

DOI: 10.20911/21799024v15n1p10/2024

Deveríamos ser mais compreensivos? Uma reflexão acerca da compreensibilidade das injustiças advindas da aplicação das IAs

**SHOULD WE BE MORE UNDERSTANDING?
A reflection on the comprehensibility of injustices arising from the
application of AIs**

Vinicius Schoenell dos Santos¹

Resumo: A rápida evolução da Inteligência Artificial (IA) levanta questões éticas sobre como interagimos com essas entidades digitais em constante desenvolvimento. Entretanto, é possível, desde já, como observam Himmelreich e Lim, que as ferramentas de IA possam perpetrar preconceitos e injustiças já conhecidas. Assim, o desenvolvimento e melhorias das IAs passam por um crivo relativo ao tempo da aplicação dessas ferramentas, que, muitas vezes, por consequência do prazo de adaptação desses sistemas, acabam sendo utilizadas como motores de difusão desses preconceitos já existentes, por uma contingência relativa ao período dessa implementação. Este artigo propõe que não deveríamos adotar uma abordagem mais compreensiva em relação às IAs, especialmente durante as fases iniciais de seu desenvolvimento, e que deveríamos, desde seus primeiros modelos, pensar em cenários complexos e considerar o máximo de cenários negativos que gerem injustiças, com a finalidade de tentar remediar que as IAs acabem por concretizar tais situações.

¹ Mestrando em Filosofia na Universidade do Vale do Rio dos Sinos (UNISINOS); bolsista Capes/Proex. E-mail: bep.vss@hotmail.com.

Palavras-chave: Inteligência Artificial; Ética; Compreensão; Injustiça; Injustiça Estrutural.

Abstract: The rapid evolution of Artificial Intelligence (AI) raises ethical questions about how we interact with these constantly developing digital entities. However, it is already possible, as Himmelreich and Lim observe, that AI tools can perpetrate prejudices and injustices that are already known. Thus, the development and improvements of AIs go through a sieve relating to the time of application of these tools, which, often, because of the adaptation period of these systems, end up being used as engines for the diffusion of these already existing prejudices, due to a relative contingency the period of this implementation. This article proposes that we should not adopt a more comprehensive approach towards AIs, especially during the initial phases of their development, and that we should, from the first AI models, think about complex scenarios and consider as many negative scenarios that generate injustices as possible, with the purpose of trying to remedy the fact that AIs end up creating such situations.

Keywords: Artificial Intelligence Ethic; Understanding; Injustice; Structural Injustice.

Introdução

O campo da Inteligência Artificial (IA) está em constante evolução, com algoritmos e modelos sendo aprimorados continuamente. Nesse contexto, é razoável assumir que as IAs estão em um processo de aprendizado constante, muitas vezes semelhante ao aprendizado humano. Sistemas que operam dados de diversas formas são, a cada dia, implementados na vida cotidiana, nas repartições públicas e nas redes privadas, usualmente com a intenção de que se obtenha uma maior eficiência nos resultados desejados. Os sistemas que estão em itens básicos e que são considerados (por muitas vezes) essenciais à vida humana como computadores, celulares, televisões etc. já fazem uso da IA para desempenhar suas funções de forma mais eficiente.

Possivelmente, a algumas décadas atrás, quando se pensava em IAs, principalmente no campo especulativo da ficção científica, se imaginava algo semelhante a um assistente, com certa *persona* humana e características como voz, ou um corpo metalizado semelhantes ao corpo humano. Isaac Asimov (2014, p. 65), na obra *Eu, robô*, estabeleceu que haveria três leis fundamentais à robótica que devem ser contempladas²³:

2 Asimov pensou em um cenário ficcional onde robôs coexistiriam com humanos, auxiliando nas tarefas e nos ofícios que são, historicamente, responsabilidades humanas. Entretanto, as leis de Asimov podem, especulativamente, ser expandidas ao domínio aplicado da IA, uma vez que parecem funcionar, também, em certas situações pertinentes advindas desse tipo de sistema.

3 Na ficção, um robô não necessariamente precisa contar com uma capacidade deliberativa. Podem existir autômatos,

1. Um robô não pode ferir um ser humano, ou, por omissão, permitir que um ser humano sofra algum mal;
2. Um robô deve obedecer às ordens que lhe sejam dadas por seres humanos, exceto nos casos em que tais ordens contrariem a Primeira Lei;
3. Um robô deve proteger sua própria existência enquanto tal proteção não entrar em conflito com a Primeira ou Segunda Leis.

