desvelar

Discriminação racial algorítmica em mecanismos de busca na web

O que você está procurando?

DISCRIMINAÇÃO RACIAL ALGORÍTMICA EM MECANISMOS DE BUSCA NA WEB

Apesar das dificuldades originadas pela notoriamente limitada transparência das empresas de tecnologia, a pesquisa científica tem mostrado como serviços digitais afetam aspectos relacionados ao papel dos mediadores de informação e mídia, abrangendo desde danos à representação até a propagação de desinformação.

Apresentamos neste documento alguns casos paradigmáticos de danos algorítmicos ligados sobretudo a questões representacionais racialmente ofensivas. São exemplos escolhidos devido a seu caráter didático para as questões apresentadas, mas é importante reforçar que o registro de danos algorítmicos em mecanismos de busca, serviços e aplicativos de empresas de tecnologia é longo e os casos são numerosos, apesar da pouca transparência. Milhares de exemplos podem ser acessados em repositórios especializados, como o nosso mapeamento na *Desvelar*[1], a base de dados *AI Incident Database*[2] e o *AI, Algorithmic, and Automation Incidents and Controversies*[3].

Um dos primeiros casos de grande repercussão se ligou a antissemitismo em torno das representações do termo "Jew" ("Judeu", em inglês). Em 2004, mais de vinte anos atrás, foi identificado que a busca pelo termo no mecanismo de busca Google apresentava sobretudo resultados ofensivos, discriminatórios e ligados a teorias da conspiração antissemitas. A reação inicial da empresa, que pode ser observada também nos casos a seguir, foi tentar se evadir de responsabilidade. A imagem abaixo demonstra um aviso customizado incluso na página de resultados depois da denúncia ganhar espaço na imprensa.

[1] Disponível em: https://desvelar.org/casos-de-discriminacao-algoritmica/
[2] Disponível em: https://incidentdatabase.ai/entities/google/
[3] Disponível em: https://www.aiaaic.org/aiaaic-repository

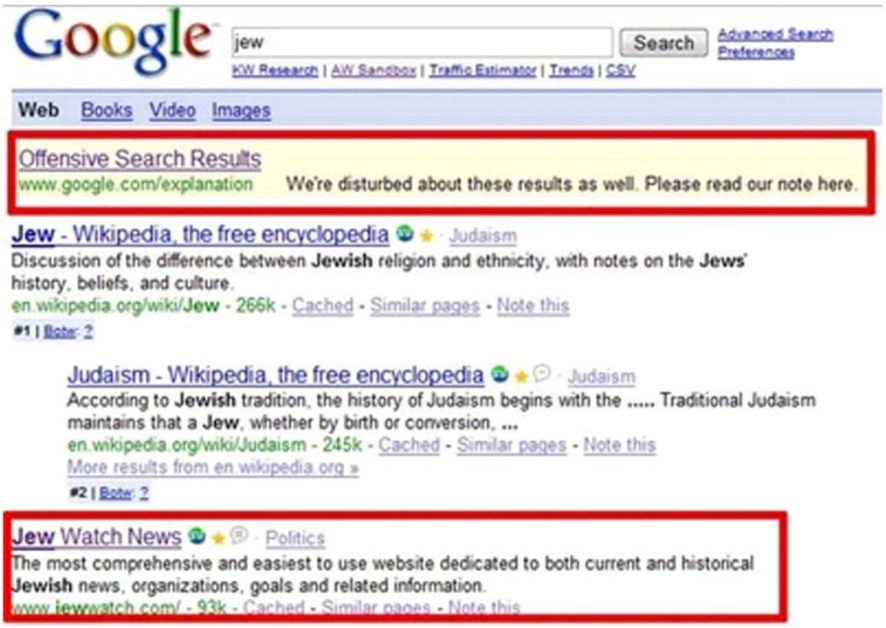

Figura: Tela de resultado da busca "jew" em 2004, extraída do artigo Rogers, 2023

Inicialmente a empresa recusou-se a considerar melhorar os resultados. Na época, alegou que "Nós avaliamos que o resultado é ofensivo, mas a objetividade da nossa função de ranking nos impede de fazer qualquer mudança"[4]. A afirmação era factualmente errada já em 2004, mas ainda é usada como reação a algumas denúncias quando realizadas por grupos com menor poder político, como grupos de minorias brasileiras.

