beneficência, da não maleficência e de justiça, percebemos que a IA não dá garantias sobre a observância de tais princípios.

---

4 “Livrar-se da responsabilidade pelas consequências é o ganho mais cobiçado e ansiado que a nova mobilidade propicia ao capital sem amarras locais, que flutua livremente. Os custos de se arcar com as consequências não precisam agora ser contabilizados no cálculo da ‘eficácia’ do investimento.” (Bauman, 1999, p. 16-17).

5 Agradeço ao prof. Virgílio Almeida as contribuições à primeira versão da seção sobre IA Responsável.

Se, quando falamos de responsabilidade, temos implicitamente o chamado à consciência ética de não causar dano, em que medida poderemos fazer com que os princípios de transparência, interpretabilidade, explicabilidade, auditabilidade e prestação de contas na IA serão observados por empresas e instituições?

No contexto brasileiro, podemos pensar em que medida tais princípios se adequam às políticas institucionais quando o tema é transparência e prover explicações sobre decisões. Enquanto os relatórios e recomendações são produzidos no Norte seguindo os interesses dos países mais ricos, para eles poderá fazer sentido buscar aplicar as regras para proteger os seus cidadãos enquanto suas empresas lucram no Brasil, onde a observância da justiça é falha e seletiva, dentro de uma política com influências coloniais que ainda é coronelista e clientelista.

Assim, se esses princípios norteiam o desenvolvimento de aplicações de IA em diversas áreas, como saúde, educação, finanças ou mídias sociais, a IA Responsável deve – em tese – combinar a governança de IA (dados e modelos) com treinamento das pessoas certas para implementá-la. E o que se entende por ‘pessoas certas’ deverá contemplar diversidade de especialidades, áreas do conhecimento, representatividade setorial, assim como pluralidade de visões. Entendemos que, por razões pragmáticas, há uma tendência a evitar a inclusão de perspectivas ‘dissonantes’ do que alguns campos já vêm discutindo, porém os desafios atuais requererem um maior entendimento sobre a ética da hospitalidade<sup>6</sup> como um desafio humano que deve ser mais importante do que a corrida das aplicações e regulamentações da IA. Se os objetivos mais nobres dos desenvolvimentos de IA forem, como sempre evocados, os de melhorar a qualidade de vida para todos e amplificar a potência da inteligência humana, será possivelmente o caso de propormos uma IA Lenta (*Slow AI*)<sup>7</sup>. Isto é, reformular os modos de desenvolvimento da IA documentando cada etapa do processo, as fontes de informação, as decisões na criação do algoritmo, e fundamentalmente ouvir a sociedade, as partes impactadas, as partes que mais sofrem com discriminações e leituras errôneas. A revisão dos métodos também exige uma mudança curricular na formação de cientistas da computação e programadores em geral. E esta formação, mais do que ética, poderá ganhar novos contornos críticos através da bioética, desde que não se trate apenas de entender o que é ética do ponto de vista teórico, e sim que perspectivas sobre a vida estas teorias trazem e qual seria a melhor abordagem para garantir a observância dos princípios de justiça com atenção à inclusão e diversidade.

Os avanços na pesquisa em IA ao longo das últimas duas décadas têm sido atribuídos à combinação do aumento da potência computacional, disponibilidade de grandes volumes de dados e progressos em algoritmos de aprendizado de máquina (Mohamed; Png; Isaac, 2020). A abordagem decolonial da inteligência artificial identifica nas práticas contemporâneas de coleta de dados o potencial surgimento de uma nova forma de governar e distribuir poder em sociedades e economias (Couldry; Mejias, 2023). Considerando que o colonialismo histórico foi um modelo extrativista,

6 A ética da hospitalidade pode ser encontrada em diversos autores, como no pensamento pós-estruturalista de Jacques Derrida, assim como podemos fazer um paralelo com o conceito africano *ubuntu*, em que o humano se complementa, comunica e exerce sua potência de maneira relacional. Este também pode ter um paralelo com a clássica ética das virtudes, sendo ‘fazer o bem’ um modo de exercitar e lapidar as virtudes.

7 A ideia de uma ‘*slow AI*’ apareceu-me intuitivamente como um elemento novo e original, porém, ao pesquisar na web, encontrei um pequeno artigo em um blog que, além de se alinhar com as referências deste texto, reforça o chamado de que os sistemas de informação, para serem éticos, devem priorizar as vozes das pessoas de modo a promover uma sociedade mais justa e inclusiva (Conroy, 2023).

que reorganizou as sociedades em vários níveis, a abordagem decolonial para a IA não se concentra apenas nas grandes plataformas digitais, como redes sociais e motores de busca, mas também nos hábitos mais amplos de coleta de dados em todos os aspectos da vida social, econômica e política. A expressão “colonialidade algorítmica” tem sido usada para expandir a noção de colonialismo de dados no contexto das interações de algoritmos em sociedades (Mohamed; Png; Isaac, 2020). Algoritmos tomam decisões, definem recursos e moldam o comportamento sociocultural e político individual e coletivo globalmente (Mendonça; Almeida; Filgueiras, 2024). A linguagem da decolonialidade trazida para o contexto da IA oferece uma nova leitura para conceitos fundamentais a serem buscados para os sistemas algorítmicos, que são justiça, responsabilidade e transparência nas decisões tomadas por esses sistemas. No contexto da IA Decolonial, os algoritmos se encaixam numa taxonomia de uma visão colonial, como: opressão algorítmica institucionalizada, exploração algorítmica e despossessão algorítmica (Zuboff, 2019).

