

## ARTIGOS

Valmir Luis Saldanha da Silva<sup>1</sup>

### Identidade algorítmica: imagens da negritude e viés racista

Algorithmic identity: images of blackness and racist bias



#### RESUMO:

As inteligências artificiais generativas e seus algoritmos são responsáveis por cada vez mais decisões em nossa sociedade. No entanto, por mais que sejam tidos como ferramentas imparciais, os dados de algoritmos com viés racial que prejudicam pessoas negras em diversas áreas da vida não param de crescer. A este fenômeno demos o nome de Identidade Algorítmica. Com base nisso, e com o entendimento de que a oposição real X fictício não se sustenta, tendo de ser trocada pela tríade real-fictício-imaginário, neste estudo demonstramos os problemas relativos aos vieses cognitivos que norteiam a construção de identidades que saem do mundo off-line e adentram o mundo on-line, principalmente pela subalternização das imagens de pessoas negras. Além disso, indicamos como as identidades algorítmicas estereotipadas colocam-se como uma barreira à efetiva inclusão sociodigital. Como possibilidade de enfrentamento, propomos a criação de um material didático voltado para o letramento sobre argumentação, elementos implícitos e vieses algorítmicos, com foco nos anos finais do Ensino Médio e nos anos iniciais de cursos universitários.


**Palavras-chave:** Identidade algorítmica; Racismo algorítmico; Letramento digital; Inclusão sociodigital; Identidade negra

#### ABSTRACT:

Generative AI and its algorithms are responsible for more and more decisions in our society. However, despite being considered impartial tools, the data from algorithms with racial bias that harm black people in various areas of life continues to grow. We have given this phenomenon the name Algorithmic Identity. Based on this, and with the understanding that the opposition between real and fictional does not hold up and must be replaced by the triad real-fictional-imaginary, in this study we demonstrate the problems related to the cognitive biases that guide the construction of identities that leave the offline world and enter the online world, mainly due to the subordination of images of black people. In addition, we indicate how stereotypical algorithmic identities pose a barrier to effective social and digital inclusion. As a possible way to address this, we propose the creation of teaching material focused on literacy about argumentation, implicit elements and algorithmic biases, focusing on the final years of high school and the initial years of university courses.

**Keywords:** Algorithmic identity; Algorithmic racism; Digital literacy; Social and digital inclusion; Black identity

<sup>1</sup> Doutor em Estudos Literários pela Universidade Estadual Paulista "Júlio de Mesquita Filho"; Professor, Instituto Federal de Educação Ciência e Tecnologia de São Paulo, Campos do Jordão, SP, Brasil.

valmir.saldanha@ifsp.edu.br,  <https://orcid.org/0000-0002-5605-4409>

## INTRODUÇÃO

Talvez a questão mais importante do século XXI seja a identidade. Pensemos desde os ataques às Torres Gêmeas, em 11 de setembro de 2001, nos Estados Unidos, que tanto modificaram a narrativa sobre pessoas muçulmanas e árabes quanto revolucionaram a maneira como o chamado “ocidente” (Said, 2013) passou a repensar as noções de segurança, controle de informação e privacidade. Ou a criação da União Europeia em janeiro de 2002, quando 12 países do bloco deixaram de ter uma moeda própria e passaram a fazer transações com o Euro. Como consequência, os cidadãos europeus mantiveram suas nacionalidades, mas agora também se compreendiam como grupo.

Ainda, a *Primavera Árabe* do fim de 2010, no Oriente Médio e no norte da África, como resistência civil a governos totalitários também tinha um forte componente identitário: a afirmação de modos de ser árabe em disputa. Esses embates geraram guerras civis como as do Líbano e da Síria, em 2011, que, dentre as várias consequências nefastas, fizeram aumentar o número de refugiados no mundo (Silva, 2017). Ou a chegada do Papa Francisco ao Vaticano, como primeiro pontífice sul-americano. A aprovação de uma lei na Holanda que permitia o chamado “casamento gay”. A eleição do primeiro presidente negro nos Estados Unidos. A aprovação da lei de cotas para pessoas pre-

tas, pardas e indígenas, bem como para pessoas oriundas de escolas públicas poderem ingressar em universidades públicas, no Brasil. Bem como as reações a essas políticas, com a eleição de políticos de extrema-direita nos Estados Unidos e no Brasil, bem como a ascensão do partido radical Alternativa para a Alemanha (AfD), com políticas antirrefugiados, anti-LGBTQ e antigoverno (DW, 2024). Todas essas são ações cujo ponto de partida e/ou de chegada tem relação direta com a identidade, o que, se não comprova, ao menos indica de que forma pensar nas questões identitárias e de identidade é uma demanda urgente do século XXI.

Contudo, componentes ainda miticamente enxergados como neutros e imparciais, tais quais os algoritmos que regem o mundo digital e as inteligências artificiais, vêm passando longe de serem analisados sistematicamente por um vocabulário oposto às práticas sociais racistas, machistas, capacitistas etc. Há um largo caminho a ser trilhado para enfrentar este problema, a despeito de trabalhos pioneiros e bem estruturados, como o representado pelas pesquisas de Tarcízio Silva (2021; 2022) que culminaram, por exemplo, na criação de uma *timeline* com o mapeamento dos danos causados pela “discriminação algorítmica” (Silva, 2024). A ideia do autor foi compilar casos, reportagens e reações sobre essa discriminação, a fim de fomentar o debate para, em seguida, proporcionar uma reação coletiva aos danos observados e po-

tenciais de sistemas algorítmicos.

A partir disso, neste trabalho discutiremos de que forma essas discriminações e esses racismos algorítmicos acabam por propor uma *identidade algorítmica*. Em seguida, identificaremos possíveis formas de promover a equidade neste contexto, principalmente pela organização, tradução e disponibilização de materiais didáticos, cuja finalidade seja a de proporcionar a capacitação das pessoas para melhor compreender a lógica da comunicação analógica e da comunicação mediada por inteligência artificial e algoritmos.

## OBJETIVOS

Como apresentamos, o ponto nodal deste trabalho deriva da identidade, mas encontra sustentação em um outro evento também do século XXI que, do ponto de vista aqui assumido, mudou a configuração do que se entende por comunidade e comunicação e, por conseguinte, também do que se entende por identidade: a criação do Facebook, em 2004. A ideia dessa rede social era a de conectar os amigos da *vida real* também no *mundo digital*, ou seja, ampliar as possibilidades de contato entre as pessoas que já se conheciam. Cris Hughes, co-fundador da plataforma, em 2006, deu a seguinte declaração: “Você não está indo online para conhecer, de forma aleatória, alguém que vive a 8 mil milhas de você. Você está entrando para ver

informações sobre pessoas que já são importantes para você.” (BBC, 2011). Como veremos, principalmente com os algoritmos que estão na base da rede social TikTok (Su *et al.*, 2021), buscar pessoas que eram importantes parece não ser mais o centro de atenção das redes, o que ficou claro desde o artigo *Attention is all you need* (Vaswani *et al.*, 2017), publicado por pesquisadores da Google.

Esses e outros algoritmos são, na verdade, modelos matemáticos. Ou, como define Cathy O’Neil (2020), “armas de destruição matemáticas” (ADMs). Dessa forma, boa parte da interação que define nossos relacionamentos e, em última instância, a identidade que carregamos são constituídos tanto por nossas individualidades biopsicofisiológicas quanto por dados estatísticos com os quais nos relacionamos em forma de textos, imagens e elementos multimodais e que temos de compreender e interpretar, atribuindo-lhes sentido. O’Neil não indica que essas ADMs são criadas com o propósito único de excluir ou segregar. Na verdade, o que ela demonstra é que estes modelos matemáticos muitas vezes não têm, por óbvio, todos os dados necessários para entender o comportamento que estão colocando em análise:

Então, trocam por dados substitutos – ou proxies, indicadores aproximados. Eles traçam correlações estatísticas entre o CEP residencial de alguém ou padrões de linguagem e seu potencial de pagar um empréstimo ou conseguir

dar conta de um emprego. Essas correlações são discriminatórias e, algumas delas, ilegais. (O’Neil, 2020, p. 33).

Tomando isso como verdade, os modelos matemáticos que compõem a formulação de algoritmos atuam de maneira aproximada a como a Teoria dos Modelos Organizadores do Pensamento (TMOP) compreende o raciocínio, já que “Não há, portanto, raciocínio sem modelos, porque isso suporia que se raciocina no vazio, sem dados significativos, e sem eles é impossível estabelecer relações, já que não há nada a relacionar.” (Moreno Marimón, Sastre, 2014, p. 110). No livro *Algoritmos de destruição em massa*, Cathy O’Neil (2020, p. 41), por exemplo, compara o racismo a um modelo matemático binário simples:

Racismo, no nível individual, pode ser visto como um modelo de previsão zunindo em bilhões de mentes humanas ao redor do mundo. É construído a partir de dados incompletos, defeituosos ou generalizados. Quer venha de experiências próprias ou de boatos, os dados indicam que certos tipos de pessoa se comportaram mal. Isso gera uma previsão binária de que todas as pessoas daquela raça irão se comportar da mesma forma.

Desnecessário dizer, racistas não gastam muito tempo buscando dados confiáveis para treinar seus modelos deturpados. E uma vez que esse modelo se transforma em crença, torna-se programado, encucado. [...] Portanto, o racis-

mo é o mais desleixado dos modelos de previsão. Ele é alimentado por coleta irregular de dados e correlações espúrias, reforçado por injustiças institucionais e contaminado por viés de confirmação. Desse modo, curiosamente, o racismo opera como muitas da ADMs [...] descritas neste livro.

Com base nesses dados, o objetivo central de discussão neste trabalho é, portanto, a relação entre a constituição das identidades, principalmente a de pessoas negras, a ascensão do mundo digital e tecnológico como repositório dessas mesmas identidades e a necessidade de termos um letramento voltado para interpretação de textos, a fim de sermos capazes de perceber a presença e a atuação de sistemas algorítmicos nos aplicativos, plataformas e serviços digitais e de saber como esses sistemas funcionam, observar seu alcance, suas possibilidades e sermos capazes de moldá-los. Como proposta prática derivada desse objetivo central, sabendo da dificuldade em eliminar situações de racismo cotidiano que nos assolam e sabendo que as noções de inclusão social e inclusão digital estão intimamente relacionadas na contemporaneidade (Bolzan; Löbler, 2016), propomos a criação de um material didático voltado para a formação de leitores, leitoras e intérpretes de texto que levem em consideração não apenas os aspectos gramaticais que estruturam as mensagens veiculadas, mas também a formação ideológica implí-

cita que se pode ler e compartilhar.

Para que essa proposta seja compreendida e posta em prática, há que se realizar os seguintes objetivos específicos: a) indicar que não há verdadeira neutralidade em projeções algorítmicas, b) caracterizar o conceito que explica a relação da identidade com as mídias sociais e denominá-lo de *identidade algorítmica*, c) demonstrar que o racismo no interior do mundo virtual é reflexo das ações no mundo “real”, d) especificar que é fundamental formar os criadores dos códigos algorítmicos como intérpretes de texto e de mundo capazes de compreender como os variados tipos de vieses interferem em tomadas de decisão e favorecem preconceitos, e) argumentar que a inclusão digital só será uma realidade quando o sujeito perceber-se como íntegro e digno tanto socialmente quanto virtualmente, f) concluir que não haverá mudança social no estado de coisas enquanto não aumentarmos nossa capacidade de compreensão dos dados da realidade, por meio da mudança de nossos próprios paradigmas e modelos de pensamento.