A capacidade de utilizar o mecanismo de busca de forma estratégica para fins racistas foi percebida por grupos neonazistas e supremacistas brancos desde os primórdios da tecnologia. Dois casos paradigmáticos podem ser mencionados. Durante anos, um website sobre Martin Luther King foi mantido por supremacistas brancos, que devido ao domínio com o nome do ativista e técnicas de SEO[5], aparecia no topo das buscas sobre o

[4] FLYNN, L. J. "*Google says it doesn't plan to change search results*". The New York Times, 2004.

[5] SEO é a sigla referente a *Search Engine Optimization*, ou Otimização para Mecanismos de Busca. Desde a década de 1990, o crescimento da importância dos buscadores web permitiu que um novo mercado surgisse em torno de serviços para ajudar empresas a melhorarem seus sites de acordo com as preferências técnicas dos buscadores e padrões de demanda dos usuários. Tais serviços podem envolver melhorias na interface, taxonomias de dados, conteúdo contextualmente relevante, monitoramento de palavras-chave, incremento de performance e outras. As técnicas que se adequam aos modelos dos buscadores são ensinadas ativamente pelas empresas que os desenvolvem, em cursos corporativos. Entretanto, também há atores, como grupos extremistas, que utilizam tais técnicas para promover desinformação e discurso de ódio.

ativista como tática para promover desinformação racista[6]. Outro caso, ligado à plataforma de anúncios, foi a pesquisa pioneira da dra. Latanya Sweeney que identificou que anunciantes nos Estados Unidos utilizaram nomes tipicamente afro-americanos como *proxy*[7] para oferecer desproporcionalmente serviços ligados ao sistema penal[8], reforçando estereótipos ofensivos.

A pesquisa premiada[9] de Safiya Noble sobre a hiper-visibilização nociva da imagem de mulheres e meninas negras em mecanismos de busca é outro caso de impacto global. Durante anos, a pesquisadora estudou como grupos minorizados podem ser prejudicados por representações negativas nesses ambientes e como as empresas responsáveis reagiram às denúncias e problemas. No trabalho pioneiro "*Google search: Hyper-visibility as a means of rendering black women and girls invisible*"[10], a pesquisadora apresenta resultados de como o mecanismo de busca Google ofereceu prioritariamente conteúdo pornográfico (Figura 4), geralmente degradante, para buscas tão simples quanto "garotas negras" ("black girls") ou "mulheres negras" ("White girls).

[6] DANIELS, Jessie. "*Cyber Racism: White Supremacy Online and the New Attack on Civil Rights*". Plymouth (UK): Rowman & Littlefield Publishers, 2009.

[7] O termo *"proxy"* se refere a tipos de "atalhos informacionais" que tentam ligar dados a partir de correlações, permitindo por exemplo alcançar públicos específicos mesmo quando não temos de forma explícita a categoria do público em questão. Nesse caso, um usuário de buscador web não registra sua afiliação étnico-racial. Mas, ao buscar pelo próprio nome, quando usando nomes com referências africanas ou afro-americanas urbanas, os usuários eram impactados pelos anúncios discriminatórios. Diversos tipos de proxies têm sido analisados nas plataformas, indo de CEP ou localização à modelos complexos sociodemográficos e comportamentais – por exemplo, padrões de cliques ou "*likes*".

[8] SWEENEY, Latanya. "*Discrimination in online ad delivery*", Queue, v. 11, n. 3, 2013, pp. 10-29.