A parte que permanece invisível na opressão algorítmica são os sistemas de conhecimento entendidos como neutros, científicos e eficazes. Ainda, a parte fantasiosa que eleva a IA ao nível super-humano e sobrenatural, como uma força inevitável e além de nossas capacidades de compreender ou controlar. Ela carrega um imaginário sobre o futuro como libertação das limitações humanas através da ciência. Assim, para Timnit Gebru, para que a IA contemple os valores de um futuro com equidade e humanidade, ela precisa ser trazida de volta para a Terra (DAIR, 2021). E as fantasias e romantização das tecnologias também fazem parte de um sistema de privilégios que se sustentam por extremas desigualdades, exploração de povos e territórios, desde a indústria 1.0.

Se desfazer a história dos genocídios e escravização dos povos em territórios colonizados não é possível, ainda podemos mudar definitivamente o curso dos acontecimentos entendendo que os riscos já evidenciados nos usos da IA mostram fragilidades mais profundas do que é possível ‘consertar’ através de um e outro ‘remendo’ no algoritmo. O futuro que se delineia é que, na esfera política e social, nada irá mudar e ainda pode piorar se ações enérgicas não forem tomadas para desacelerar as implementações de IA em favor da responsabilidade.

A Declaração de Princípios para Sistemas Algorítmicos Responsáveis (ACM, 2022) considera que as IAs estão cada vez mais sendo utilizadas em todas as esferas da sociedade com potencial de grande impacto. Em uma perspectiva bioética, vamos falar de impactos sobre as vidas e não apenas nos danos explícitos, como também naqueles invisíveis que transformam as visões que se têm sobre si mesmo, sobre as relações e os modos de ver e entender o mundo. Novamente, o que temos como indicativos para regulamentar são propostas genéricas como a necessidade de uma IA responsável atender aos princípios de legitimidade e competência; minimização de danos; segurança e privacidade; transparência; interpretabilidade e explicabilidade; manutenibilidade; contestabilidade e auditabilidade, com recomendações para que os construtores e operadores de sistemas de IA comparem como seriam as decisões humanas nos contextos aos quais a IA pretende responder; que os desenvolvedores conduzam testes com avaliações de impacto abrangentes; que os formuladores de políticas invistam em auditorias para avaliar as aplicações da IAR nos processos de desenvolvimento e implementação da IA; e que operadores de sistemas de IA desenvolvam consciência sobre suas decisões no processo de elaboração dos algoritmos (Almeida; Nas, 2024).



Em termos práticos, tais recomendações alcançam um nível de sofisticação na medida em que até agora todo treinamento envolvendo a arquitetura dos sistemas é focada fundamentalmente em aspectos técnicos, seguindo metodologias e lógicas que resolvem dúvidas e incertezas com respostas aproximadas, muitas vezes lhes conferindo *status* de verdade. Para mudar estes modos de conhecer, desenvolver e realizar projetos, será necessário investir em significativa mudança paradigmática sobre os modos de entender a tecnologia, suas relações com o humano, a natureza e a sociedade. Ainda, apresentar aos desenvolvedores os conhecimentos éticos, filosóficos de modo não apenas prescritivo, mas crítico, ou seja, através do exercício da reflexão, algo que os grandes modelos de linguagem em IA falham em prover com precisão.

Assim, a implantação da IA Responsável enfrenta diferentes tipos de obstáculos, como a dificuldade para a realização de auditorias nos sistemas de IA que sejam capazes de fazer valer a necessidade de prestar contas à sociedade (Raji; Chock; Buolamwini, 2023). O primeiro passo das empresas para se proteger das auditorias externas é promover a prática internamente contratando profissionais dedicados aos debates éticos em torno da IA Responsável. Porém, tais implementações estão aquém das necessidades e revelam casos de conflitos internos, como da cientista da computação Timnit Gebru, demitida do Google (Hao, 2020) em virtude da coautoria da pesquisa “On the dangers of stochastic parrots: Can language models be too big?” (Bender *et al.*, 2021). O trabalho apontou problemas sobre o custo ambiental do treinamento dos modelos de linguagem, os vieses racistas e sexistas, entre outros perigos como a geração de desinformação.

Nos últimos anos, os debates éticos em torno do desenvolvimento das tecnologias emergentes são frequentemente apontados como obstrutores na corrida das inovações de IA e robótica. Tem havido, portanto, um processo lento para que as empresas adotem novas visões que reflitam contratações de pessoal capazes de propor metodologias e reflexões para atender as demandas sobre uma IA ética e responsável. Na ausência de um trabalho estrutural que se consolide dentro das empresas através de equipes internas qualificadas para entender os impactos sociais das IAs, uma segunda alternativa é contratar consultorias para auditar os sistemas em desenvolvimento. Estas consultorias apenas indicam sua visão sobre as possíveis formas de mitigar os problemas identificados, sem, contudo, ter autoridade sobre quaisquer modificações que devem ser feitas pelas próprias empresas. Além disso, não há garantias de que as empresas forneçam todas as informações em virtude da proteção de sigilos e privacidades e das informações que escolhem abrir. A isso se acrescenta o limite contratual de sigilo para as consultorias.