Dessas leis de Asimov, não necessariamente se segue que tais normas são necessárias para a existência de robôs ou IAs, mas tais ideias do autor sugerem que há certos princípios nos quais se deve pensar quando nos relacionamos (como seres humanos), com entidades artificiais. Os seres humanos tomam, para si, a ideia de que são a espécie dominante do planeta a muito tempo. Essa centralização do entendimento humano como espécie “superior” ou “subjugadora” das demais evidencia uma série de perspectivas que são próprias da humanidade. James Rachels (1990, p. 184), defendeu que os interesses dos seres humanos como espécie são considerados prioritários não simplesmente por sua humanidade, mas devido a características morais exclusivas que não são comuns com outros animais. Peter Singer (1994, p. 87) enfatizou, como uma possível solução a esse viés *especista*⁴, que “a ética não exige que eliminemos as relações pessoais e as afeições parciais, mas exige que, em nossas ações, levemos em conta as reivindicações morais dos que são afetados por elas, e que o façamos com um certo grau de independência” (SINGER, 1994, p. 87), *i.e.*, que não é porque a humanidade possui certa isonomia em relação a outras espécies que não se deva agir com responsabilidade quando as ações humanas afetam as demais espécies, mesmo que julgadas inferiores.

As normas humanas, pressupondo que a humanidade, de fato, aja de tal forma, como considerado por James e Singer, de modo a preservar os indivíduos humanos (não considerando aqui as normas que visam preservar outros indivíduos, pois, embora tais normas existam, não vem ao caso compará-las ao número de normas e regras morais que são voltadas exclusivamente para a humanidade), são constituídas nos âmbitos jurídicos, moral, interrelacional, social etc., de modo a garantir aos humanos a possibilidade de uma existência. Dessa forma, ao longo do desenvolvimento humano e da sociedade, surgiram diversas doutrinas filosóficas e teorias, muitas delas norteadas por fortes princípios, que se propuseram a explicar ou indicar a forma mais harmoniosa ou correta dos

como o *The writer* de Jaquet Droz, criado no século XVI (IGOR, 2018), que seriam formas mais primitivas de robôs, que funcionam com mecanismos similares aos de relógio, e que podem ser movidos a cordas, pilhas ou bateria, e executam tarefas simples, de forma repetitiva e objetiva. Entretanto, também, há o que se tomou por robôs mais “conscientes”, *i.e.*, robôs que possuiriam certa capacidade deliberativa, que executariam tarefas nas quais envolveria certos “raciocínios”, como é o caso do robô *C-3PO*, que, no universo ficcional de *Star Wars*, é um robô (*droid*) de protocolo, cuja função é facilitar a comunicação entre diferentes espécies e culturas, atuando como intérprete e facilitador em negociações e encontros diplomáticos, pois, geralmente, são fluentes em várias línguas e programados para entender nuances culturais, o que os torna úteis em situações em que a comunicação eficaz é crucial. O robô *R2-D2*, também de *Star Wars*, cuja função é ser um droide astromecânico, tem por designação fazer manutenção de naves, atuar na navegação e computação e possuem habilidades mecânicas e versatilidade em operar em situações diversas. Também é um exemplo de robô imaginado como um ente deliberativo, porém com uma função menos social, mais voltada para a execução de tarefas objetivas de reparo e controle.

4 O especismo, pode ser definido, segundo Singer (2004, p. 6) como “um preconceito ou atitude de favorecimento dos interesses dos membros de uma espécie em detrimento dos interesses dos membros de outras espécies”.

seres humanos agirem. A *ética das virtudes*, o *utilitarismo* e a *deontologia* Kantiana são exemplos dessas teorias (CRUZ, 2020, p. 592).

Entretanto, com o surgimento emergente das tecnologias baseadas em IA, essas teorias não parecem se fazer suficientes para explicar todos os fenômenos relacionados com esse novo tipo de aparato tecnológico. Savulescu *et al* (2021) propôs que o princípio que se deveria usar perante essas tecnologias emergentes é a de um *equilíbrio reflexivo coletivo prático*, que teria como objetivo principal fornecer resultados mais equilibrados, quando as teorias principialistas mais relevantes são divergentes e não conseguem apresentar uma resposta imediata ao problema.