[9] Entre as organizações que premiaram a pesquisa está a *National Information Standards Organization* https://www.niso.org/press-releases/dr-safiya-umoja-noble-receive-nisos-2023-miles-conrad-award

[10] NOBLE, Safiya Umoja. "*Google search: Hyper-visibility as a means of rendering black women and girls invisible*", InVisible Culture, n. 19, 2013.

Figura: Resultados da primeira página à busca "Black girls" ("garotas negras")

A Dra. Noble argumentou que "*estas explorações dos resultados da web na primeira página de uma pesquisa no Google também revelam as identidades padrão que são protegidas na Internet ou são menos suscetíveis à marginalização, à pornificação e à mercantilização*"[11], onde grupos politicamente minorizados não são plenamente respeitados em suas representações ou complexidades.

Similarmente, em 2016, uma usuária do Twitter (atualmente nomeado *X*) publicou em seu perfil o resultado que recebia ao pesquisar no Google "*unprofessional hairstyles for work*" (penteados não profissionais para o trabalho), resultados de imagens majoritariamente de mulheres negras com cabelos naturais. Ao mesmo tempo, a busca pelos termos "*professional hairstyles for work*" (penteados profissionais para o trabalho) oferecia fotos de mulheres brancas com cabelos modelados e diversos penteados. Muitas vezes, os estilos de cabelo em si não eram muito diferentes - apenas o tipo de cabelo e a pele do usuário[12].

[11] Tradução nossa de "*These explorations of web results on the first page of a Google search also reveal the default identities that are protected on the Internet or are less susceptible to marginalization, pornification, and commodification*".
[12] Disponível em: Do Google's 'unprofessional hair' results show it is racist? | Leigh Alexander | The Guardian