As pesquisadoras Raji, Costanza Chock e Buolamwini (2023) enfatizam que é necessária uma “terceira via” para auditar os sistemas de modo que um setor independente das empresas possa atuar de maneira mais assertiva, pontuando os possíveis problemas, dúvidas e até denunciando, se necessário. Desse modo, as autoras sugerem que tais auditores devem ter amparo e proteção legal para que possam compartilhar com transparência suas conclusões. Mais do que isso, para que representantes da sociedade participem de auditorias, deve-se viabilizar um processo de qualificação em etapas. Portanto, são proposições também educativas que demandam debate, investimentos e tempo. As autoras afirmam que o contexto atual não possibilita a participação efetiva de *third-party auditors* e fornecem propostas de regulamentação para garantir acreditação, proteção e apoio a auditores externos e independentes em

favor de políticas equitativas para a implementação de uma IA justa e responsável. A liberdade de revelar publicamente os resultados de uma auditoria deve ser defendida e não coibida, ainda que a transparência da informação se confunda com o caráter de denúncia quando os modelos averiguados já se encontram em uso e estão potencialmente causando danos, como os casos apresentados pela *Pro Publica*, que revelaram vieses nas previsões da IA que, além de errôneas, eram tendenciosas, pois consideraram indivíduos de pele negra como de maior risco a reincidir em atos ilegais do que aqueles de pele clara (Angwin *et al.*, 2022).

Em resumo, para que as aplicações de IAR sejam efetivas e causem real impacto, elas são dependentes de mudanças estruturais desde a pesquisa (P&D) e seus métodos, fato que confronta os territórios e fronteiras do conhecimento. Positivamente a IAR, se implementada conforme proposto nesta abordagem, é uma oportunidade para promover debates interdisciplinares e estimular mudanças.

### 3 – Decolonialidade da IA

O debate sobre a decolonialidade da IA é recente e se identifica com os problemas éticos recorrentes com o avanço das tecnologias de IA. Por exemplo, quando a IA reforça estereótipos de beleza e a mulher negra é frequentemente identificada como ‘feia’ (Araújo; Meira; Almeida, 2016) e com adjetivos pejorativos (Noble, 2018) – assim como as *latinas* nos EUA –, há um reforço de preconceitos antigos que no contexto colonial representaram ideias utilizadas para justificar a subalternização e exploração dos grupos que mais sofrem as consequências do colonialismo industrial moderno.

Os mecanismos de busca automatizados frequentemente reproduzem preconceitos contra as mulheres (Noble, 2018) e tendem a enfatizar estereótipos em que a imagem do belo é branca (Magno *et al.*, 2016), ainda que a busca na internet tenha sido feita no Brasil, um país onde a maioria da população é negra e parda. A perspectiva decolonial, portanto, busca atingir os elementos menos evidentes que impactam as subjetividades entre perspectivas éticas e estéticas que não são facilmente visíveis dentro dos desenvolvimentos tecnológicos. Tais elementos invisíveis nos produtos e artefatos são frequentemente suprimidos, como desastres ambientais e danos à vida, em geral (Nas, 2021).

Desse modo, o conceito de decolonialidade se aplica à inteligência artificial na medida em que os desvios éticos da IA impactam e causam danos à vida. A pergunta fundamental em que o conceito da decolonialidade emerge é: quando há falhas de interpretação, reconhecimento e recomendação dos algoritmos, quais são as consequências e para quem? (Nascimento, 2019).

As falhas comumente ocorrem porque os modelos são treinados e desenhados dentro das perspectivas dos países desenvolvidos, seus valores, modos de ver e entender o humano e o mundo. Fala-se de um tecnocolonialismo quando tudo caminha em uma direção de dependência de tecnologias que são criadas, administradas e que possuem toda infraestrutura nos territórios do Norte Global. São tecnologias que carregam o conhecimento de uma tecnoutopia etnocêntrica, antropocêntrica e especista quando anunciam que a IA poderá melhorar as vidas de mulheres e crianças em situação de vulnerabilidade em todo o mundo (Global Grand Challenges, 2023). Na prática, este discurso corrobora a imposição de um ideário ocidental que confere um

*status* subalterno aos conhecimentos e perspectivas de outras culturas, e, no caso da IA, a colonialidade do conhecimento se reflete no conjunto de dados (*datasets*) que são centrados na cultura do Ocidente (Mboa Nkoudou, 2023).

As características de uma monocultura aparecem quando a IA apresenta com frequência indivíduos de pele clara como biotipos característicos dos territórios do Norte Global. Dessa maneira, tende-se a criar um padrão do que seria o “humano” ou “pessoa”, que em geral também se refere ao gênero masculino. Isto gera, em termos de produção de imagens, uma monocultura ou supremacia cultural que invisibiliza a diversidade, potencialmente influenciando bloqueios sobre a alteridade (Nas, 2023b).

Se queremos elencar quais as características fundamentais da decolonialidade da IA, ela se refere à sub-representação de dados relacionados à diversidade cultural e epistêmica. Assim, as injustiças experienciadas pelas pessoas negras e indígenas na era colonial e de extração monopolista de recursos tornaram-se protótipos de futuras explorações aplicadas globalmente através dos algoritmos de IA (Miller, 2022).

Atualmente, quando ferramentas de IA são utilizadas para contratar ou demitir alguém, são aplicados critérios que na prática tendem a subjugar pessoas e grupos que historicamente fazem parte da parcela oprimida pelas estruturas de poder. Desse modo, a perspectiva decolonial propõe um olhar que vá além de simplesmente ver uma falha da IA como um caso isolado ou como representação de tipos de preconceitos criticáveis pela ética humanista do mundo industrial moderno. A fissura entre teoria e prática operada por esta ética elaborou exceções julgadas por sistemas de autoridade em que outros modos de conhecer e explicar o mundo são obstruções deliberadamente ignoradas.