## MATERIAL E MÉTODOS

Esta pesquisa desenvolveu-se fundamentalmente por levantamento bibliográfico, em todas as instâncias. Principalmente, partimos dos estudos de Wolfgang Iser (2013, p. 31), para quem não há uma oposição binária entre o que é considera-

do “real” e o que é “fictício” (no nosso caso, o mundo digital), mas sim uma tríade que se compõe de real, fictício e imaginário. “Como o texto ficcional contém elementos do real sem que se esgote na descrição deste real, seu componente fictício não tem o caráter de uma finalidade em si mesma, mas é, enquanto fingido, a preparação de um imaginário.” Esse modo de pensar nos leva a concluir que a constituição de um imaginário por meio das conclusões fictícias proporcionadas pelas redes sociais, em vez de serem tomadas socialmente como engodo ou mentira, são, na verdade, muitas vezes compreendidas como se fossem exatamente o real.

Outro ponto de que partimos é a ideia das psicólogas sociais e educadoras espanholas Moreno Marimón e Genoveva Sastre (2014, p. 92), de que “toda pessoa constrói *modelos organizadores* dos fenômenos, situações ou acontecimentos que ocorrem em seu mundo, de acordo com sua particular maneira de contemplá-lo [...]”. Isso

leva à construção de um sistema organizado de representações que podemos chegar a conhecer a partir dos *dados que seleciona* de um determinado fenômeno, do *significado* que atribui a tais dados, de como os relaciona ou *organiza* e das *implicações ou consequências que faz derivar de tudo isso. Dados, significado, implicações e organização são*, portanto, os elementos fundamentais detectáveis em todo modelo organizador. (Moreno Marimón; Sas-

tre, 2014, p. 93).

A conclusão é que os referentes das ideias, por mais científicas que elas nos apareçam, não são os “fenômenos” do mundo real, senão as próprias ideias (dados, significado, implicações e organização) que temos sobre estes mesmos fenômenos.

Valemo-nos também de um ponto de vista teórico-metodológico baseado na constituição de identidades linguísticas discursivas. Maria do Rosário Gregolin (2008, p. 87) indica que “A maioria dos autores que tem tratado a questão das identidades enxerga na atualidade elementos em comum: a descontinuidade, a fragmentação, a ruptura e o deslocamento.” Dentre estes estudiosos, ela cita Zygmunt Bauman, com *Identidade* (2005), Stuart Hall, com *Identidade e pós-modernidade* (2002) e Boaventura de Sousa Santos, com “Modernidade, identidade e a cultura de fronteira” (2000). Contudo, é a partir da leitura do artigo “Toxicômanos de identidade – subjetividade em tempo de globalização” (1997), de Suely Rolnik, que ela desenvolve a afirmação mais determinante para o desenvolvimento deste trabalho, a seguinte:

As redes sociais eletrônicas convidam à produção de identidades *fakes*: fazer-se passar por outro, criar codinomes e pseudônimos, inserir falsas imagens de si, criando comunidades como o *Orkut* ou as “guildas” [clãs] dos jogos

eletrônicos. Entretanto, como afirmar que determinada identidade é “falsa” se aceitamos que não há uma única verdadeira? O sujeito da contemporaneidade é um consumidor de identidades cujos *kits* lhes estão disponíveis nas formas midiáticas (Rolnick, 1997). (Gregolin, 2008, pp. 89-90).

Ora, se a relação entre identidades “falsas” e “verdadeiras”, por intermédio do mundo digital e das redes sociais, é mediada por formas midiáticas que se adicionam a cada sujeito, então talvez a própria noção de identidade seja um mito, uma ficção produtora de realidade.

Como último ponto de partida, tomamos o estudo da professora da UCLA e membra da Comissão de Inteligência Artificial da Universidade de Oxford Safiya Umoja Noble, em *Algoritmos de destruição em massa: como o Google fomenta e lucra com o racismo* (2021). Neste trabalho, ela investigou a política que guiava/guia mecanismos de busca e suas representações de “mulheres e meninas negras”. Diz ela:

Documentei como certas buscas por palavras-chave guiam usuários procurando por informações a uma abundância de pornografia, usando a configuração padrão “moderada” do Google Search, e ofereci outros exemplos de como o Vale do Silício defende a si mesmo ao continuar subempregando pessoas que têm experiência nos importantes campos de estudos étnicos e de gênero. O valor dessa exploração está em

demonstrar como gênero e raça são socialmente construídos e mutuamente constituídos através da ciência e da tecnologia. A própria noção de que tecnologias são neutras deve ser diretamente contestada como sendo enganosa. Quer uma pessoa se importe ou não com as más representações de mulheres e meninas não brancas ou considere as representações conceituais problemáticas de adolescentes, professoras, enfermeiras ou médicas, há evidência de que a forma como as plataformas de mídia digital e os algoritmos controlam a narrativa sobre pessoas pode ter consequências atrozadas quando levadas ao extremo. (Noble, 2021, p. 191).

A pesquisa de Noble não só corrobora a noção de que precisamos trabalhar em prol da construção de um modelo de educação de linguagem e leitura que revele os pressupostos por trás dos textos, quanto também demonstra que as narrativas algorítmicas continuam encerradas no *paceto da branquitude* de que fala Cida Bento (2022).

## RESULTADOS E DISCUSSÃO

### O mito da neutralidade

“No princípio era o mito. Depois surge a ficção. Mais tarde ainda aparece a ciência.” (Fiorin, 2016, p. 9). É com essas palavras que o professor e pesquisador José Luiz Fiorin inicia sua obra *As astúcias da enunciação*. A ideia de Fiorin não é a de

demonstrar uma crítica à ciência e/ou à tecnologia e uma descrença na capacidade científica de explicar ideias, instrumentos e processos do mundo. Pelo contrário, o que o autor busca é compreender que “o mito irriga o pensamento científico e a realização artística”, de modo que “ele continua a alimentar todas as formas de apreender a realidade.” (2016, p. 9). O vaticínio embasado de Fiorin, a seguir, é fundamental:

O mito é uma explicação das origens do [ser humano], do mundo, da linguagem; explica o sentido da vida, a morte, a dor, a condição humana. Vive porque responde à angústia do desconhecido, do inexplicável; dá sentido àquilo que não tem sentido. Enquanto a ciência não puder explicar a origem das coisas e o seu sentido, haverá lugar para o pensamento mítico. Será que esse ideal se tornará realidade um dia? Dificilmente. Como se dará conta dos novos anseios, dos novos desejos do ser humano? Precisamos de utopias, que, sendo uma espécie de mito pré-construído, têm a função de organizar e orientar o futuro. (Fiorin, 2016, pp. 9-10).

Quiséssemos discorrer sobre a origem dos mitos, tal qual Mircea Eliade, Lévi-Strauss, Frazer, Vernant, Betty Mindlin e tantos outros o fazem, ou sobre como nossa civilização se vale do mito adâmico para explicar a origem da Linguagem, como o faz Fiorin, e teríamos já aí material suficiente. No entanto, a ideia aqui é perceber como as novas

tecnologias digitais de informação e comunicação apropriam-se de uma espécie de mito sobre o que são *peças negras*, ao mesmo tempo em que criam e disseminam imagens capazes de encapsular estereótipos, principalmente por meio do domínio da Inteligência Artificial Generativa.

Não deve soar estranho que a constituição imagética de pessoas negras no interior do mundo digital seja o alvo de uma pesquisa científica exploratória e qualitativa como esta. Se partirmos do pensamento do filósofo e cientista político camaronês Achille Mbembe (2022, p. 43), que entende que a “tecnologia é hoje uma realidade ao mesmo tempo material e imaterial, psíquica, pessoal e interior. Ela já não pertence apenas ao mundo exterior, membrana que define a fronteira entre um interior (a humanidade) e um exterior (a natureza)”, e se compreendermos com ele e com o psicólogo Tristan Dagron (2019) que é pela tecnologia que muitos de nós experienciamos hoje as “três realidades do mundo vivo”, a saber, a realidade biológica do corpos, “a realidade psíquica dos afetos e a realidade social das trocas, da linguagem e das interações” (Mbembe, 2020, p. 43), então perceberemos a necessidade de se aliar a discussão das teorias humanas – aquelas que não se curvam à predição exata e são chamadas de *soft theory* – à *hard-core theory*, governadas por lei e eventos previsíveis (Iser, 2006).

Pierre Lévy indica de forma muito precisa

em que sentido nos movemos na relação entre esses dois campos de estudo. Para ele, “Os paladinos do *big data* alimentam a ilusão epistemológica de que poderiam prescindir da teoria e que lhes seria possível fazer surgir o conhecimento de uma ‘simples’ análise estatística dos dados” (Lévy, 2013, p. 14). Ora, prescindir da teoria é justamente permanecer na *ilusão* de que a ciência, principalmente aquela construída nas e pelas ciências exatas e engenharias – ou seja, as teorias *hard-core* – é destituída de ideologias e, portanto, *completamente imparcial*. Contra isso, Eduardo Magrani em *A internet das coisas* (2018, p. 23) cita Maike Wile dos Santos e conclui que “*big data* é mais que um emaranhado de dados, pois é essencialmente relacional”, ou seja, o *big data* não tem valor intrínseco, em si, sem que os humanos atribuam sentido a isso. “Para bem entender essas transformações”, segundo o autor, “precisamos compreender que o *big data* somos nós”.

Sem discutir a questão tecnológica, mas sim a constituição dos sujeitos e das comunidades pela linguagem, as análises dos discurso vão pelo mesmo caminho. Por exemplo, desde pelo menos os anos 1960-1970 e a argumentação de Louis Althusser (1976) e Michel Pêcheux (1988), por um lado, e do anos 1990 com os estudos de Norman Fairclough (2001), por outro, sabemos que somos todas e todos ideologicamente comprometidos com determinadas visões de mundo que podem se



alterar, evidentemente, mas que são sempre enviadas. Pêcheux assim indica:

Poderíamos resumir essa tese dizendo: *as palavras, expressões, proposições, etc., mudam de sentido segundo as posições sustentadas por aqueles que as empregam*, o que quer dizer que elas adquirem seu sentido em referência a essas posições, isto é, em referência às formações ideológicas [...] nas quais essas posições se inscrevem. (Pêcheux, 1988, p, 160).

Fairclough, por sua vez, observa o seguinte:

Pode-se considerar uma ordem de discurso como a faceta discursiva do equilíbrio contraditório e instável que constitui uma hegemonia, e a articulação e rearticulação de ordens de discurso são, conseqüentemente, um marco delimitador na luta hegemônica. Além disso, a prática discursiva, a produção, a distribuição e o consumo (como também a interpretação) de textos são uma faceta de luta hegemônica que contribui em graus variados para a reprodução ou a transformação não apenas da ordem de discurso existente (por exemplo, mediante a maneira como os textos e as convenções prévias são articuladas na produção textual), mas também das relações sociais e assimétricas existentes. (Fairclough, 2001, p.123)

Ressalvadas as diferenças entre as teorias e as bases que as sustentam, uma coisa fica patente: a ilusão da neutralidade epistemológica de que

fala Pierre Lévy ao abordar os “paladinos dos *big data*” não se sustenta e é também uma espécie de mito, visto sermos sujeitos discursivos, isto é, seres de linguagem.