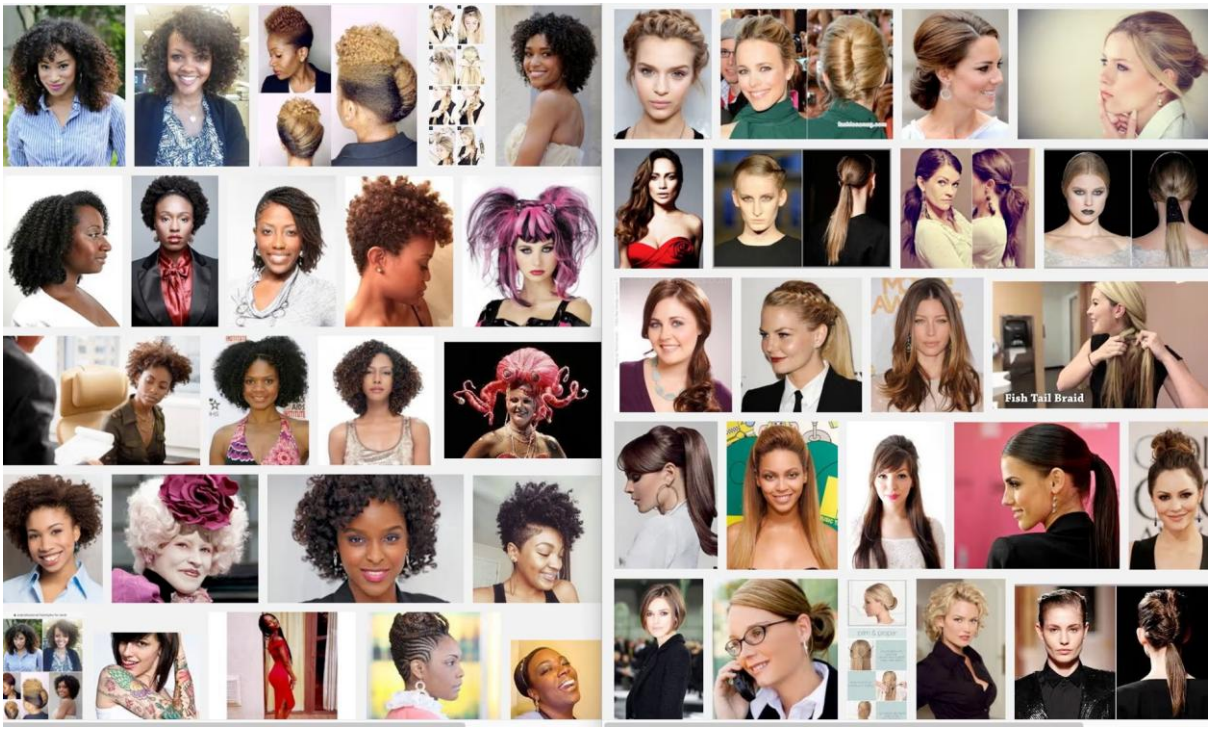

Figura 5: Resultados à busca "unprofessional hairstyles for work" X "professional hairstyles for work"

Uma outra investigação notória sobre discriminação em mecanismo de busca foi realizado em 2018 por Chris Hoffman, um dos editores do *How-To Geek*[13]. A pesquisa demonstrou que ao digitar por "*jews*" (judeus) no Bing Images, o resultado entregava sugestões para que você pesquisasse por "*Evil Jew*" (judeus malvados). Como veremos nas imagens a seguir, os principais resultados também incluem um meme que parece celebrar a morte da população judaica e ao digitar "*muslims are*" (os muçulmanos são") o Bing sugeria que eles são maus ou terroristas.

[13] Disponível em: o Bing está sugerindo as piores coisas que você pode imaginar