A perspectiva decolonial da IA ultrapassa os limites das suas aplicações nos mais diversos contextos, como saúde, trabalho, educação, arte ou segurança pública, entre outros. Em todos estes contextos há vieses demonstrando que a tecnologia serve melhor a alguns e pior a muitos, refletindo, assim, os grandes problemas de concentração de riqueza *versus* situação de extrema pobreza em configurações de forças que separam os que estão mais próximos e os que estão mais afastados de certos privilégios. A crítica decolonial é política e também diz respeito à cultura; por esse motivo, ela não é tecnofóbica, pois não há como eliminar da cultura o que representam hoje as tecnologias na vida. Assim, ela não se contrapõe aos desenvolvimentos tecnocientíficos, mas propõe reflexão sobre o que se quer, a quem atende, quais os impactos e sobre quem eles incidem. E, a partir de tal consciência, descobrir os caminhos possíveis de coexistência, a partir de reestruturações conceituais e de modos criativos de pensar sobre as tecnologias emergentes.

Os desafios de pesquisa para uma IA Decolonial incluem a necessidade de infraestruturas descentralizadas que colem dados através de pesquisa etnográfica e qualitativa e que sejam base para novos desenvolvimentos locais de tecnologias de IA. É necessário coletar e organizar dados de maneira responsável para que os ambientes de rede representem diversidade e atendam aos anseios éticos, sociais e ambientais, que valorizem ações afirmativas e busquem sintonia com os objetivos do desenvolvimento sustentável, combatendo racismo, injustiças sociais e problemas ambientais (DecolonizAI, [2022]). Igualmente importante é estabelecer colaborações com pesquisadores do Sul Global, da diáspora africana e latino-americana no Norte.

Mboa Nkoudou (2023), do Centro de Expertise Internacional de Montreal para a Inteligência Artificial (CEIMIA), propõe: priorizar necessidades locais incentivando a participação das comunidades na identificação dos problemas que desejam solucionar; investir em coleta e desenvolvimento de bases de dados que representem as diversas

culturas do território; que instituições culturais, acadêmicas e os desenvolvimentos de IA devem considerar os conhecimentos dos povos originários; construir bases de dados com participação proativa dos grupos marginalizados; que educadores e mídias devem resistir às narrativas dominantes, engajando-se em perspectivas críticas e oferecendo programas para letramento tecnológico questionadores das narrativas tecnológicas e fazendo com que elas se alinhem às experiências e culturas locais; que instituições educacionais e governamentais devem investir em talentos locais para os desenvolvimentos da IA; que Comitês de Ética e organizações de IA devem adotar práticas éticas que incluam as comunidades de modo a garantir que tais desenvolvimentos reflitam os valores sociais e princípios éticos locais; incentivar colaborações internacionais com equidade nas tomadas de decisões; que o Legislativo e Judiciário devem garantir a soberania digital com políticas e leis adequadas ao contexto de modo a garantir que os dados permaneçam sobre controle da comunidade e para o seu benefício; que a comunidade de IA global deve celebrar as conquistas locais através das mídias de modo a evidenciar os avanços de representatividade frente às narrativas dominantes.

As recomendações abrangem e contemplam de forma concisa a inter-relação entre os desenvolvimentos científicos e a atenção à cultura dos povos, territórios, com respeito a saberes distintos do fazer hegemônico. Sobretudo, um convite a entender que não apenas os conhecimentos e modos de vida mais distantes do ocidental, mas que grande parte da população das terras colonizadas também possui em seu DNA as marcas de etnias que foram extintas por genocídios, sequestros e escravização que resultam em melancolia em consequência de um luto por separação e perda de um mundo substituído por outro ao qual não pertencem com dignidade (Nas, 2023a). A incapacidade de exames de DNA revelarem ancestralidades apagadas por ausência de dados ilustra também como o método científico embarcado na IA cria falsas expectativas, super-representando grupos em que há maior riqueza e quantidade de dados e sub-representando outros para os quais não há informação suficiente, colaborando, deste modo, para a invisibilidade de etnias que desapareceram por genocídios e políticas de mestiçagem. Lewis Gordon argumenta que o conceito de raça é fabricado e lembra que Antenor Firmin questionou o rigor das alegações das ciências humanas euromodernas, pois “a ciência rigorosa se adequa às demandas de seu objeto. Não tenta forçar a realidade a encaixar-se em suas pressuposições” (Gordon, 2023, p. 107). Temos aí o problema epistemológico de base que é também como se define um algoritmo: tudo parte de uma fórmula (forma) que irá determinar uma sequência de ações e resultados.

#### **4 – Discussão**

Sistemas de IA podem ter efeitos transformadores e de longo prazo em indivíduos e na sociedade. Para gerenciar esses impactos de maneira responsável e direcionar o desenvolvimento de sistemas de IA para o benefício público, os princípios da IA Responsável estabelecem regras e orientações sobre como projetar, desenvolver, implementar e auditar os serviços baseados em tecnologias de inteligência artificial, enquanto a proposta de uma ‘IA Decolonial’ é atentar para os riscos que envolvem uma IA opressora que reforça a monocultura e epistemicídios. Quando Cathy O’Neal revela como operam os “algoritmos de destruição em massa” (O’Neil, 2017, 2021), podemos imaginar a metáfora do Cavalo de Troia como uma IA que se apresenta com os melhores

propósitos – na aparência imponente e ‘mágica’ da técnica que a faz um ‘presente dos deuses’ – sob o ponto de vista do imaginário, enquanto existem elementos invisíveis capazes de gerar danos irreparáveis.