Assim, seria plausível que, quando pensamos nas IAs, cheguemos a conclusões “personalistas”, do tipo que possam defender que assim como somos compreensivos com o aprendizado humano e suas imperfeições iniciais, deveríamos aplicar uma abordagem semelhante às IAs emergentes. Porém, a isso, argumentar-se-á que não devemos incorrer nesse tipo de raciocínio, a pena de acabarmos por perpetrar mais ainda as injustiças sociais (como o racismo, a LGBTfobia, a misoginia, a xenofobia etc.) já colocadas, prejudicando, no processo, grupos que já são vulneráveis *per se* na atual sociedade em que vivemos.

1. A Injustiça Estrutural e a IA

Himmelreich e Lim (2022) procuraram compreender como as IAs interagem com as injustiças estruturais da sociedade. Como definição de injustiça estrutural, os autores afirmam que as desigualdades não são meramente individuais, mas são profundamente enraizadas nas estruturas da sociedade e requerem uma análise crítica e ação para mudar essas estruturas a fim de alcançar uma sociedade mais justa e igualitária⁵.

Os autores explicitam que há no norte global (*i.e.*, países como EUA e países europeus) uma crescente tomada de consciência que o *status quo* se baseia e manifesta várias injustiças normativas passadas e atuais. Essa tomada de consciência está levando os pesquisadores e políticos a prestarem atenção no papel que a IA desempenha na manutenção, consolidação ou mesmo exacerbação deste status quo injusto.

Um software ou algoritmo pode ser apenas uma ferramenta objetiva, mas, se aplicado a um contexto injusto, é provável que esse tipo de ferramenta perpetue ou agrave injustiças. Uma IA “interage com estruturas sociais injustas – a IA agrava a injustiça estrutural” (HIMMELREICH e LIM, 2022, p. 2).

Os autores pretendem mostrar que a injustiça estrutural é conceito normativo, empiricamente respeitável e, como tal, o conceito de injustiça estrutural deve ser aplicado e estar presente nas investigações e na elaborações de leis

⁵ Essa definição fornecida por Himmelreich e Lim provém, principalmente, da obra de Iris Marion Young, *Structural Injustice and the Politics of Difference*, de 2011.

que tratam de questões referentes a IA, como a forma que a IA interage com as identidades sociais (como gênero ou raça), ou como o desenvolvimento da IA reflete os interesses econômicos existentes. Uma teoria da injustiça estrutural permite aos investigadores e profissionais identificar, articular e talvez até antecipar fenômenos e problemas que de outra forma poderiam passar despercebidos. Alguns elementos teóricos básicos que compõem a injustiça estrutural estendem-se facilmente do gênero e raça às questões de capacidade, identidade sexual ou poder econômico (HIMMELREICH e LIM, 2022, p. 3).

Os autores, então, procuram compreender quais considerações teóricas e morais servem de fundamento para uma preocupação com os valores de diversidade, igualdade e inclusão (DEI) e, a partir de exemplos baseados no contexto norte-americano, procuram exemplificar e fornecer uma base conceitual para mostrar como a IA se relaciona com a injustiça estrutural. Com isso, procuram contribuir para as discussões sobre DEI ao conceituar os apelos atuais por diversidade, equidade e inclusão na esfera da IA como exigências de justiça, em vez de uma tentativa de conter danos ou tentativas de cumprir códigos de conduta éticos que são frequentemente elaborados por corporações de tecnologia.

Um dos exemplos analisados por Himmelreich e Lim é o dos hospitais que utilizam regras de decisão (algoritmos) para identificar pacientes com grandes necessidades, (*i.e.*, pacientes com doenças crônicas e necessidades médicas complexas) e convidá-los para programas especiais de tratamento ambulatorial.

Em 2019, os pesquisadores descobriram que um algoritmo para identificar esses pacientes com grandes necessidades apresentou “viés racial significativo” (Obermeyer *et al.*, 2019): embora mais de 46% dos pacientes negros devessem ter recebido a ajuda de tais programas de tratamento, apenas 17,7% receberam. Os pacientes negros tinham menos probabilidade de serem selecionados para programas de tratamento especial, apesar de não haver diferença nas condições de saúde subjacentes (HIMMELREICH e LIM, 2022, p. 4).

O algoritmo fez uma correlação entre saúde e despesa. O sistema de saúde, em geral gasta menos com a saúde dos negros, logo, o algoritmo desenvolveu a tendência de subestimar a necessidade de tratamento para negros em casos em que há grandes necessidades médicas. Dessa forma, a IA ignorou a necessidade de tratamentos de saúde de pessoas negras e, consequentemente, acabou contribuindo (ou até mesmo agravando) para um cenário de que gerou (ou reproduziu) uma injustiça estrutural (no caso, injustiça racial).