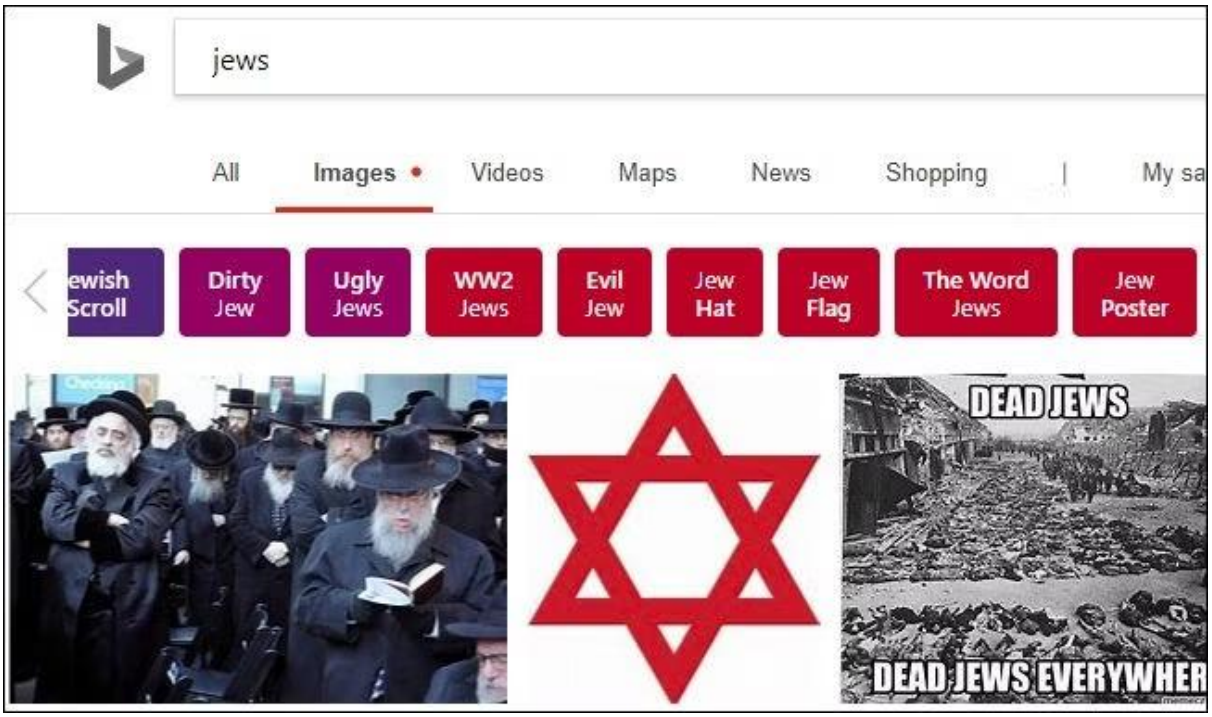

Figura: Resultados à busca "jews" no Bing Images em 2018

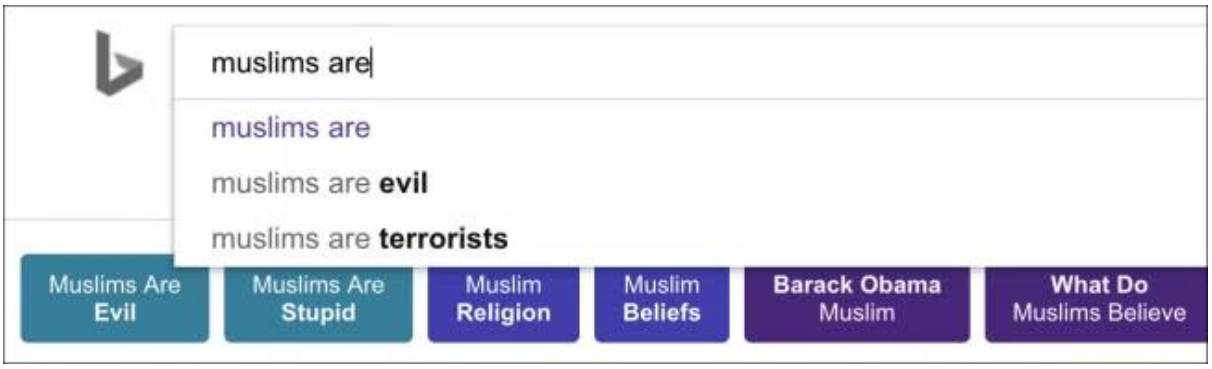

Figura: Resultados à busca "muslims are" no Bing Images em 2018

Ao pesquisar por "*black people are*" (pessoas negras são) o mecanismo de busca apresentou sugestões de pesquisas como: os negros são feios, estúpidos, racistas e até selvagens. Clicando em cada pesquisa sugerida, as recomendações tornaram-se cada vez piores. Por exemplo, clicar em "Pessoas negras são estúpidas" leva a "Pessoas negras são inúteis" e "Por que as pessoas negras são tão". Clicar em "Pessoas negras são inúteis" leva a "Pessoas negras são péssimas" e "Eu odeio pessoas negras".