A possibilidade do dano vem trazendo o debate ético para as tecnologias, de modo que, ao pesquisar os termos “AI Ethics” AND “ethics in computing” na plataforma Dimensions ([2024]), os resultados em 3/01/2024 foram modestos, com 27 artigos publicados nos últimos quatro anos. Considerando que se trata de um diretório de busca que acessa as bases de dados de revistas e universidades (majoritariamente) em língua inglesa e do Norte Global, entendemos que o debate ético em torno das tecnologias emergentes vem apenas se adensando nos últimos anos e, enquanto os novos desenvolvimentos de IA vêm chamando atenção para os possíveis riscos e como mitigá-los, a produção a este respeito aparece em publicações que não são acessadas pelos diretórios de pesquisa, como relatórios extra-acadêmicos, institucionais e não governamentais.

Diante de tal escassez sobre um histórico reflexivo a respeito de uma ética para as tecnologias, coloca-se o argumento de que, diferentemente de entender a ética como uma capacidade unicamente humana, é preciso torná-la parte dos sistemas, de modo que a IA possua um agenciamento ético (Bertoncini; Serafim, 2023). O questionamento se baseia no entendimento de que entender a tecnologia como unicamente instrumental é uma visão antiquada correspondente ao primeiro momento da revolução tecnológico-industrial, em que as características são distintas das tecnologias de informação contemporâneas, quando softwares e dispositivos condensam um volume maior de informações que alimentam sistemas de aprendizado contínuo através da interação humano-computador e outros meios.

O que se observa, portanto, é a existência de grandes lacunas nos ‘territórios do conhecimento’ desde interseções entre a Filosofia para o entendimento da técnica em suas dimensões existenciais e políticas e as Ciências Sociais e Exatas. Uma ética capaz de reconhecer o direito das coisas como agentes morais (Nas, 2023c) é uma perspectiva metaética que tende a se alinhar com cosmovisões distintas da eurocêntrica, como o perspectivismo ameríndio (Maciel, 2019), que advoga a defesa de direitos para rios, montanhas e tudo que existe.

Diante das lacunas criadas por territórios disciplinares cercados, estas são ainda epistemologias marginais que causam desconforto nos territórios de conhecimento que regem os desenvolvimentos tecnológicos.

O debate meramente técnico a respeito de como os desafios para tornar uma IA ética *by design* assemelha-se à ‘lavagem ética’ (Bietti, 2021)<sup>8</sup>, pois proposições éticas são de difícil aplicação em perspectivas pragmáticas. Do mesmo modo, demandar explicabilidade sobre as decisões da IA é tão complexo quanto explicar as decisões humanas e dar-lhes transparência.

No debate ético não existe um consenso entre especialistas sobre o tipo de abordagem a ser aplicada caso a intenção seja traduzir uma ética através de algoritmos em sistemas computacionais.

8 A discussão sobre *ethical washing* (Bietti, 2021) enfatiza que o olhar sobre a ética não deve ser meramente técnico, mas – de outro modo – deve ser entendido como um exercício filosófico de fazer perguntas como forma de atingir um conhecimento que ainda não está pronto e vai sendo construído pela reflexão. Sobretudo, que contratar filósofos ou consultorias éticas para adequar as empresas às demandas tende a reforçar a ideia do conhecimento exclusivo de poucos técnicos e especialistas que estão distantes do mundo em ‘torres de marfim’.

Inicialmente, os conceitos éticos tradicionais dizem respeito a condutas individuais, enquanto os sistemas de informação reúnem responsabilidades de um conjunto de ‘atores’, entre os quais designers, engenheiros, desenvolvedores, empresas e usuários (Taddeo; Floridi, 2018). Portanto, o contexto apresenta desafios para o esboço do que seria uma ‘ética distribuída’ (Floridi, 2014; Floridi; Sanders, 2004). Ou seja, é necessário o debate aberto sobre novas perspectivas éticas quando a IA avança com grande rapidez, e tais debates são ainda modestos e demandam um outro tempo que envolve reflexão, escutas e encontros multissetoriais, interdisciplinares e inclusivos.

Para que se saia da inércia provocada pelas lacunas entre teoria e prática, necessidades e viabilidades, a solução aponta para a necessidade de uma reforma profunda no sistema educacional. Assim, nos perguntamos: até quando os departamentos de Filosofia seguirão ignorando os desafios postos por um mundo mediado pelas tecnologias da informação, incentivando os alunos pelo caminho tradicional e conservador de aprofundar-se em autores vistos e revistos milhares de vezes, formando filósofos que terão tanta dificuldade de conversar com cientistas da computação como o inverso? A bioética segue como um campo novo e especulativo que discute as disputas entre perspectivas éticas para avaliar contextos específicos. As ciências ditas ‘exatas’, entre outras áreas do conhecimento, poderão se beneficiar com a aproximação dos debates bioéticos se os incorporarem ao seu currículo, de modo a abrir campos de reflexão sobre a ética em pesquisa desde a origem de um projeto. Igualmente, as artes e o design tendem a contribuir na busca dos possíveis caminhos da inovação na esfera do conhecimento.

Por vícios epistemológicos, as tecnologias intangíveis são frequentemente comparadas com as tangíveis<sup>9</sup>; porém, diferentemente, elas são um conjunto de informações medidas e difundidas, modelos de aprendizado humano aplicados à computação, e usam o conceito de inteligência como um modo de treinar, reconhecer e replicar conhecimento existente sedimentado. Além disso, representam estatísticas acerca de repetições, identificadas como padrões, por conseguinte, ‘modelos de verdade’ que na prática são falseáveis.