O caso de Catherine Taylor, descrito por Cathy O’Neil, auxilia neste entendimento. Taylor teve um emprego na Cruz Vermelha perdido, um pedido de auxílio habitacional negado e ao menos mais dez outras menções negativas a seu nome por ser homônima e ter nascido no mesmo dia de uma mulher acusada de “fabricar e vender metanfetaminas” (O’Neil, 2020, p. 245). Por conta dos processos automatizados e algorítmicos, Catherine não teria chances de provar sua inocência e restabelecer sua identidade, a não ser que um ser humano se incumbisse de uma avaliação que escapasse aos modelos matemáticos predeterminados. Para a sorte de Catherine Taylor, uma funcionária experiente chamada Wanda Taylor (sem parentesco) resolveu seu caso. Com a experiência de sua formação acadêmica em Harvard e no MIT (Massachusetts Institute of Technology), aliada ao tempo dedicado ao mercado financeira em grandes empresas de Wall Street e à sua atuação carreira de professora, Cathy O’Neil (2020, pp. 247-249) conclui:

Nos sobra uma questão: quantas Wanda Taylors existem por aí esclarecendo identidades falsas e demais erros em nossos dados? A resposta: não o suficiente. Humanos na economia

dos dados são pontos fora da curva e algo retrô. Os sistemas são feitos para rodar automaticamente tanto quanto possível. Essa é a maneira mais eficaz; é aí que estão os lucros. Os erros são inevitáveis, como em qualquer programa estatístico, mas a forma mais rápida de reduzi-los é fazer o ajuste fino dos algoritmos rodando as máquinas. Humanos no local apenas obstruem os trabalhos.

[...]

Mas não podemos contar com sistemas automatizados para abordar o problema. Mesmo com seu poder surpreendente, as máquinas não conseguem ainda se ajustar por mais justiça, ao menos não sozinhas. Analisar dados e julgar o que é justo é absolutamente estranho e enormemente complicado a elas. Apenas seres humanos podem impor essa restrição.

Nesse sentido, pensarmos sobre o modo como o mundo digital propõe e propaga determinadas identidades digitais relativas às pessoas pretas e pardas torna-se um imperativo necessário. Achille Mbembe denomina “brutalismo” esta era na qual os elementos digitais não humanos compõem indissociavelmente o ser humano biopsicofisiológico. Para ele, esta é “a era do ser fabricável em um mundo fabricado” (Mbembe, 2022, p. 45). Nós, por outro lado, a essa forma de constituição identitária por meio da qual estão diluídas as fronteiras entre religião, ciência e mitologia e que só encontra as “condições para um reecantamento do mundo [...] em discursos sobre as nanotecnolo-

gias, as tecnologias da informação ou as ciências cognitivas” (2022, p. 45), propomos a denominação de *identidades algorítmicas*. Ao longo deste trabalho, discorreremos sobre essa terminologia e sua aplicabilidade. O próximo ponto, porém, é indicar problemas na também “mítica” inclusão digital.

Todos estamos sujeitos a essa constituição identitária, por certo, se temos boa parte de nossos destinos relacionados às formas como os algoritmos nos alcançam: desde a recomendação de um filme e/ou uma série a ser assistida em um canal de streaming, passando pela indicação de um podcast em uma plataforma de áudio ou do próximo vídeo curto a ser visto nas redes sociais, até maior ou menor probabilidade de conseguir um empréstimo ou um financiamento em um banco (O’Neil, 2020). Apesar do discurso corrente de inclusão digital, há um duplo problema nessa ideia: o primeiro, é o de que as pessoas negras e/ou pertencentes às classes D/E são sistematicamente excluídas socialmente o que gera uma exclusão digital; o segundo, é o de que, quando se consegue ultrapassar a barreira da exclusão social, há uma identidade algorítmica estereotipada que produz novas barreiras à mudança social efetiva. A seguir, abordaremos estes dois problemas.

## Inclusão às avessas

A avaliação do modo como as pessoas negras têm ou não têm acesso à internet e de como isso afeta a possibilidade de alteração das identidades algorítmicas é um ponto a ser discutido. Para embasar nossa discussão, tomemos os seguintes dados, amplamente divulgados na imprensa:

a) Durante o período mais crítico da pandemia de COVID-19, a revista Forbes (2020) divulgou a matéria “Negros e pobres sofrem com exclusão digital durante a pandemia”. Os dados indicavam que 85% da população mais pobre só tinha acesso à internet por meio de telefones celulares. Já quando o recorte era por raça, tínhamos que 65% da população negra só acessava a internet por telefones celulares em relação à 51% da população branca. (Mari, 2020).

b) O jornal Le Monde Diplomatique (2021), na matéria “Dificuldades no acesso à internet: expressões do racismo estrutural”, indicou que 77% das pessoas enquadradas nas classes D/E nunca tinham utilizado um computador de mesa. Pelo recorte racial, apenas 48% dos indígenas e 55% das pessoas pretas já utilizaram computador pelo menos uma vez na vida. Entre as pessoas brancas, o índice é de 63%. (Melo; Terso, 2021).

c) Na matéria “Alunos negros têm menos acesso a recursos tecnológicos na escola” (2024), a Folha de S. Paulo trouxe um estudo do Insper que

comparou o acesso a um pacote de itens de tecnologia (banda larga, laboratórios e um computador a cada cinco alunos) nas escolas em 2019 e 2023. A conclusão foi a de que, em 2019, 51% dos alunos brancos tinham essa estrutura disponível, já em 2023 o número subiu para 57%. Entre os estudantes pardos, esse percentual era de 41% e subiu para 49%, em 2023. Para os alunos pretos, o índice passou de 44% para 50% no mesmo período. Isso demonstra que em 2023 os alunos pretos e pardos ainda não tinham alcançado o patamar de 2019 dos alunos brancos. (Folha, 2024).

d) O jornal O Globo (2024) divulgou na coluna de Ancelmo Gois uma pesquisa feita em parceria do Itaú com o Cebrap sobre o uso de recursos digitais entre pessoas com mais de 50 anos. A pesquisa concluiu que “do ponto de vista da inclusão digital, mulheres e homens negros acessam 14% menos a internet quando comparados com mulheres e homens brancos.” (Gois, 2023, s/p).

e) O Jornal Nacional (2022), reproduzindo os dados de um estudo feito pelo Instituto Locomotiva com a multinacional Pwc, indicou que “No Brasil, a desigualdade digital reproduz a social. Menos de um terço da população está conectada à internet e o acesso pleno é privilégio dos mais ricos. Os mais pobres, em maioria negros, chegam a ficar desconectados metade do mês.” (Jornal Nacional, 2022, s/p).

A observação e a análise mais pormenori-

zada desses dados revela como há uma barreira entre os discursos que pregam a internet e as inteligências artificiais como base para solucionar problemas sociais e as pessoas que mais necessitam que estes problemas sejam resolvidos. Um outro estudo mais incisivo ainda neste sentido é o de Gonzales e Araújo (2021), que demonstraram de que forma a exclusão digital impactou negativamente no recebimento do auxílio emergencial, durante a pandemia, a parcela mais pobre da população brasileira. A conclusão dos pesquisadores da FGV foi que as pessoas das classes D/E em sua maioria não têm celular nem acesso à internet ou, quando têm, não há um plano de dados suficiente para executar funções simples e/ou espaço na memória para baixar novos aplicativos. Tal entrave tecnológico foi responsável, neste sentido, por uma dupla exclusão, já que as pessoas que mais precisavam receber o auxílio emergencial naquele período eram justamente as que mais estavam excluídas digitalmente e, conseqüentemente, as que mais continuaram excluídas socialmente.

Se pessoas negras e pobres, das mais diversas idades, são as que mais sofrem com o acesso à internet, é de se perguntar de que forma podemos realmente esperar que elas interfiram na produção do conhecimento *na* internet. Isto, pois os dados anteriores indicam que tanto jovens quanto pessoas mais velhas, caso sejam negros no Brasil, têm acesso à internet limitado pelo uso de smartpho-

nes. Por mais que saibamos da potência de uso desse tipo de aparelho, não é razoável concluir que por meio dele fica facilitada a possibilidade de desenvolver programas, estudar algoritmos, criar interfaces etc. Tal forma de pensar leva a uma resultante reveladora: a discussão que fazemos acerca das identidades algorítmicas tende a permanecer disponível apenas para o grupo de pessoas identificadas com a branquitude (Bento, 2022) enquanto a inclusão digital permanecer no nível do discurso destituído da prática. Em outras palavras, o que se quer dizer é que, como as pessoas negras têm acesso primitivo à internet e às novas tecnologias digitais de informação e comunicação, a inclusão digital é uma falácia, já que esta é feita de modo desigual e, pior ainda, sem que antes tenha ocorrido a inclusão social.

A fim de não desviarmos do tema deste trabalho, é importante dizer que nosso ponto de vista toma a relação da inclusão com as identidades algorítmicas da seguinte maneira:

1. Falta de inclusão social → Falta de representatividade dos excluídos em locais de decisão. | 2. Desigualdade de inclusão digital → Dificuldade na aquisição e no uso de ferramentas de produção de conhecimento. | 3. Falta de inclusão social + Desigualdade de inclusão digital → Perpetuação de identidades algorítmicas estereotipa-

das + Aumento de barreiras para a inclusão sociodigital. | 4. Aumento de barreiras para a inclusão sociodigital → Falta de inclusão social.

Com esse esquema, derivado da análise de todos os dados que compilamos ao longo deste trabalho e do que Tarcízio Silva (2022) chama de “racismo algorítmico”, compreendemos que as identidades algorítmicas são, na verdade, efeito da reprodução de uma modelagem social. Assim, como pessoas negras são histórica e socialmente mais vulneráveis – haja vista o esforço do Brasil para inclusão na ONU do Objetivo de Desenvolvimento Sustentável (ODS) 18, cujo propósito é a promoção da igualdade étnico-racial (Brasil, 2023) –, sua invisibilidade social reflete na invisibilidade digital. Dessa forma, pessoas negras pertencentes às classes C/D/E tendem a não participar da constituição da internet e dos produtos a ela associados, a não ser como usuárias e em condições precárias, como as reportagens indiciam. Não conseguindo participar efetivamente da criação de modelos de linguagem e de algoritmos que reflitam suas próprias identidades, essas pessoas ficam à mercê do *conhecimento* de outras pessoas, o que gera a perpetuação de identidades algorítmicas estereotipadas, com prejuízo aos negros.