Entretanto, o racismo estrutural não é algo necessariamente deliberado, podendo se manifestar sem que aquele que fala tenha consciência disso, pois advém de uma estrutura social e de práticas que se perpetuam ao longo do tempo:

Muitas pessoas ainda acreditam que a raça é uma essência biocomportamental que explica as nossas disposições comportamentais e culturais (Zack, 2014) e operam com base neste pressuposto quando lidam com membros de outros grupos raciais. Outros insistem que a raça é constru-

ída socialmente, e não natural ou biológica. No entanto, tal como outras construções sociais (por exemplo, o dinheiro), a raça tem efeitos materiais profundos na nossa experiência vivida. “A raça é”, para citar Charles W. Mills, “uma realidade contingentemente profunda que estrutura o nosso universo social particular” (1998, p. 48) (HIMMELREICH e LIM, 2022, p. 4).

Desta forma, mesmo quando o *locus* do mecanismo causal que explica as diferenças nos gastos médicos (a partir de concepções raciais) está localizado nos indivíduos (os profissionais médicos) as propriedades que são fornecidas a esse mecanismo são estruturais: suposições, hábitos de pensamento, estereótipos, preconceitos, normas discursivas, ou generalizações sobre pacientes negros.

Outra hipótese ressaltada pelos autores é a de que pessoas negras confiariam menos no sistema de saúde. Logo, não frequentariam tanto clínicas e hospitais e, por consequência, os gastos com pessoas negras seriam menores. Um dos argumentos usados para justificar essa falta de confiança seria o *estudo de Tuskegee*⁶ (que é uma explicação não estrutural). Porém, em contrapartida, Himmelreich e Lim (2022, p. 6) afirmam que outros pesquisadores apontam que essa justificativa não é suficiente para explicar o fenômeno da falta de confiança, uma vez que o estudo não é suficientemente conhecido pela população negra no geral. Outras hipóteses consideram, também, estatuto socioeconômico, normas de gênero, a falta de consciência das necessidades de cuidados de saúde, as atitudes religiosas e espirituais em relação à medicina, bem como os antecedentes criminais (CHEATHAM *et al.*, 2008).

Esses exemplos da IA aplicado no contexto da saúde sugere que há sim uma conexão relevante entre a as injustiças estruturais (no sentido amplo) e a forma como se operacionaliza as IAs.

2. Deveríamos ser compreensivos com as injustiças geradas no período de implementação das IAs?

Mesmo considerando o exemplo fornecido por Himmelreich e Lim, parece necessário uma melhor consideração sobre a aceitabilidade do processo de implementação das IAs e da compreensão necessária que se demanda frente a isso. Pode-se, inicialmente, considerar algumas afirmações para pensarmos sobre essa questão, sendo elas: (a) as IAs são implementadas na sociedade den-

6 Estudo de Tuskegee foi um estudo clínico realizado nos Estados Unidos entre 1932 e 1972 pelo Serviço de Saúde Pública dos Estados Unidos (USPHS), na cidade de Tuskegee, Alabama, e envolveu principalmente homens afro-americanos de baixa renda. O objetivo do estudo era observar a progressão da sífilis não tratada em seres humanos, no intuito de entender melhor a doença. Como consequência, o estudo foi considerado profundamente antiético e imoral, uma vez que os participantes não foram informados adequadamente sobre a sua natureza e não receberam tratamento adequado para suas condições de saúde. Os participantes foram enganados a acreditar que estavam recebendo cuidados médicos gratuitos quando, na verdade, não estavam sendo tratados para a sífilis. O estudo continuou por 40 anos, mesmo depois que a penicilina se tornou um tratamento eficaz para a sífilis na década de 1940. Como resultado, muitos dos participantes sofreram danos graves à saúde, incluindo danos neurológicos, cegueira, deformidades e morte. Esse estudo foi amplamente condenado quando suas práticas antiéticas foram reveladas publicamente na década de 70, o que acarretou mudanças significativas nas regulamentações e práticas de pesquisa médica nos Estados Unidos (GAMBLE, 2006; CDC..., 2022).

tro de um tempo que começa em t_1 e se estende até o presente contínuo, (b) existem preconceitos e injustiças que são estruturais da sociedade, bem como outros que são criados a partir de contingências, (c) as IAs não tem, por natureza, um caráter injusto, porém a forma com que esses sistemas são manejados, usados e implementados podem ser pervertidos a gerar injustiças de modo *a posteriori*, (d) ninguém deve ser vítima de injustiça como um “sacrifício” para que as IAs sejam melhoradas futuramente.