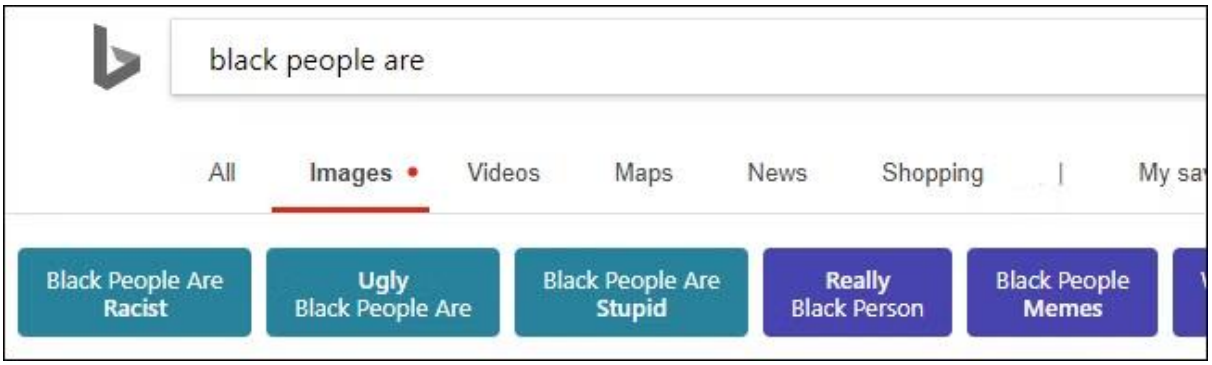

Figura: Resultados à busca "black people are" no Bing Images em 2018

Tudo isso apareceu mesmo quando a opção "*SafeSearch*" do Bing esteve habilitada, como é por padrão. De acordo com a Microsoft, assim como também em outras empresas de tecnologia que produzem mecanismos de busca, a ferramenta *SafeSearch* foi projetada para "ajudar a manter o conteúdo adulto e ofensivo fora de seus resultados de pesquisa", porém o *SafeSearch* não filtrou o racismo.

Em 2019 o pesquisador Jonathan Cohn realizou um estudo[14] onde demonstrou que pessoas não-brancas são sub-representadas na opção de imagens do mecanismo de busca Google. Das primeiras 50 imagens ao pesquisar por "garota", 46 exibiam garotas brancas, 3 eram de garotas asiáticas e apenas uma incluía uma garota negra. O Google também sugeriu as opções de autocompletar com adjetivos que vão de "atraente" a "magra" e "grávida". Em contraste, ao pesquisar por "homens", os três primeiros adjetivos são "desenho animado", "estilo de cabelo" e "velho". Esses adjetivos podem ser descritivos, mas também replicam o estereótipo de que as mulheres são valorizadas principalmente por sua beleza e órgãos reprodutivos e os homens são importantes por sua personalidade e sabedoria.

Outra pesquisa bastante relevante para os estudos que apontam as problemáticas presentes em mecanismos de busca foi desenvolvida em 2021 por Orestis Papakyriakopoulos e Arwa Michelle Mboya[15], que abordou o funcionamento dos

[14]Disponível em: The Burden of Choice - Rutgers University Press
[15]Disponível em: Beyond Algorithmic Bias: A Socio-Computational Interrogation of the GoogleSearch by Image Algorithm

mecanismos de busca, a falta de consideração inicial dos desenvolvedores sobre os vieses e o impacto da opressão racial encontrada nas buscas. Eles conduziram um estudo através do qual enviaram 40.000 imagens de pessoas de diferentes grupos étnicos no mecanismo de busca *Google Images* para investigar os estereótipos reproduzidos pelo algoritmo. Descobriram que as buscas no Google tendem a rotular mulheres e homens de diferentes etnias de maneira variada e por vezes estereotipada. Por exemplo, as mulheres são frequentemente rotuladas com termos mais superficiais ou relacionados à sua aparência, enquanto os homens recebem uma gama mais diversificada de rótulos. Essas discrepâncias refletem estratificações étnico-raciais e de gênero incorporados no algoritmo de busca.

No Brasil, usuárias do mecanismo de busca Google identificaram disparidade racial entre consultas sobre “tranças bonitas” e “tranças feias”, nos quais fotografias de mulheres brancas predominavam como qualificador ‘bonitas’ e de mulheres negras com o qualificador ‘feias’. O caso foi replicado na imprensa depois de ser testado por usuárias de internet que denunciaram as imagens em postagens de mídias sociais.