Uma consequência disso é a relação entre racismo e algoritmos na área médica, como discorreremos a seguir. Outra, é a questão relativa às

tecnologias de reconhecimento facial que rotineiramente são usadas contra pessoas negras e dificultam suas vidas quer seja na abertura de uma conta em um banco digital ou sendo acusadas injustamente de algum crime (Santos *et al.*, 2023).<sup>1</sup> Tudo isso leva ao que o pensador jamaicano Stuart Hall chamou de “O Espetáculo do Outro”, em seu livro *Cultura e representação* (2016). Para Hall, a representação de si mesmo e do outro é um ativo de demonstração de poder em nossa sociedade. Por isso, há um agrupamento dado por características que os sujeitos detentores da capacidade e da possibilidade de produzir o mundo digital têm em comum. Essas características, por se repetirem no interior daquele grupo são mais facilmente aceitas como “normais”. De um ponto de vista simbólico, todos os “outros” passam a ser anormais, exóticos ou, no mínimo, diferentes. O diferente, por certo, não é estudado em suas características próprias, mas sempre tomando como base aquilo em que ele difere da norma. Daí o esquema que acabamos de mostrar indicar a ciclicidade de seus fundamentos.

Assim, se por questões sociais poucas pessoas negras podem *construir* o mundo algorítmico, suas tentativas de produzir identidades de si condiscentes com suas próprias experiências não chegam a tornar-se realidade. Se, além disso, as pessoas negras estão alijadas até mesmo de um uso efetivo das ferramentas digitais de mais alto im-

pacto, por estarem restritas ao uso de smartphones e de planos de dados insuficientes, como vimos, então sua “insignificância” estatística permite que os desmandos das identidades algorítmicas racializadas não tragam maiores consequências a quem age dessa forma. Se, de outra maneira, um contingente relevante de pessoas negras das classes A/B aos poucos consegue atingir patamares de alguma influência social e de postos de algum poder de decisão, essa posição não apaga o racismo, e casos como os detectados por Luiz Trindade (2018; 2022) de racismo nas redes sociais continuam demonstrando que há um embate entre *nós, os normais X vocês, os negros* que perpetua a construção das identidades algorítmicas.

Isso quer dizer que a inclusão social falha gera uma inclusão sociodigital falha e que, mesmo quando há uma quebra dessa barreira, a acolhida nas redes é inexistente, fazendo com a pessoa negra não se veja representada naquele contexto ou, caso se veja, perceba-se meramente como um estereótipo de si mesma. A conclusão disso é que mesmo quando pode ocorrer algum tipo de inclusão digital a barreira da identidade algorítmica racializada se ergue impedindo-a. Devemos então compreender em que consiste propriamente esta denominação.

## Identidades algorítmicas

Do ponto de vista do conceito, o algoritmo remete primeiramente ao nome do estudioso matemático persa Muhammad ibn Mûsâ al-Khowârizmi, cujo sobrenome é pronunciado como Alcuarismi (Leavitt, 2009). Foi Alan Turing (1937), no entanto, quem definiu algoritmo como um conjunto não ambíguo e ordenado de passos executáveis que definem um processo finito. Na contemporaneidade, podemos falar que os algoritmos constituem-se como um conjunto de regras e procedimentos lógicos de análise, comparação e tratamento de informações cujo intuito é levar à solução de um problema em um número limitado de etapas. Para tal, existem tanto os algoritmos que podemos chamar de “simples”, como códigos computacionais escritos para resolver problemas específicos, tal qual ativar o despertador em determinado horário, quanto algoritmos “inteligentes”, que são programados para solucionar problemas a partir do conhecimento que adquirem com os dados que recebem.

Assim, com a programação tradicional de algoritmos “simples” alguém sabe a resposta a ser dada para determinado problema e, com a ajuda de um código, essa resposta é inserida em um ambiente, podendo ser alcançada caso sejam seguidas etapas corretas em um determinado processo. Já com os algoritmos “inteligentes”, que também

podemos denominar como “aprendizado de máquina”, “inteligência artificial” e/ou “aprendizado estatístico”, a ideia é fazer a própria máquina aprender o que deve enviar de resposta ao usuário. Tomando quer seja as noções de Pêcheux quer seja as noções de Fairclough, é possível perceber o possível problema embutido nesse “aprendizado”: se os dados que compõem as bases a partir dos quais os sistemas algorítmicos funcionam são socialmente e/ou ideologicamente comprometidos, então não se pode supor que as respostas dadas pelos sistemas de inteligência artificial a quaisquer questões sejam socialmente e ideologicamente neutras.

Dentre as formas como a máquina aprende, podemos citar o *aprendizado supervisionado*, a partir do qual são inseridos dados no sistema marcados como sendo os padrões do que é certo e válido, de um lado, e errado e inválido, por exemplo, de outro. Outro modo é o *aprendizado não supervisionado*, em que não há etiquetas no banco de dados que indiquem os caminhos a serem seguidos e, portanto, a máquina pode agrupar e/ou segmentar dados de acordo com critérios e arranjos que ela percebe no grupo de dados disponível. Uma terceira forma de *machine learning* é a *aprendizagem por reforço*, que tem relação direta com a psicologia comportamental e ocorre quando a máquina aprende pela interação que os usuários têm com ela (Fontana, 2020). Um estudo de Su et

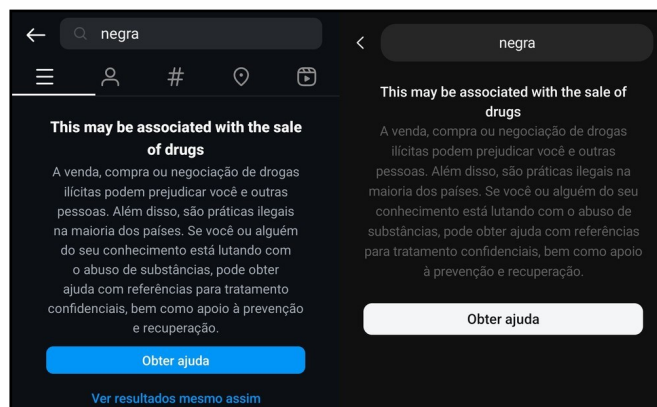
al. (2021), denominado “*Viewing personalized video clips recommended by TikTok activates default mode network and ventral tegmental area*”, por exemplo, descobriu que quando as pessoas recebiam vídeos na plataforma do TikTok que, após a aprendizagem por reforço o algoritmo entendia que eram específicos para elas, havia o aumento da atividade em duas áreas do cérebro: a *default mode network* e a *ventral tegmental area*. Enquanto a primeira diz respeito às reflexões que o sujeito faz de si mesmo, esta última está ligada à liberação de dopamina, o neurotransmissor de busca da recompensa. Ou seja, o que este estudo revelou é que o TikTok conseguiu criar um algoritmo reforçador do que somos enquanto seres sociais e discursivos. Não é difícil perceber, portanto, por qual motivo é tão difícil sair das redes sociais.

Sabendo que a atenção requerida pelos algoritmos das redes sociais e pelos mecanismos de buscas se orienta em torno dessa ideia recompensadora que temos de ver a nós mesmos ali representados, bem como a nossos desejos, o ambiente digital torna-se uma *bomba-relógio* para as populações negras. Ao se preocupar com o “exame da cooptação comercial de identidades, experiências e comunidades negras nas mais poderosas companhias tecnológicas, e em especial no Google”, Noble (2021, p. 20) revelou como uma busca que ela fez em 2009 causou-lhe um impacto profundo. Um amigo um dia lhe disse

“casualmente: ‘Você deveria ver o que acontece quando se procura “meninas negras” no Google’. Eu assim fiz, e fiquei aturdida. Presumi que era um erro absurdo que potencialmente mudaria com o tempo.” (2021, p. 43). Depois, procurando por algo que fosse condizente ao interesse de uma enteada e duas sobrinhas com quem passaria um tempo, Noble foi ao Google. “Minha busca pelas palavras-chave ‘meninas negras’ [*black girls*] apresentou XotaNegraQuente.com [*SugaryBlackPussy.com*] como meu primeiro resultado” (2021, p. 22). A fim de dar o benefício do equívoco ao algoritmo do Google, Noble refez a pesquisa. “Dois anos depois desse incidente, realizei essas pesquisas novamente, apenas para encontrar resultados semelhantes” (2021, p. 22).<sup>2</sup>

A experiência de Noble, ainda que traumática, poderia ser minimizada por algumas pessoas dizendo que os algoritmos, por meio de sua capacidade de autorregulação, já teriam evoluído em relação ao racismo. A realidade, todavia, não permite concluir a mesma coisa. No dia 5 de setembro de 2024, por exemplo, as redes Instagram e Threads, controladas pela empresa Meta, em seus mecanismos de busca relacionavam a palavra “negra” com venda de drogas.

Figura 1— Reprodução de busca pela palavra “negra” no Instagram e no Threads



Fonte: G1. Disponível em: <https://g1.globo.com/tecnologia/noticia/2024/09/05/instagram-e-threads-relacionam-palavra-negra-com-venda-de-drogas.ghtml>

Em reportagem, o G1 (2024, s/p) indicou que a pesquisa pelo termo “negra” levava a uma aviso, no qual as “plataformas [ofereciam] ajuda sobre como denunciar conteúdo. Por outro lado, as redes sociais não mostraram alertas em buscas por 'branca' ou por drogas como 'maconha' e 'cocaína'.” Os avisos das redes alertavam para os prejuízos pessoas relacionados à “venda, compra ou negociação de drogas ilícitas”, bem como continham atalhos para os usuários pedirem ajuda para si mesmos, denunciarem algo ou verem os resultados mesmo assim. Neste caso, a Meta afirmou “que evita mostrar nas pesquisas conteúdos que violem suas diretrizes de recomendações, mas não explicou em seu site por que a palavra ‘negra’ entra nesse filtro.”

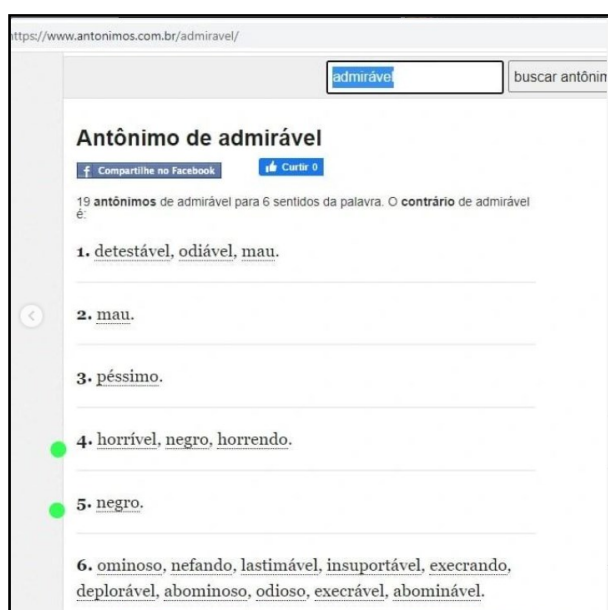
O questionamento seguinte é, necessariamente, o de pensar sobre o que leva a máquina a



responder dessa maneira à entrada da expressão “negra”. Como sabemos, os sistemas são abastecidos por dados minerados na própria rede e também por dados incluídos por pessoas. Um simples exemplo é capaz de demonstrar porque essas associações são “naturalizadas” e só se alteram quando há um clamor nessa direção. Em janeiro de 2023, ao buscarmos em um dicionário online bastante usado um antônimo para a palavra “admirável”, encontramos como enorme surpresa, por duas vezes, a palavra “negro”:

Figura 2 - Captura de tela do site

www.antonimos.com.br



Fonte: Autor (2024)

Após cobrança feita por intermédio das redes sociais, a equipe da empresa Dicio, responsável pelo dicionário de antônimos, respondeu-nos dizendo: “Enquanto dicionários, registramos os

significados e buscamos sempre atualizar e adaptar nossas definições de acordo com a evolução das palavras.”<sup>3</sup> O que todos esses dados podem nos levar a concluir é que há a possibilidade de que os algoritmos que relacionam “meninas negras” a “pornografia” e a expressão “negra” a drogas ilícitas não fazem mais do que repetir padrões que de tal forma se estabeleceram em nossas sociedades a ponto de indicarem como uma palavra cujo significado é contrário a “admirável” a palavra “negro”, em duas entradas distintas.