Assim, a IA se torna uma prótese da inteligência humana (e suas lacunas), e os usuários dos sistemas deixam de ser ‘clientes’ para se tornar ‘produtos’ e cobaias de experimentos algorítmicos (Garcia-Vigil, 2021), ameaçando os valores humanos dentro de princípios éticos e de justiça, em consideração aos significados de biopolítica e necropolítica aqui tratados. Desse modo, os saberes condensados e consolidados das tecnopolíticas permanecem em curso, ainda que criticados, abrindo precedentes para que tecnofascismos impactem vidas, seus corpos, corações e mentes.

## 5 – Considerações Finais

O problema do viés no conhecimento, que define e separa territórios, pode ser examinado quando pensamos que questões éticas são mediadas por avaliações das consequências dos atos, do julgamento de valor sobre o que é “bom”, “aceitável”, “ruim”, “inaceitável”. Uma vez que estas são questões dependentes de narrativas, pontos de vista e contextos, elas também replicam dinâmicas de poder e disputas baseadas em visões de mundo. Desse modo, esforços de governança da IA em direção à regulamentação podem ser indicativos, porém não definitivos, dado que são

---

9 Por exemplo, em comparação com a regulamentação da IA: “no início não havia regulamentação de trânsito para os carros, depois foi feito e tudo foi resolvido”.

incapazes de mitigar problemas estruturais de injustiças sociais embarcados em algoritmos, dados enviesados, e os possíveis maus usos de IA. Uma vez que os que mais sofrem as consequências são os sujeitos excluídos de decisões e oportunidades, os sistemas computacionais apenas replicam os problemas anteriores à existência da inteligência artificial, como o acesso universal aos direitos considerados essenciais à dignidade humana (ONU, 1948).

Assim, neste artigo, propõe-se que o entendimento sobre o que seria uma 'IA Responsável' seja mais amplo do que esforços técnicos na esfera jurídica, legislativa e governamental, já que apenas estabelecer normas não garante o respeito às mesmas e tampouco sua ampla aplicação. A existência de legislação não impede que injustiças ocorram, assim como a possibilidade de punição não impede que muitas empresas desrespeitem direitos essenciais de cidadania.

Soluções meramente técnicas incorrem no risco de incentivar um *ethical washing* por parte das empresas, tornando sua abordagem ética performativa e ineficiente.

Assim, para tornar os princípios da IA Responsável efetivos, são necessários mecanismos de governança que agreguem representações de diversos setores de atuação e conhecimento, com aplicação de metodologias criativas, alternativas e inovadoras, lembrando que não há inovação sem risco, e parte do risco é acolher o que é desconhecido ou que 'destoa'. Evoca-se tratar a alteridade com hospitalidade. E aqui não estamos falando estritamente de indivíduos na sua diversidade de cores, orientações e preferências, mas sobretudo do que eles trazem incorporado na sua expressão no mundo, com ideias e modos de existir que não se definem por rótulos. Finalmente, é preciso agregar os que ousam divergir de maneira produtora, pela não conformidade aos modos com os quais todos os setores da sociedade se tornam máquinas de produção contínua, onde cada elemento deverá cumprir uma utilidade de caráter alienante. Urge acolher as vozes criativas, dissidentes e combatentes da opressão nos territórios do conhecimento, entendendo que bioética, IAR e decolonialidade são campos do saber em construção.

## Referências

- ACM – ASSOCIATION FOR COMPUTING MACHINERY. Statement on principles for responsible algorithmic systems. *ACM Bulletins*, New York, 1 nov. 2022. Disponível em: <https://www.acm.org/articles/bulletins/2022/november/tpc-statement-responsible-algorithmic-systems>. Acesso em: 15 mar. 2024.
- AGAMBEN, Giorgio. *Homo sacer: o poder soberano e a vida nua 1*. Tradução de Henrique Burigo. Belo Horizonte: Editora da UFMG, 2002.
- ALMEIDA, Virgílio; NAS, Elen. Desafios da IA responsável na pesquisa científica. *Revista da USP*, São Paulo, n. 141, p. 17-28, jun. 2024. DOI: <https://doi.org/10.11606/issn.2316-9036.i141p17-28>.
- ANGWIN, Julia; LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren. Machine bias. In: MARTIN, Kristen. *Ethics of data and analytics*. Sebastopol: Auerbach Publications, 2022. p. 254-264.
- ARAÚJO, Camila Souza; MEIRA, Wagner; ALMEIDA, Virgílio. Identifying stereotypes in the online perception of physical attractiveness. In: SOCINFO – SOCIAL INFORMATICS INTERNATIONAL CONFERENCE, 8., 2016, Bellevue. *Anais [...]*. Bellevue: Springer International Publishing, 2016. p. 419-437. Disponível em: <https://repositorio.ufmg.br/bitstream/1843/ESBF-ALLHYX/1/camilasouzaaraujo.pdf>. Acesso em: 15 mar. 2024.
- BAUMAN, Zygmunt. *Globalização: as consequências humanas*. São Paulo: Zahar, 1999.
- BENDER, Emily M.; GEBRU, Timnit; MCMILLAN-MAJOR, Angelina; SHMITCHELL, Shmargaret. On the dangers of stochastic parrots: can language models be too big?. In: ACM CONFERENCE

- ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, 4., 2021, Barcelona. *Anais* [...]. Barcelona: ACM, 2021. p. 610-623. DOI: <https://doi.org/10.1145/3442188.3445922>.
- BENJAMIN, Ruha. Assessing risk, automating racism. *Science*, Washington, D.C, v. 366, n. 6464, p. 421-422, 2019. DOI 10.1126/science.aaz3873.
- BERTONCINI, Ana Luize Corrêa; SERAFIM, Mauricio C. Ethical content in artificial intelligence systems: a demand explained in three critical points. *Frontiers in Psychology*, Lausanne, v. 14, p. 1074787, mar. 2023. DOI: <https://doi.org/10.3389/fpsyg.2023.1074787>.
- BIETTI, Elettra. From ethics washing to ethics bashing: a moral philosophy view on tech ethics. *Journal of Social Computing*, Piscataway, v. 2+, n. 3, p. 266-283, 2021.
- CONROY, Maggie. The ethics of slow AI: why taking time to develop technology matters. *Data Lab Notes*, Online, Appleton, 2023. Disponível em: <https://datalabnotes.com/slow-ai/>. Acesso em: 14 mar. 2024.
- COULDRY, Nick; MEJIAS, Ulises Ali. The decolonial turn in data and technology research: what is at stake and where is it heading? *Information, Communication & Society*, UK, v. 26, n. 4, p. 786-802, 2023. DOI: 10.1080/1369118X.2021.1986102.
- DAIR - DISTRIBUTED AI RESEARCH INSTITUTE. *Timnit Gebru launches independent AI research institute on anniversary of ouster from Google*. Oakland: DAIR, 2021. Disponível em: <https://www.dair-institute.org/press-release/>. Acesso em: 14 mar. 2024.
- DECOLONIZAI. *Sobre o projeto*. [2022]. Disponível em: <https://www.decolonizai.com/sobre-o-projeto/>. Acesso em: 30 jan. 2024.
- DIMENSIONS. *Linked research data from idea to impact*. [2024]. Disponível em: <https://www.dimensions.ai/>. Acesso em: 14 jan. 2024.
- FLORIDI, Luciano. *'Distributed morality': the ethics of information*. Oxford: Oxford Academic, 2014.
- FLORIDI, Luciano; SANDERS, Jeff W. On the morality of artificial agents. *Minds and Machines*, Berlin, v. 14, p. 349-379, 2004. Disponível em: <https://link.springer.com/article/10.1023/B:MIND.0000035461.63578.9d>. Acesso em: 14 mar. 2024.
- FOUCAULT, Martins. *Nascimento da biopolítica: curso dado no Collège de France (1978-1979)*. São Paulo: Martins Fontes, 2008.
- FOUCAULT, Michel. *História da sexualidade I: a vontade de saber*. Tradução de Maria Thereza da Costa Albuquerque e J. A. Guilhon Albuquerque. Rio de Janeiro: Edições Graal, 1988.
- FOUCAULT, Michel. *Microfísica do poder*. Rio de Janeiro: Edições Graal, 1993.
- FOUCAULT, Michel. Outros espaços. In: FOUCAULT, Michel. *Ditos e escritos*. Barueri: Forense Universitária, 2001. v. 3, p. 411-422.
- GARCIA-VIGIL, José L. Reflexiones en torno a la ética, la inteligencia humana y la inteligencia artificial. *Gaceta Médica de México*, Ciudad de México, v. 157, n. 3, p. 311-314, maio/jun. 2021. DOI: <https://doi.org/10.24875/gmm.20000818>.
- GLOBAL GRAND CHALLENGES. *Catalyzing equitable artificial intelligence (AI) use*. Seattle: Bill & Melinda Gates Foundation, 2023. Disponível em: <https://gcgh.grandchallenges.org/challenge/catalyzing-equitable-artificial-intelligence-ai-use>. Acesso em: 30 jan. 2024.
- GORDON, Lewis Ricardo. *Medo da consciência negra*. Tradução de José Geraldo Couto. São Paulo: Todavia, 2023.
- GORDON, Lewis Ricardo. *Freedom, Justice, and Decolonization*. New York: Taylor & Francis. 2021.
- HAO, Karen. We read the paper that forced Timnit Gebru out of Google: here's what it says. *MIT Technology Review*, Massachusetts, 4 dez. 2020. Disponível em: <https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/>. Acesso em: 4 jan. 2024.
- IPEA - INSTITUTO DE PESQUISA ECONÔMICA APLICADA. *Atlas da violência*. Brasília, DF: IPEA, [2023]. Disponível em: <https://www.ipea.gov.br/atlasviolencia/>. Acesso em: 14 mar. 2023.



MACIEL, Lucas da Costa. Perspectivismo ameríndio. In: ENCICLOPÉDIA de Antropologia. São Paulo: USP, 2019. Disponível em: <https://ea.fflch.usp.br/conceito/perspectivismo-amerindio>. Acesso em: 4 jun. 2024.

MAGNO, Gabriel; ARAÚJO, Camila Souza; MEIRA JUNIOR, Wagner; ALMEIDA, Virgílio. Stereotypes in search engine results: understanding the role of local and global factors. arXiv preprint, [s.l.], nov. 2016. Disponível em: [https://arxiv.org/search/?query=Stereotypes+in+search+engine+results%3A+understanding+the+role+of+local+and+global+factors&searchtype=all&abstracts=show&order=-announced\\_date\\_first&size=50](https://arxiv.org/search/?query=Stereotypes+in+search+engine+results%3A+understanding+the+role+of+local+and+global+factors&searchtype=all&abstracts=show&order=-announced_date_first&size=50). Acesso em: 4 jun. 2024.

MBEMBE, Achille. Necropolítica. *Arte & Ensaios*, Rio de Janeiro, n. 32, p. 122-151, 2016. DOI: <https://doi.org/10.60001/ae.n32.p122%20-%20151>

MBOA NKOUDOU, Thomas Hervé. We need a decolonized appropriation of AI in Africa. *Nature Human Behaviour*, London, v. 7, n. 11, p. 1810-1811, 2023. DOI: [10.1038/s41562-023-01741-3](https://doi.org/10.1038/s41562-023-01741-3).