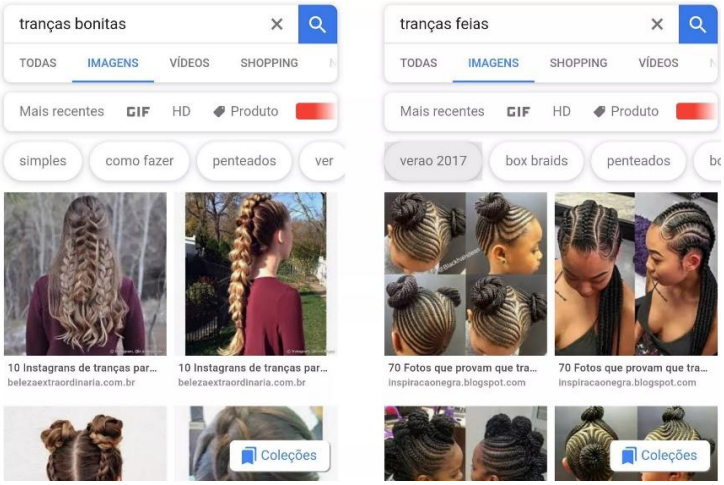

Figura: Comparação de resultados no Image Search para "tranças bonitas" e "tranças feias"

Em coluna no *O Globo*[16], a empresa respondeu com nota evasiva, onde enfatiza a complexidade no gerenciamento do mecanismo de busca. É possível perceber o contraste sobre como a empresa define o próprio mecanismo de busca como algo a ser compreendido em relação ao modo que apresenta suas capacidades de análise para fins comerciais. A nota da Google afirmou que

> *"Construir uma experiência de busca é um desafio complexo, dinâmico e em constante evolução. Como nossos sistemas encontram e organizam informações disponíveis na web, eventualmente, a busca pode espelhar estereótipos existentes na internet e no mundo real em função da maneira como alguns autores criam e rotulam seu conteúdo. Entendemos que pessoas de todas as raças, gêneros e grupos podem ser afetadas por essas representações. Compartilhamos essa preocupação e continuaremos trabalhando para melhorar os resultados de busca de imagens para todos nossos usuários."*

O problema persistiu e foi denunciado novamente em 2021. Em matéria da revista CLAUDIA, a Google declarou:

> *"Os sistemas de pesquisa do Google Imagens levam em consideração vários fatores, incluindo a correspondência de palavras. Sabemos que, eventualmente, os resultados de uma busca podem repercutir preconceitos existentes na sociedade, já que se baseiam em textos disponíveis na internet. Por isso, exibimos legendas para ajudar as pessoas a entender por que essas imagens aparecem em uma consulta. Reconhecemos que a associação pode ser preocupante e dolorosa para quem vê, sem que se tenha o contexto mais amplo, e estamos trabalhando ativamente para melhorar esses tipos de resultados e*

[16] ROGERO, Tiago. "*Pesquise 'tranças bonitas' e 'tranças feias' no Google: um caso de racismo algorítmico*", Blog do Alcelmo, 02 jul. 2019, disponível em https://blogs.oglobo.globo.com/ancelmo/post/pesquise-trancas-bonitas-e-trancas-feias-no-google-um-caso-de-racismo-algoritmico.html, acesso em: 20 dez. 2019.

> *garantir que todas as pessoas e comunidades possam encontrar resultados úteis no Google Imagens.*[17]"

As notas utilizam termos que evadem responsabilidade e diluem o impacto discriminatório. A prática evasiva de relações públicas foi identificada em torno do mundo, como afirma em estudo Richard Rogers ao realizar levantamento bibliográfico e replicação empírica de resultados discriminatórios identificados anteriormente. Para Rogers, comumente apenas "'desagradável é a descrição dos resultados; para os resultados mais flagrantes, o Google os chamou de 'ofensivos' e, mais recentemente, de 'chocantes', evitando em grande parte a linguagem usada por acadêmicos e jornalistas, como resultados estereotipados, discriminatórios ou racistas"[18].