Precisamos nos atentar a como a resposta da Dicio tem uma relação direta com a formação de identidades algorítmicas. Ora, em primeiro lugar, se a função dos dicionários é registrar as palavras em contextos reais de uso, isso significa que durante um momento significativo de nossa organização sociocultural a palavra “negro” era usada em contextos nos quais se opunha ao significado atribuído à palavra “admirável”. Muito provavelmente, havia textos em que isso estava registrado, além das interações orais que ficaram guardadas na memória popular. Isso nos leva ao segundo ponto, pois, como explica George F. Luger (2013, p. 476), dentre as várias razões para a comunidade de Inteligência Artificial ter encontrado tantas dificuldades na compreensão e no processamento de raciocínio automatizado e linguagem natural estão

a grande quantidade de conhecimento, de habilidades e de experiência necessária para dar

suporte ao uso da linguagem. A compreensão bem-sucedida da linguagem requer um entendimento do mundo natural, da psicologia humana e de convenções sociais. Ela se vale de habilidades tão variadas quanto o raciocínio lógico e a interpretação de metáforas. Por causa da complexidade e da ambiguidade da linguagem humana, a compreensão de linguagem natural tem motivado, em grande parte, a pesquisa em representação de conhecimento. [...] As articulações de linguagem não são criações aleatórias de sons ou palavras, mas ocorrem em padrões.

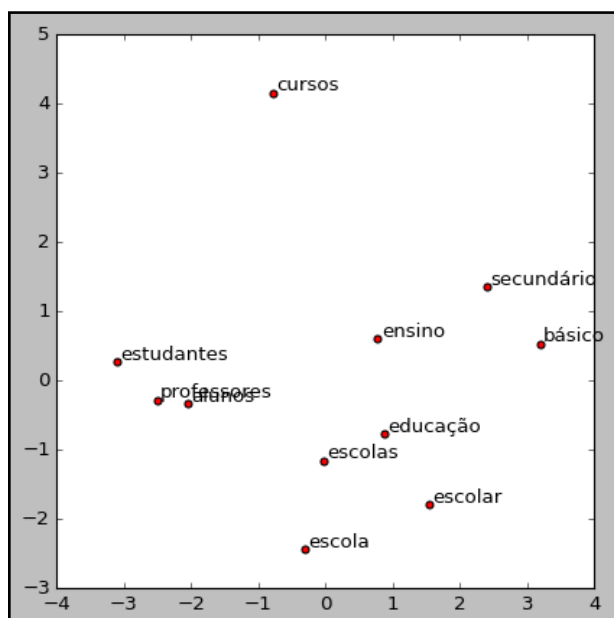
Esses “padrões” são identificados, dentre outras formas, por vetores de palavras (*word vectors* ou *word embeddings*). Dos exemplos mais usados constam o Word2Vec, FastText e o GloVe (além do Elmo, do Bert e do GPT), que são uma classe de algoritmos cujo propósito é o de encontrar uma representação vetorial para palavras que, quando ocorrem em contextos parecidos, tenham vetores também parecidos (Pennington; Socher; Manning, 2014). Algoritmos desse tipo usam redes neurais para fazer previsões sobre uma palavra levando em consideração seu contexto ou sobre o contexto tendo como base uma palavra (Levy; Goldberg; Dagan, 2015). Como essa *semântica distribucional* é disposta também com base na filosofia analítica de Wittgenstein, é possível compreender o procedimento por trás dessa organização, já que para o filósofo “o significado [*Bedeutung*] de uma

palavra é o seu uso na linguagem” (Wittgenstein, 2002, §43). Sendo assim, os vetores das palavras são constituídos quando a máquina “lê” uma série de textos e identifica em que contextos aquela palavra aparece, que tipo de palavra normalmente a antecede e, depois, a sucede etc. Com dados como esses, por exemplo, é que é possível que tenhamos mecanismos de autocompletar em processadores de textos e buscadores, pois há certa previsibilidade sobre qual palavra e/ou segmento de palavra deve completar a sentença dado o contexto específico em que ela se inscreve.

Em consequência disso, o uso da representação vetorial semântica é um padrão de representação muito usual em Processamento de Linguagem Natural (PLN). O vetor semântico representa um ponto dentro de um *espaço vetorial multidimensional* que, por sua vez, é caracterizado na álgebra linear pelo número de direções independentes no espaço. Ou seja, cada palavra do “mundo real”, quando compreendida como um vetor semântico, tem seu significado construído a partir da distribuição densa ou esparsa das palavras “vizinhas”. Por isso é que ganha destaque a multidimensionalidade, já que os vetores podem retratar vários aspectos do significado das palavras, tais como a similaridade, a orientação e associação, entre outros. Vê-se similaridade entre “comércio” e “negócio”, por exemplo. A orientação de sentimento, ou polaridade, pode ser percebida, por

exemplo, em “admirável”, que conota uma avaliação positiva, e “detestável”, que conota uma avaliação negativa. Ainda, a associação entre palavras pode ser comprovada por “vôlei” e “bola”, que são claramente relacionados, uma vez que vôlei se joga com uma bola, etc. Veja-se o exemplo da representação do espaço vetorial semântico da palavra “ensino”, gerada com o modelo GloVe, disponibilizado pelo Núcleo Interinstitucional de Linguística Computacional (NILC) da USP:

Figura 3 - Espaço vetorial semântico da palavra “ensino”



Fonte: NILC. <http://nilc.icmc.usp.br/nilc/index.php/repositorio-de-word-embeddings-do-nilc>. In: SENO *et al.*, 2023. <https://brasileiraspln.com/livro-pln/1a-edicao/parte5/cap10/cap10.html>

Por essa lógica que estamos investigando, palavras como “escola”, “educação”, “estudantes”,

“curso”, “professores” e “alunos”, em maior ou menor grau, ajudam a estabelecer o significado da palavra “ensino”. Ou seja, o contexto e o campo semântico determinam a conotação de determinada palavra. Justamente por isso, se voltamos à entrada do dicionário e percebemos a estrutura de significados que ele recolheu, então conseguimos compreender que há uma chance de que a inteligência artificial em questão tenha aprendido com os textos que nós produzimos, enquanto sociedade, e feito a associação mais “lógica”, em que “meninas negras” se relaciona diretamente com “XotaNegraQuente.com”, e “negra” se relaciona diretamente com “maconha e cocaína”. Isso, pois pornografia e drogas ilícitas, para o padrão médio da sociedade, é visto como algo negativo. Do mesmo modo, “negro” associa-se/ associou-se a “detestável, odiável, mau, péssimo, horrível, horrendo, ominoso, nefando, lastimável, insuportável, execrando, deplorável, abominoso, odioso, execrável, abominável”, já que pela semântica distribucional (Harris, 1954; Firth, 1957) o significado de uma palavra é dado pelo contexto de seu uso, isto é, pelo significado das palavras que mais frequentemente ocorrem por perto dessa primeira. Portanto, como nessa acepção “negro” é antônimo de “admirável”, escapa do campo semântico relativo à “admirável”, opondo-se completamente.

Se sabemos, com os já citados Pêcheux, Gregolin, Fairclough, Mbembe, Moreno Marimón

e Sastre, que as identidades se constituem linguisticamente de acordo com os significados atribuídos a cada sujeito em relação aos outros e a si mesmo, então as identidades algorítmicas de que falamos se constituem, por exemplo, a partir de identificações que ora incluem e ora excluem esses mesmo sujeitos de determinados campos semânticos. No caso, alija-se a “pessoa negra” do campo semântico de “admirável”. De um ponto de vista histórico, essas relações lógicas auxiliam inclusive a compreender por que motivo na década de 1960 ganhou muita atenção nos Estados Unidos o movimento *Black is beautiful*, “O negro é lindo”, em tradução literal. A “surpresa” estava justamente na aproximação semântica de duas palavras que *costumeiramente* não faziam parte do mesmo espaço vetorial nos textos publicados nas grandes mídias, nos livros e nas conversas banais da sociedade letrada.<sup>4</sup>

Essas escolhas no campo das relações semânticas organizadas pelos códigos algorítmicos, no entanto, não são apenas um “bug” no sistema, isto é, uma falha cuja procedência não é identificada, e o exemplo de Anna Everett é o suficiente para demonstrá-lo. No início dos anos 2000, enquanto ligava seu computador, Everett percebeu que os *prompt* (conjunto de palavras que fornece as informações necessárias para que o modelo de linguagem saiba o que deve gerar como resultado) de comando do DOS mostrados na tela indicavam a

hierarquia virtual do sistema (Santos, 2022, p. 454; Daniels; 2018, p. 62). O problema é que as palavras que os responsáveis pela programação usaram eram “*Pri. Master Disk, Pri. Slave Disk, Sec. Master, Sec. Slave*”, que em tradução livre significam “Primeiro Disco Mestre, Primeiro Disco Escravo, Segundo Disco Mestre, Segundo Disco Escravo”. O grande questionamento que se segue é de Everett tentando compreender por qual motivo, para demonstrar que um sistema era *dependente* do outro para o correto funcionamento, os pesquisadores *escolheram* estas palavras, “senhor” e “escravo”, como as representativas? (Everett, 2002). O ponto crítico, como é possível perceber, não está “apenas” no resultado do que os programadores fazem, isto é, no código, mas também nas *palavras* que eles usam para organizar o mundo, descrevê-lo, imaginá-lo e reimaginá-lo, pois serão elas que se transformarão em *prompts* que, por sua vez, direcionarão o mundo digital disponível.

Tendo em vista que nosso argumento se arvora nessas relações lógicas organizadas pela semântica distribucional e pela análise do discurso, faz sentido termos explicitado cada passo como o fizemos até então. Há certamente algo por ser dito sobre como uma pessoa negra que se move pelas redes sociais e pelo mundo digital e recebe definições como essa na construção de sua própria identidade e, também, sobre como isso interfere, ainda

que de modo implícito, na própria inclusão digital. Abordaremos isso em seguida. No momento, convém indicar como a alteridade digital também contribui para a construção de uma identidade algorítmica negra em posição socialmente rebaixada.