Quando afirma-se (a), está se pressupondo que as IAs, como qualquer outro sistema, ou novidade tecnológica, não existia (na vida social) antes do tempo t_1 e passa a existir consequente a t_1 . Disso, pode-se inferir, de forma muito breve, que como todo novo recurso que passa a existir na sociedade as IAs demandam um momento de implementação e, consequentemente, a partir dessa implementação em t_1 , um momento subsequente de adaptação. Assim, as IAs, uma vez implementadas, entram em contato com a sociedade e, por consequência, com o produto gerado pela interação social e pela tecitura estrutural que a permeia. Logo, é possível afirmar que (b), uma vez que reconhecemos que existem injustiças, principalmente, que estão intrínsecas a estrutura da sociedade em que se vive, mas que tais injustiças não são eternas e perenes, e que podem se dissolver ao longo do tempo, da mesma forma que podem surgir novas formas de injustiça ao decorrer do tempo, não garantindo um status imutável às injustiças. Assim, é natural, ou minimamente orgânico, que a nova tecnologia, no caso as IAs, se permeiem por essas injustiças e as reproduzam, uma vez que estamos falando de sistemas que são capazes de aprender e cruzar dados de forma (até mesmo) similar a forma humana.

Destarte, como as IAs “aprendem” a perpetrar essas injustiças, é possível que se afirme (c), uma vez que não possuem natureza hobbesiana⁷, isto é, não são más ou egoístas por natureza, mas permeadas por preconceitos humanos que se interpelam a partir dos dados e comandos que são fornecidos a essas IAs.

Então, se considerarmos a ideia de temporalidade expressa em (a), a ideia de que há injustiças sociais estruturais expressa em (b) e que as IAs não possuem natureza hobbesiana, como expresso em (c), podemos, por último, pensar em (d). Pessoas e grupos que são socialmente mais vulneráveis, como apresentado nas considerações de Himmelreich e Lim, podem ser vitimadas no processo de adaptação dessas IAs.

Por vitimadas, não necessariamente se afirma que esses indivíduos ou grupos passam a ser vítimas de novas injustiças, mas que são “revitimadas” por um sistema que é, na grande maioria das vezes, produzido por aqueles que possuem, quando em relação ao grupo minoritário, o status de vitimadores, e acabam por ser mantidas em uma posição que pode ser considerada negativa nesse novo processo.

7 A expressão “natureza hobbesiana”, aqui, se refere a uma visão de mundo inspirada nas ideias do filósofo político inglês Thomas Hobbes, conhecido por sua obra “Leviatã”, na qual ele descreve a natureza humana como inerentemente egoísta, competitiva e propensa ao conflito. De acordo com Hobbes, na “natureza do homem”, em um estado de natureza não regulamentado pelo governo, as pessoas agiriam em busca de seus próprios interesses, o que levaria a um estado de guerra de todos contra todos.

Considerações finais

As IAs devem, desde sua criação, serem tomadas como um tipo de tecnologia que necessita de certa atenção social e ética, para que não incorra na reprodução das injustiças estruturais que já existem na sociedade. Os indivíduos e grupos minoritários, por vez, não deveriam ser “revitimados” por uma tecnologia que é implementada de maneira nova ou revolucionária, a fim de que, apenas “futuramente” se observe essas injustiças na prática e se calcule o dano produzido por tal tecnologia. Neste caso, as IAs (ou tecnologias novas de grande impacto na vida social cotidiana) devem ser entendidas, desde o início, como possíveis perpetradoras de injustiças estruturais e devem ter o olhar atencioso dos eticistas e teóricos desde quando potência, ainda antes de ser colocada em execução na sociedade ou em seus estágios iniciais.

Quando se implementa uma nova ferramenta de IA ou uma nova tecnologia ou sistema sem se preocupar com esse tipo de impacto reprodutor (ou gerador) de injustiças, mesmo que involuntariamente, os indivíduos que fazem parte de grupos minoritários historicamente violentados se tornam “sacrifícios”⁸. Por sacrifícios, entende-se que os indivíduos estão em uma posição não de serem descartados, mas oferecidos como oferenda a uma ideia de progresso que, futuramente, pode gerar frutos positivos para esses grupos minoritários. Se hoje, como no exemplo abordado por Himmelreich e Lim, os negros recebem menos convites para participar dos tratamentos que necessitam, eles estão sendo sacrificados, para que, futuramente, o problema seja identificado postumamente e resolvido, mesmo que hoje já se tenha a ciência que as IAs acabam por reproduzir comportamentos discriminatórios ou injustos humanos.