MENDONÇA, Ricardo Fabrino; ALMEIDA, Emeritus Virgílio; FILGUEIRAS, Fernando. *Algorithmic Institutionalism: the changing rules of social and political life*. Oxford: University Press, 2024.

MILLER, Katharine. The movement to decolonize AI: centering dignity over dependency. *HAI - Institute for Human-Centered AI*, Stanford, 21 mar. 2022. Disponível em: <https://hai.stanford.edu/news/movement-decolonize-ai-centering-dignity-over-dependency>. Acesso em: 30 jan. 2024.

MOHAMED, Shakir; PNG, Marie-Therese; ISAAC, William. Decolonial AI: decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy and Technology*, Stanford, v. 33, n. 4, p. 659-684, 2020. Disponível em: <https://doi.org/10.1007/s13347-020-00405-8>. Acesso em: 4 jun. 2024.

NAS, Elen. *Arte eletrônica: elo perdido*. São Paulo: Amazon Kindle, 2020. *E-book*.

NAS, Elen. *Bioethics of nonpresence: body, philosophy and machines*. São Paulo: Amazon Kindle, 2021. *E-book*.

NAS, Elen. Descolonizar o conhecimento: a perspectiva de Lewis Gordon. *Desenvolvimento Social*, Montes Claros, v. 29, n. 2, p. 189-199, 2023a. DOI [10.46551/issn2179-6807v29n2p189-199](https://doi.org/10.46551/issn2179-6807v29n2p189-199).

NAS, Elen. O Manifesto das Coisas: apontamentos para liberalização das vozes suprimidas. *Aurora*. Revista de Arte, Mídia e Política, v. 16, n. 48, p. 5-20, 2023c. DOI: <https://doi.org/10.23925/1982-6672.2023v16i48p5-20>.

NAS, Elen; AZEVEDO, Telma; LONGH, Fernando; TERCEIRO, Luciana; VALENTE, Tânia. Future visions for a decolonized future of HCI: thick descriptions of a survey interaction to discuss the colonization of imagination. INTERNATIONAL CONFERENCE ON HUMAN-COMPUTER INTERACTION, Cham, p. 109-116, 2023b. DOI: <https://doi.org/10.17613/30bc-6j76>.

NASCIMENTO, Elen Cristina Carvalho. Reflexões bioéticas na era da inteligência artificial. In: CASTRO, João Cardoso; NIEMEYER-GUIMARÃES, Márcio; SIQUEIRA-BATISTA, Rodrigo (ed.). *Caminhos da bioética*. Teresópolis: Editora Unifeso, 2019. p. 345-362.

NASCIMENTO, Elen Cristina Carvalho. *Arte eletrônica: elo perdido entre ciência, design e tecnologia*. 2017. Dissertação (Mestrado em Design) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2017. Disponível em: <https://www.maxwell.vrac.puc-rio.br/30067/30067.PDF>. Acesso em: 14 mar. 2024.

NOBLE, Safiya Umoja. *Algorithms of oppression*. New York: New York University Press, 2018.

OECD – ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. *AI principles overview*. Paris: OCDE.AI, 2019. Disponível em: <https://oecd.ai/en/ai-principles>. Acesso em: 14 mar. 2024.

O'NEIL, Cathy. *Algoritmos de destruição em massa*. São Paulo: Editora Rua do Sabão, 2021.

O'NEIL, Cathy. *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: Crown Publishers, 2016.

ONU – ORGANIZAÇÃO DAS NAÇÕES UNIDAS. *Declaração universal dos direitos humanos*. Paris: ONU, 1948. Disponível em: <https://declaracao1948.com.br/declaracao-universal/declaracao-direitos-humanos/>. Acesso em: 14 mar. 2016.

RAJI, Inioluwa Deborah; CHOCK, Sasha Costanza; BUOLAMWINI, Drjoy. Change from the outside: towards credible third-party audits of AI systems. *MacArthur Foundation*, Chicago, 7 jun. 2023. Disponível em: <https://www.macfound.org/press/grantee-publications/outside-scrutiny-to-change-ai-systems>. Acesso em: 14 mar. 2024.

SIMONDON, Gilbert. *L'individu et sa genèse physico-biologique*. Grenoble: Éditions Jérôme Millon, 1995.

TADDEO, Mariarosaria; FLORIDI, Luciano. How AI can be a force for good. *Science*, Washington, D.C, v. 361, n. 6404, p. 751-752, 2018. DOI 10.1126/science.aat5991.

UNESCO – THE UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION. *Recommendations on the ethics of artificial intelligence*. Paris: UNESCO, 2023. Disponível em: <https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence>. Acesso em: 14 mar. 2024.

ZUBOFF, Shoshana. *The age of surveillance capitalism: the fight for a human future at the new frontier of power*. New York: Public Affairs, 2019.

#### Minicurrículo da Autora:

**Elen Nas**. Doutora em Bioética, Ética Aplicada e Saúde Coletiva pela Universidade Federal do Rio de Janeiro (2021). Pesquisadora de pós-doutorado da Cátedra Oscar Sala do Instituto de Estudos Avançados da Universidade de São Paulo. Pesquisa financiada pelo Comitê Gestor da Internet no Brasil/IEA-USP (Projeto 642). E-mail: [elennas@usp.br](mailto:elennas@usp.br).

Avaliador 2: Icaro Ferraz Vidal Junior, [Orcid](#);  
Editora de Seção: Raquel Kritsch, [Orcid](#).