Em estudo sobre os dados do Facebook no Brasil, por exemplo, Luiz Trindade (2018, p. 195) identificou que, entre 2012 e 2016, 81% de todas as vítimas de racismo na plataforma eram mulheres que tinham alguma relevância social – eram modelos, cantoras, estudantes universitárias, esportistas etc. – e idades entre 20 e 35 anos. Os de- tratores eram em sua maioria homens e jovens e os motivos para a exposição pública do racismo eram variados, com 64,3% sendo uma reação a algo postado nas redes, como: “uma foto na piscina”, “um texto sobre ter sido aceita em uma facul- dade de medicina no exterior”, “a rejeição de ami- zade na rede com um homem estranho”, “uma foto grávida”, “uma foto com o cabelo natural”, etc. Ou seja, a existência dessas mulheres negras de modo “livre” nas redes sociais já era o bastante para que elas fossem objeto de manifestações de ódio. Mais uma vez, as palavras que se associam às imagens constroem um imaginário social no qual é recorrente e “naturalizado” que, após a publicação de uma postagem com uma foto de um corpo ne- gro, por exemplo, haja uma série de comentários negativos, ofensivos e de baixo calão como respos- ta. Não se pode duvidar que a máquina também

“aprenda” esse padrão e o identifique como uma reação “normal” de nossa sociedade, visto que a repetição desse modelo é constante.

Para que não se pense que a questão que levantamos em relação à identidade algorítmica é “meramente estética” e, por isso, de menor rele- vância – o que por si só já merecia uma análise mais apurada –, tomemos um último exemplo de como o corpo social é primariamente responsável pela construção de um campo semântico cuja for- matação será base para os vetores de palavras e de como isso impacta de fato a vida das pessoas. Em 2011, o Instituto de Pesquisa Econômica Apli- cada (IPEA) publicou o estudo “Desigualdade de transplantes de órgãos no Brasil: análise do perfil dos receptores por sexo e raça ou cor”. É impor- tante indicar que, de modo análogo ao que prati- camos nesta pesquisa específica, também o IPEA teve de recorrer a muitas fontes estrangeiras em sua revisão de literatura. “Apesar da relevância, o tema ainda carece de bibliografia no Brasil. Como encontramos um número maior de referências na literatura norte-americana, nossa revisão biblio- gráfica contém muitas citações do assunto refe- rente aos Estados Unidos.” (Marinho *et al.*, 2011, p. 9). O estudo concluiu que “a grande maioria dos receptores são homens brancos”, mas isso não é devido ao fato de esse grupo ser mais suscetível à necessidade de transplantes, já que “tanto as mu- lheres, quanto os negros e os pardos, apresentam

necessidades bem semelhantes às dos homens brancos”. As razões para isso, de acordo com o IPEA, são muitas e complexas, mas “perfeitamente identificáveis”. Para os/as pesquisadores/as, os elementos que levam a essa desigualdade nos transplantes de órgãos no Brasil e no exterior são: “Preconceitos, medos, desinformação, biologia humana, subfinanciamento da saúde, racismo e vieses desfavoráveis às mulheres e às minorias e diversos tipos de disparidades [...]” (Marinho *et al.*, 2011, p. 37).

Sabendo que o trabalho do IPEA foi publicado a quase quinze anos, era de se esperar que a sociedade tivesse encontrado formas de otimizar esse importante trabalho e retirado os vieses racistas e machistas desse processo. No entanto, com o trabalho “*Racial disparities in kidney transplant waitlist appearance in Chicago: Is it race or place?*”, “Disparidades raciais no aparecimento em listas de espera para transplante renal em Chicago: uma questão de raça ou de lugar?”, em tradução literal, Robert B. Peng *et al.* (2018) identificaram uma correlação algorítmica entre os itens “raça” e “endereço” e a maior ou menor prioridade para entrada na lista de espera para um transplante. Segundo o estudo, se a pessoa for negra e periférica, então menor é a chance de receber maior prioridade no programa digital que classifica e ordena os possíveis receptores de órgãos.

O estudo “*Examining the Potential Impact of*

*Race Multiplier Utilization in Estimated Glomerular Filtration Rate Calculation on African-American Care Outcomes*”, em tradução literal “Examinando o impacto potencial da utilização do multiplicador de raça no cálculo da taxa de filtração glomerular estimada sobre os resultados do atendimento a afro-americanos”, de Salman Ahmed *et al.* (2021), por sua vez, foi ainda mais incisivo. Após examinar o histórico de mais de 57 mil pacientes com doenças crônicas nos rins, em Boston, EUA, o estudo concluiu que o cálculo padrão CKD-EPI usado para converter os exames de sangue em um conceito capaz de avaliar a condição dos rins continha um aumento de 15,9% no número final quando o/a paciente era negro/a. De acordo com o estudo, no entanto, a questão racial não interfere de modo algum no funcionamento dos rins. O mais sério problema dessa fórmula é que para alcançar um lugar mais à frente na fila de transplantes é necessário que a pontuação conquistada após os cálculos seja baixa. Dessa maneira, o estudo em questão demonstrou que caso o algoritmo utilizado não considerasse a variável raça, então mais de 700 pacientes negros, ou seja, aproximadamente um terço deles, teria sido classificado como portador de doença grave e, por isso, aumentado sua chance de sobrevivência (Ahmed *et al.*, 2021). Neste caso, por conta de determinada identidade do sujeitos, o algoritmo decidiu quem tinha mais chance de viver e quem tinha menos.

Uma variante desse mesmo tema foi encontrada no artigo *“Hidden in Plain Sight – Reconsidering the Use of Race Correction in Clinical Algorithms”*, “Escondido à vista de todos – Reconsiderando o uso da correção de raça em algoritmos clínicos”, em tradução literal, de Darshali A. Vyas *et al.* (2020). Os autores ilustram seus argumentos indicando que em áreas como cardiologia, nefrologia, obstetrícia, oncologia, endocrinologia, urologia e pneumologia, os algoritmos que auxiliam os médicos na tomada de decisão sobre tratamentos e cirurgias se organizam também em torno de preconceitos raciais que são alheios às questões biológicas:

Uma inserção sutil da raça na medicina envolve algoritmos de diagnóstico e diretrizes práticas que ajustam ou “corrigem” seus resultados com base na raça ou etnia do paciente. Os médicos usam esses algoritmos para individualizar a avaliação de risco e orientar as decisões clínicas. Ao incorporar a raça aos dados básicos e às decisões de assistência médica, esses algoritmos propagam a medicina baseada em raça. Muitos desses algoritmos ajustados por raça orientam as decisões de forma que podem direcionar mais atenção ou recursos para pacientes brancos do que para membros de minorias raciais e étnicas. (Vyas *et al.*, 2020, p. 874).

Quer sejam as questões médicas, estéticas, sociais, culturais, o que todos esses dados revelam é que os algoritmos, ao produzirem e reproduzi-

rem imagens e identidades preconceituosas de pessoas negras, na verdade parecem estar “apenas” dando continuidade a um modo de ser e de pensar que se estabeleceu na sociedade. Esse modo de ser identifica pessoas racializadas e coloca, nas primeiras páginas, o que a sociedade em organizada em torno de um viés de branquitude espera que essas pessoas sejam. Em suma, a problemática é tanto de organização cultural quanto de arranjo e relevância de determinados dados. Safiya Noble (2021, p. 240), refletindo a necessidade de uma organização interdisciplinar da ciência de dados e da biblioteconomia, questiona:

Se o processo que guia o desenvolvimento de sistemas tecnológicos é a prioridade que se dá à aquisição eficiente e **rápida** de informações em registros, de bancos de dados a mecanismos de busca pela internet, então quais são os traços digitais singulares que definem e distinguem a informação sobre pessoas racializadas e mulheres nos Estados Unidos?

Um pouco mais à frente, Noble (2021, p. 247) continua: “Temos que nos perguntar o que quer dizer, em termos práticos, buscar por conceitos sobre gênero, raça e etnicidade apenas para descobrir informação faltante ou mal representa-

da, seja no banco de dados da biblioteca ou na rede aberta.” E se nos perguntamos, junto com a pesquisadora, percebemos que a situação brasileira não é muito distinta da norte-americana, mesmo porque as ações tomadas pelos algoritmos não só estão no mundo e o organizam, mas verdadeiramente inventam um mundo (Silva, 2021; Silva, 2022) que, em larga medida, emula as mesmas relações sociorraciais em que os sujeitos negros são rotineiramente subalternizados.

Sabendo-se que as identidades que temos como sujeitos dependem das relações sociais e linguísticas que mantemos com o entorno das comunidades “virtuais” e/ou “reais” de que participamos, logo se vê que os processos de identificação com o outro, com o mundo e consigo mesmo, para as pessoas negras, é enviesado de tal forma que sua identidade algorítmica é fatalmente diminuída. Quando se tem a compreensão de que as fronteiras entre o digital e o não digital se liquefazem, a identidade algorítmica diminuída significa a pessoa inteira diminuída.

Sabendo-se também que “Não existe mais a tecnologia de um lado e, do outro, aquilo que a filosofia ocidental [de Martin Heidegger] costumava chamar de a ‘verdade do ser’.” (Mbembe, 2022, p. 44), então a tecnologia dos algoritmos não só constrói imagens que se identificam ou desidentificam com os sujeitos, sem maiores consequências. Pelo contrário, as identidades algorítmicas consti-

tuem os sujeitos, quer seja como avatares virtuais de si mesmos quer seja como indivíduos de carne e osso, de tal maneira que sua própria existência parece ter como destino uma automatizada escolha algorítmica supostamente neutra, como já discutimos anteriormente. Por isso é que “Parte do desafio de compreender a opressão algorítmica é perceber que as formulações matemáticas que guiam as decisões automatizadas são feitas por humanos” (Noble, 2021, p. 18). Por isso é que, antecipando já elementos que pensamos serem necessários para começar o enfrentamento desses problemas, mais do que “atacarmos” os algoritmos e as “inteligências artificiais”, o que precisamos é formar mais e melhores pessoas de acordo com esses entendimentos, isto é, compreendendo a tendência de espriamento dos resultados algorítmicos como um fator decisivo para a existência factual de determinadas pessoas.

## CONCLUSÃO

Conforme desenvolvemos ao longo deste estudo, há sérios entraves à inclusão digital de pessoas racializadas, em especial pessoas negras, e isso ou produz ou mantém distorções sobre as imagens sociais dessas pessoas em seus grupos. A esse processo de estereotipação imagética das identidades demos o nome de *identidades algorítmicas*. A fim de combater esse desvio que vai des-



de o âmbito das engenharias e da informática, passando pelas ciências sociais e econômicas e chegando à comunicação e às linguagens, vários são os processos que precisam ser postos em prática e muitos já o estão sendo, principalmente com os programas de transferência de renda proporcionados pelo Governo Federal. Este tipo de programa é fundamental, já que, como vimos, a relação entre exclusão digital e exclusão social é simbiótica.

Outro movimento importante que precisa ser continuado e ainda mais incentivado é o de cotas raciais para ingresso em universidades e institutos federais, pois só tendo capacidade técnica adequada é que os sujeitos negros poderão propor seus próprios algoritmos e alterar o estado de coisas racista que hoje se encontra espalhado em vários ramos sociais. Este é um trabalho lento, de resultado demorado, mas que deve ser fomentado com vistas a uma futuro mais igualitário.