Dessa forma, parece ser coerente uma defesa mais “radicalizada” de uma posição que não tolere ou que não seja muito compreensiva com a possibilidade de haver vítimas no processo de implementação de uma IA. Se hoje é possível identificar que as IAs podem acabar sendo induzidas ou acabarem reproduzindo injustiças, preconceitos, estereótipos negativos e etc., então não devemos ser compreensivos que elas, em seu processo de adaptação social, executem essas injustiças, uma vez que já somos capazes de ter consciência de que raça, gênero, orientação sexual, diversidade étnica e status econômico são parte de problemas estruturais da sociedade e que a uma notável parcela das pessoas é capaz de julgar que são problemas injustos, então deve-se atentar a esses problemas desde o momento que se gesta a tecnologia, e se atentar para as possíveis consequências práticas de injustiça que esses grupos podem acabar sofrendo.

Assim, não deveríamos ser compreensivos, como sociedade, com a negligência desses problemas estruturais quando implementadas novas tecnologias, principalmente tecnologias com funções deliberativas, como no caso das IAs, e, principalmente, não devemos ser compreensivos com o sacrifício de uma dignidade pertencente aos grupos minoritários em detrimento de um desenvolvimento de sistemas que optam por não se preocupar com esse tipo de injustiça, mes-

8 Sacrifício, nesse contexto, deve ser tomado da forma mais religiosa e tribalista do termo.

mo que as vezes o discurso hegemônico possa ser o de ser mais compreensivo com as imperfeições iniciais dessas entidades digitais, pois elas pode resultar em avanços mais éticos e responsáveis, proporcionando benefícios significativos à sociedade à medida que se estabelecem e evoluem. Como ressalta Pereira (2010, p. 11), “a tolerância necessita de policiamento”. Se não formos capazes (como sociedade) de ser menos tolerantes com a produção de injustiças que decorrem do processo de adaptação da nova tecnologia, deveríamos, ao menos, ser bastante vigilantes acerca dos resultados negativos que essa tecnologia nova produz socialmente.

Referências

ASIMOV, Isaac. *Eu, robô*. São Paulo: Aleph, 2014.

CDC - CENTERS FOR DISEASE CONTROL AND PREVENTION. The Untreated Syphilis Study at Tuskegee Timeline. *CDC - Centers for Disease Control and Prevention*. [S.l.], 2022. Disponível em: <<https://www.cdc.gov/tuskegee/timeline.htm>>. Acesso em: 20 dez. 2023.

CHEATHAM, C. T., BARKSDALE, D. J., & RODGERS, S. G. Barriers to health care and health-seeking behaviors faced by Black men. *Journal of the American Academy of Nurse Practitioners*, [s. l.], v. 20, ed. 11, p. 555–562, 2008.

CRUZ, Jorge Silva. Ética das virtudes: em busca da excelência. *Revista de Medicina*, São Paulo, v. 99, n. 6, p. 591-600, 2020.

GAMBLE, Vanessa Northington. Under the Shadow of Tuskegee: African Americans and Health Care. *Health Psychology*, [s. l.]: Routledge, 2006.

HIMMELREICH, Johannes. LIM, Désirée. *AI and Structural Injustice: Foundations for Equity, Values, and Responsibility*. The Oxford Handbook of AI Governance, Oxford, 2022.

IGOR, Leonardo. *Os primeiros robôs da humanidade: autômatos capazes de escrever, desenhar e tocar foram construídos no século 18*. 2021. Disponível em < <https://www.opovo.com.br/noticias/curiosidades/2021/03/07/os-primeiros-robos-da-humanidade--automatos-capazes-de-escrever--desenhar-e-tocar-foram-construidos-no-seculo-18.html>>. Acesso em: 17 dez. 2023.

PEREIRA, Luíz Moniz. Tolerância evolucionária. *Revista Portuguesa de Psicanálise*, Portugal, v. 30, n. 2, p. 117-147, 2010.

SAVULESCU, Julian; GYNGELL, Christopher; KAHANE, Guy. Collective Reflective Equilibrium in Practice (CREP) and controversial novel technologies. *Bioethics*, Filipinas, v. 35, p. 652-663, 2021.