Há, no entanto, um forte ponto negativo comum a essas duas propostas de resolução, a saber, a culpabilidade recaída sobre a vítima. É necessário explicar melhor: nestes dois modos de solucionar o problema, são as pessoas negras que primeiro receberão auxílio e/ou instrução e, depois disso, lutarão para modificar as estruturas racistas da sociedade. Mais uma vez, há um duplo encargo. Nossa proposição, portanto, vem no intuito de ser o mais horizontalizada possível, pois não tem a intenção de recair apenas sobre as pes-

soas negras e/ou pobres.

Diante de todos os dados que levantamos, o que mais ficou explícito é que há um viés nos algoritmos que contraria o senso comum de neutralidade das inteligências artificiais. Sendo assim, o que propomos como solução é justamente a discussão, a explicação e a apresentação dos vieses cognitivos que estão por trás da produção não só dos algoritmos, mas de toda e qualquer produção comunicativa que os seres humanos realizam. De modo específico, cremos na necessidade do desenvolvimento de um *material didático* a ser implementado no Ensino Médio e nos primeiros anos dos diversos cursos universitários.

De modo análogo ao trabalho que já vem sendo realizado pelo Instituto Palavra Aberta, por meio do Programa Educamídia, que recebe apoio do Google.org, criado para capacitar professores e organizações de ensino, e engajar a sociedade no processo de educação midiática dos jovens, nossa proposta é que um material didático voltado para os últimos anos do Ensino Médio e o primeiro ano universitário auxilie nas disciplinas de Comunicação e Expressão e de Língua Portuguesa. Sabendo que já temos a Lei nº 11.645/08, que inclui no currículo oficial da rede de ensino a obrigatoriedade da temática “História e Cultura Afro-Brasileira e Indígena”, a Lei nº 14.926/24, que inclui a educação ambiental sobre mudanças climáticas e biodiversidade no ensino superior e a Lei nº 14.986/24,

que tornou obrigatório o ensino sobre as contribuições das mulheres na educação básica do Brasil, por exemplo, cremos na necessidade de colocar em pauta e discussão uma lei que obrigue o ensino dos fundamentos dos algoritmos e das bases de interpretação de textos e vieses cognitivos que eles contêm.

A ideia deste material é realmente compreender o estudo de língua portuguesa não como mero armazenamento de funções gramaticais e/ou literárias que têm pouca relevância para a efetivação leitura dos textos e compreensão dos sentidos do mundo. O material didático em questão deve abordar os seguintes aspectos:

*Parte 1 – As bases da leitura.* Nesta parte, serão estudados os elementos básicos que devem ser trabalhados para a leitura de todo e qualquer texto. A ideia é fornecer ferramentas para o aumento da capacidade leitora dos sujeitos o que, por sua vez, facilitará o entendimento das mensagens implícitas nas formas de comunicação on-line e off-line. Os tópicos dessa primeira parte são os seguintes: 1. O que é texto e quais são as formas básicas de leitura?; 2. Quais são os elementos que conferem textualidade?; 3. O que são signos linguísticos? Qual a relação entre signos, símbolos, índices e ícones de linguagem?; 4. O que são elementos explícitos e implícitos em um texto? Qual a diferença entre pressupostos e subentendidos?; 5. Como a memória social e linguística interfere na

capacidade de ler elementos implícitos? Como ampliar o espectro de referências e melhorar a interpretação de textos?; 6. Como a constituição de subjetividade e alteridade interferem na interpretação de textos e na atribuição de sentidos ao mundo?; 7. Como descobrir elementos que revelem propósitos linguísticos e como a interlocução textual produz imagens e atribui sentidos a um texto e ao mundo?; 8. Qual a relação entre os gêneros textuais e as funções da comunicação?

*Parte 2 – Construção argumentativa e inclusão sociodigital.* Levando em consideração os aspectos já citados e também a contribuição essencial da professora Roxane Rojo no livro *Letramentos múltiplos, escola e inclusão social* (2009), nesta parte o foco de estudo são as várias maneiras de construir argumentos, a avaliação crítica de ações e a relação entre tomada de decisões, inclusão social e inclusão digital. Os tópicos dessa segunda parte são os seguintes: 1. O assunto, o fato e a opinião: como distingui-los em diversos contextos?; 2. Qual a relação entre opinião, verdade, moral e ética?; 3. Como as *Fake news* e o discurso de ódio orientam a cidadania e os modelos argumentativos?; 4. O que é letramento digital e como ele pode dar ferramentas para combater a manipulação comunicativa?; 5. Como definir tipos de problemas e formas de solução no mundo on-line e no mundo off-line?; 6. Como compreender números que traduzem ideias e avaliar enganos produzidos

pela leitura enviesada de dados, gráficos, tabelas e estatísticas?; 7. Como dados gráficos, tabelas e estatísticas ficcionalizam e produzem identidades de pessoas racializadas e/ou de comunidades minorizadas?; 8. Como as imagens e seu tratamento digital ficcionalizam e produzem identidades de pessoas racializadas e/ou de comunidades minorizadas?

*Parte 3 – Os vieses cognitivos e os algoritmos.* Nesta parte, o foco está na capacitação para a percepção da presença e da atuação de sistemas algorítmicos nos aplicativos, plataformas e serviços digitais para o entendimento de como esses sistemas funcionam, conseguindo observar seu alcance, suas possibilidades e sendo capazes de moldá-los. Para tal, valemo-nos do processo de tradução do *Algorithmic Bias Playbook (Manual sobre vieses algorítmico, em tradução livre)* da University of Chicago Booth, sob responsabilidade de Obermeyer *et al.* (2021) e também dos estudos de Kahneman e Miller (1986) e Kahneman e Tversky (2000). Os tópicos dessa terceira parte são os seguintes: 1. O que são vieses cognitivos? Como identificá-los, mensurá-los e detectar suas causas em contextos algorítmicos?; 3. Qual a relação entre o processamento de informação e a tomada de decisão por humanos e por máquinas? 4. Como a questão sociocultural implica avaliações e reavaliações de vieses? 5. Como perceber o viés de ancoragem e o viés de afinidade? 6. Como perceber o viés de percepção e o viés de confirmação? 7. Como perceber

o viés de efeito halo ou efeito auréola e o viés de grupo? 8. Como perceber o viés de grupo e combater todas as formas de vieses já descritas e relatadas?

Por certo, a implementação de um material didático que aborde os temas que apresentamos não resolve a questão por completo. Mesmo porque a constante mutação do gênero dos problemas impede que se almeje uma resolução única. Ainda assim, trabalhando de forma multidisciplinar e multifatorial, como este material didático pretende, teremos mais chance de atingir pessoas dos mais variados nichos, credos, gêneros, etnias e classes sociais, fazendo uma transformação mais horizontal das possibilidades de construções de identidades algorítmicas. Isto, pois a revolução digital que se espera, e da qual as pessoas negras devem tomar parte como agentes, não está relacionada às aparências das coisas por intermédio das IAs generativas e de seus algoritmos, mas sim do modo como nossas percepções do mundo estão sendo organizadas por essas IAs. Neste sentido, nossos pensamentos e as relações que mantemos uns com os outros e com as coisas do mundo real-fictício-imaginário que nos cerca é que podem ou não sucumbir a formas pré-determinadas e/ou estereotipadas de caracterizar todos aqueles/as homens e mulheres negros/as que, ainda hoje, são vistos como o *outro exótico*.

Enfim, se nossa proposta de solução diz

respeito ao aprimoramento das técnicas de leitura do mundo e de compreensão dos vieses algorítmicos, é porque de fato as técnicas não são apenas ferramentas, mas, isto sim, estruturas de percepção. Quem é capaz de dominar o máximo de técnicas possível é também capaz de perceber de que modo elas condicionam a forma como o mundo nos aparece e a forma como os fenômenos nos são dados a apreciar.

## REFERÊNCIAS

- AHMED, S., NUTT, C.T., ENEANYA, N.D. *et al.* Examining the Potential Impact of Race Multiplier Utilization in Estimated Glomerular Filtration Rate Calculation on African-American Care Outcomes. **Journal of General Internal Medicine**. v. 36, pp. 464-471 (2021). <https://doi.org/10.1007/s11606-020-06280-5>
- ALTHUSSER, L. **Filosofia e Filosofia Espontânea dos Cientistas**. Lisboa: Editorial Presença, 1976.
- BBC. **Como o Facebook mudou a internet, o comércio e até a política**. Reportagem de 11 julho 2021. <https://www.bbc.com/portuguese/internacional-55966587>
- BENTO, C. **O pacto da branquitude**. São Paulo: Companhia das Letras, 2022.
- BRASIL. **Resolução nº 02/CNODS, de 20 de dezembro de 2023**. Institui a Câmara Temática para o Décimo Oitavo Objetivo de Desenvolvimento Sustentável sobre Igualdade Étnico-Racial na Agenda 2030. Disponível em: <https://www.gov.br/secretariageral/pt-br/cnods/resolucoes/Resolucao2ODS18IgualdadeEtnicoRacial.pdf>
- BOLZAN, L. M.; LÖBLER, M. L. Socialização e afetividade no processo de inclusão digital: um estudo etnográfico. **Organizações & Sociedade**. 2016, v. 23, n. 76, pp. 130-149. Disponível em: <https://doi.org/10.1590/1984-9230767>. Acessado 3 Outubro 2024
- DANIELS, J. The algorithmic rise of the “alt-right”. **Contexts**, Vol. 17, Issue 1, pp. 60-65. American Sociological Association. <https://doi.org/10.1177/153650421876654>
- DAGRON, T. **Pensée et cliniques de l’identité**. Paris: Vrin, 2019.
- DW. **O partido que faz alemães temerem a volta do nazismo**. Reportagem de 9 de fevereiro de 2024. <https://www.dw.com/pt-br/o-partido-que-faz-alem%C3%A3es-temerem-a-volta-do-nazismo/video-68219970>
- EVERETT, A. **The Revolution Will Be Digitized: Af-**

rocentricity and the Digital Public Sphere. In: **Social Text**. v. 20, n. 2, pp. 125-146, 2002.

FAIRCLOUGH, N. **Discurso e mudança social**. Brasília: Editora Universidade de Brasília, 2001.

FIORIN, J. L. **As astúcias da enunciação**: as categorias de pessoa, espaço e tempo. 3. ed. São Paulo: Editora Contexto, 2016.

FIRTH, J. R. The technique of semantics. **Transactions of the philological society**, v. 34, n. 1, p. 36–73, 1957.

FOLHA DE S. PAULO. **Alunos negros têm menos acesso a recursos tecnológicos nas escolas**. 18 junho 2024. <https://www1.folha.uol.com.br/educacao/2024/06/alunos-negros-tem-menos-acesso-a-recursos-tecnologicos-nas-escolas.shtml>

FONTANA, E. **Introdução aos algoritmos de aprendizagem supervisionada**. Curitiba: Universidade Federal do Paraná (UFPR), 2020. Disponível em: [https://fontana.paginas.ufsc.br/files/2018/03/apostila\\_ML\\_pt2.pdf](https://fontana.paginas.ufsc.br/files/2018/03/apostila_ML_pt2.pdf).

GOIS, A. Idosos negros acessam 14% menos a internet quando comparados a mulheres e homens brancos. **O GLOBO**. 23 maio 2023. <https://oglobo.globo.com/blogs/ancelmo-gois/>

[post/2023/05/idosos-negros-acessam-14percent-menos-a-internet-quando-comparados-a-mulheres-e-homens-brancos.ghtml](https://post/2023/05/idosos-negros-acessam-14percent-menos-a-internet-quando-comparados-a-mulheres-e-homens-brancos.ghtml)

GONZALES, L.; ARAUJO, M. Efeitos da exclusão digital no acesso ao auxílio emergencial. In: **Centro de Estudos de Microfinanças e Inclusão Financeira da FGV**, 27 maio, 2021. Disponível em: [https://www.impacto.blog.br/site/wp-content/uploads/2021/05/Auxilio\\_InclusaoDigital\\_VersaoFinal\\_2505.pdf](https://www.impacto.blog.br/site/wp-content/uploads/2021/05/Auxilio_InclusaoDigital_VersaoFinal_2505.pdf)

GREGOLIN, M. R. V. Identidade: objeto ainda não identificado? In: **Estudos da Língua(gem): imagens de discurso**. Vitória da Conquista, UESB, v.6, n.1, p.81-97, 2008. DOI: <https://doi.org/10.22481/el.v6i1.1058>

HALL, S. **Cultura e representação**. Rio de Janeiro: Ed. PUC Rio, Apicuri, 2016.

HARRIS, Z. S. Distributional Structure. **Word**, v. 10, n. 2-3, p. 146-162, 1954.

ISER, W. **O fictício e o imaginário**: perspectivas de uma antropologia literária. 2. ed. Rio de Janeiro: EdUERJ, 2013.

ISER, W. **How to do theory**. Malden, MA: Blackwell, 2006.

JORNAL NACIONAL. **Menos de um terço da população brasileira tem acesso pleno à internet, mostra pesquisa.** 18 março 2022. <https://g1.globo.com/jornal-nacional/noticia/2022/03/18/menos-de-um-terco-da-populacao-brasileira-tem-acesso-pleno-a-internet-mostra-pesquisa.ghtml>

KAHNEMAN, D.; MILLER, D. T. Norm theory: Comparing reality to its alternatives. In: **Psychological Review**, n. 93, p. 136-153, 1986.

KAHNEMAN, D.; TVERSKY, A. (Org.) **Choices, values and frame.** New York: Cambridge University Press, 2000.

LEAVITT, D. **O homem que sabia demais:** Alan Turing e a invenção do computador. Ribeirão Preto, SP: Novo Conceito, 2011.

LEVY, O.; GOLDBERG, Y., DAGAN, I. Improving Distributional Similarity with Lessons Learned from Word Embeddings. **Transactions of the Association for Computational Linguistics**, vol. 3, pp. 211–225, 2015. <https://aclanthology.org/Q15-1016.pdf>

LÉVY, P. Préface. In: VIAL, S. **L'Être et l'Écran.** Paris: PUF, 2013.

LUGER, G. F. **Inteligência artificial.** 6. ed. São Paulo: Pearson Education do Brasil, 2013.

MARI, A. Negros e pobres sofrem com exclusão digital durante a pandemia. **Forbes**, 27 maio 2020. <https://forbes.com.br/forbes-tech/2020/05/negros-e-pobres-sofrem-com-exclusao-digital-durante-a-pandemia/>

MBEMBE, A. **Brutalismo.** 2. ed. São Paulo: n-1 edições, 2022.

MARINHO, A.; CARDOSO, S. S.; ALMEIDA, V. V. **Desigualdade de transplantes de órgãos no Brasil: análise do perfil dos receptores por sexo e raça ou cor.** Brasília: IPEA, 2011. [https://repositorio.ipea.gov.br/bitstream/11058/1491/1/td\\_1629.pdf](https://repositorio.ipea.gov.br/bitstream/11058/1491/1/td_1629.pdf)

MELO, P. V.; TERSO, T. Dificuldades no acesso à internet: expressões do racismo estrutural. **Le Monde Diplomatique**, 24 março 2021. <https://diplomatique.org.br/dificuldades-no-acesso-a-internet-expressoes-do-racismo-estrutural/>

MORENO MARIMÓN, M.; SASTRE, G. **Como construimos universos:** amor, cooperação e conflito. São Paulo: Ed. Unesp, 2014.

NASCIMENTO, Gabriel. Entre o lócus de enunciação e o lugar de fala: marcar o não-marcado e trazer o corpo de volta na linguagem. **Trabalhos em**

- Linguística Aplicada**. v. 60, n. 1, pp. 58-68, 2021. <https://doi.org/10.1590/010318139578611520210313>.
- NOBLE, S. U. **Algoritmos da opressão**: como o Google fomenta e lucra com o racismo. Santo André, SP: Rua do Sabão, 2021.
- OBERMEYER, Z.; NISSAN, R.; STERN, M.; EANEFF, S.; BEMBENECK, E. J.; MULLAINATHAN, S. **Algorithmic Bias Playbook**. Chicago: Center for Applied AI at Chicago Booth, 2021. Disponível em: <https://www.chicagobooth.edu/-/media/project/chicagobooth/centers/caai/docs/algorithmic-bias-playbook-june-2021.pdf>
- O'NEIL, C. **Algoritmos de destruição em massa**: como o big data aumenta a desigualdade e ameaça a democracia. Santo André, SP: Rua do Sabão, 2020.
- PÊCHEUX, M. **Semântica e discurso**: uma crítica à afirmação do óbvio. Campinas, SP: Editora da Unicamp, 1988.
- PENG, R. B.; LEE, H.; KE, Z. T.; SAUNDERS, M. R. Racial disparities in kidney transplant waitlist appearance in Chicago: Is it race or place? In: **Clinical transplantation**, v. 32, n. 5, 2018. Disponível em: <https://doi.org/10.1111/ctr.13195>
- PENNINGTON, J.; SOCHER, R.; MANNING, C. GloVe: Global Vectors for Word Representation. In: **Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)**, pp. 1532–1543, Doha, Qatar, 2014. Association for Computational Linguistics. <https://nlp.stanford.edu/pubs/glove.pdf>
- ROJO, R. **Letramentos múltiplos, escola e inclusão social**. São Paulo: Parábola Editorial, 2009.
- SANTOS, C. C. A emergência do ativismo digital antirracista no Brasil. **Dissonância: Revista de Teoria Crítica**, v. 6, Campinas, 2022, p. 453-500.
- SANTOS, L. G. M.; COSTA, A. B.; DAVID, J. S.; PEDRO, R. M. L. Reconhecimento facial: tecnologia, racismo e construção de mundos possíveis. **Psicologia & Sociedade**. 2023, v. 35 Disponível em: <https://doi.org/10.1590/1807-0310/2023v35e277141>. Acessado 13 Outubro 2024
- SAID, E. **Orientalismo**: O Oriente como invenção do Ocidente. São Paulo: Companhia de Bolso, 2013
- SENO, E. R. M; CLARO, D.; MOTA, L.; RODRIGUES, J. **10 Semântica Distribucional**. Publicado em: 26 set. 2023. <https://brasileiraspln.com/livro-pln/1a-edicao/parte5/cap10/cap10.html>

- SILVA, D. F. O fenômeno dos refugiados no mundo e o atual cenário complexo das migrações forçadas. In: **Revista Brasileira de Estudos de População**, Belo Horizonte, v.34, n.1, p.163-170, jan./abr. 2017. DOI <http://dx.doi.org/10.20947/S0102-3098a0001>
- SILVA, T. **Comunidades, algoritmos e ativismos digitais: olhares afrodiaspóricos**. São Paulo: LiteraRua, 2021.
- SILVA, T. **Racismo algoritmo: inteligência artificial e discriminação nas redes sociais**. São Paulo: Edições SESC, 2022.
- SILVA, T. Linha do tempo do racismo algorítmico. In.: **Desvelar. Justiça racial, IA e tecnologias digitais**. 2024. Disponível em: <https://desvelar.org/casos-de-discriminacao-algoritmica/>
- SU, C.; ZHOU, H.; GONG, L.; TENG, B.; GENG, F.; HU, Y. Viewing personalized video clips recommended by TikTok activates default mode network and ventral tegmental area. In: **Neuroimage**. V. 237, 15, August 2021. <https://doi.org/10.1016/j.neuroimage.2021.118136>
- TRINDADE, L. V. P. **It is not that funny: critical analysis of racial ideologies embedded in racialized humour discourses on social media in Brazil**. 2018. PhD Thesis (Sociology) - University of Southampton, Southampton, 2018.
- TRINDADE, L. V. P. **Discurso de ódio nas redes sociais**. São Paulo: Jandaíra, 2022.
- TURING, A. M. On computable numbers with an application to the Entscheidungsproblem. In: **Proceedings of the London Mathematical Society**, v. 2-42, Issue 1, 1937, pp. 230-265. <https://doi.org/10.1112/plms/s2-42.1.230>
- VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, L.; POLOSUKHIN, I. Attention is all you need. In: **NeurIPS**, 2017. Disponível em: <https://doi.org/10.48550/arXiv.1706.03762>
- VYAS, D, A.; EISENSTEIN, L. G.; JONES, D. S. Hidden in Plain Sight – Reconsidering the Use of Race Correction in Clinical Algorithms. **The New England Journal of Medicine**. v. 383, n. 9. 2020, pp. 874-882. <https://www.nejm.org/doi/pdf/10.1056/NEJMms2004740>
- WITTGENSTEIN, L. **Philosophical investigations**. 3. ed. Oxford: Blackwell Publishing, 2002.



## NOTAS

---

<sup>1</sup>Uma importante análise na grande mídia foi feita no podcast *O Assunto*, apresentado pela jornalista Natuza Nery, no dia 15 de fevereiro de 2024. O título do programa foi: “Inteligência artificial: mulheres negras sofrem mais erros em abordagens de reconhecimento facial do que brancos” e pode ser conferido aqui: <https://g1.globo.com/podcast/o-assunto/noticia/2024/02/15/inteligencia-artificial-mulheres-negras-sofrem-mais-erros-em-abordagens-de-reconhecimento-facial-do-que-brancos.shtml>

<sup>2</sup>Em seu livro, Noble adiciona em vários momentos, como provas, as imagens (“prints”) das pesquisas feitas.

<sup>3</sup>A postagem na rede social Instagram, e a resposta da empresa Dicio, podem ser conferidas no seguinte link: [https://www.instagram.com/reel/Cn9l41vz3h/?utm\\_source=ig\\_web\\_copy\\_link&igsh=MzRIODBiNWFIZA==](https://www.instagram.com/reel/Cn9l41vz3h/?utm_source=ig_web_copy_link&igsh=MzRIODBiNWFIZA==)

<sup>4</sup>Ainda que não seja objeto de análise neste trabalho, convém indicar que a afirmação “O negro é lindo” traz, de um ponto de vista enunciativo, uma resposta a sua própria enunciação, pois ao afirmar que o negro é lindo o movimento antirracista pre-

cisa, primeiro, assumir o discurso racista de que o negro não é lindo, como base discursiva, a fim de elaborar sua afirmação-resposta. Uma boa discussão sobre *locus da enunciação* ocorre em Nascimento (2021).