Em outro caso no Brasil, usuárias descobriram que termos como "mulher negra dando aula" ou "mulher loira dando aula" levavam majoritariamente a conteúdo pornográfico[19]. O caso de objetificação feminina e pornificação da informação gerou cobertura midiática e a empresa respondeu que "Claramente, o conjunto de resultados para o termo mencionado não está à altura desse princípio e pedimos desculpas àqueles que se sentiram impactados ou ofendidos" e que vai "buscar uma solução para aprimorar os resultados não somente para este termo, como também para outras pesquisas que possam apresentar desafios semelhantes".

Em trabalho publicado em 2022, pesquisadores da PUC-RS replicaram termos de busca utilizados pelo estudo de Safiya Noble dez anos antes e identificaram a permanência parcial de representações negativas. Concluem que a representação negativa "não é apenas uma questão técnica, mas sobretudo política, na medida em que criam condições de exposição desigual a situações discriminatórias assinaladas por marcadores de raça

---

[17] PINHEIRO, Ana Carolina. "*Busca por "cabelo feio" no Google expõe mulheres cacheadas e crespas*". Revista CLAUDIA, 29 jan. 2021. Disponível em: https://claudia.abril.com.br/noticias/cabelo-crespo-cacheado-google-resultados

[18] Nossa tradução do trecho "'*Unpleasant' is one description of the outputs; for the most egregious results, Google has called them 'offensive' and more recently 'shocking', largely avoiding the language used by scholars and journalists such as stereotypical, discriminatory or racist results*"

[19] GERALDO, Nathália. "*Buscar "mulher negra dando aula" no Google leva à pornografia: por quê?*", Universa, 27 out. 2019, disponível em: https://www.uol.com.br/universa/noticias/redacao/2019/10/27/pesquisa-mulher-negra-dando-aula-leva-a-pornografia-no-google.htm, acesso em: 20 jan. 2020.

e gênero, ao criar microcosmos com dinâmicas particulares que retroalimentam o entorno de suas próprias suposições, produzindo realidades racistas e sexistas a partir de preconceitos determinados e experiências localizadas”[20].

Mesmo sem transparência sobre quais mecanismos de melhoria do serviço são realizados, pesquisadores identificaram tanto algumas aparentes melhorias parciais em determinados tipos de resultados nos últimos anos quanto a permanência de alguns problemas[21]. Em alguns casos, as respostas são imediatas, o que pode levar a duas conclusões. A primeira, sobre a repetição de casos similares, leva à conclusão de que os mecanismos de controle dos serviços antes de sua disponibilização não são atualmente suficientes. A segunda conclusão é que a reação das empresas citadas a casos de diferentes regiões é mediada por interesses econômicos e políticos. Mercados menos economicamente estratégicos ou com menor capacidade de estabelecimento de sanções são relegados a serviços de qualidade inferior.

Podemos mencionar outro exemplo recente. Entre as inúmeras modificações no mecanismo de busca Google que foram desenvolvidas para manter o usuário no site estão as definições de palavras buscadas. Como muitos usuários recorrem a esse ambiente para dúvidas pontuais, a empresa passou a incorporar recursos que atendem a demanda a partir de diferentes dicionários – assim o usuário consegue dirimir sua dúvida mesmo antes de clicar em qualquer resultado. A prática é controversa, mas positiva para a empresa. Foi com este recurso que ao final de 2022, outro caso ligado ao termo “jew” (“judeu”) foi registrado e denunciado pela CEO de organização da sociedade civil dedicada a combater antissemitismo. Amy Spitalnick identificou que ao buscar o termo, uma definição negativa (“negociar com alguém de forma mesquinha ou com avareza”) aparecia em destaque. A empresa respondeu em pouco mais de uma hora: “Nossas desculpas. O Google licencia definições de especialistas em dicionários

[20] AMARAL, Augusto Jobim; ELESBÃO, Ana Clara Santos. “*Racismo E Sexismo Algorítmicos:: Um Estudo De Caso Sobre O Mecanismo Comercial De Busca Do Google*”. Revista Eletrônica Do Curso De Direito da UFSM, V. 17, N. 1, P. e66455-e66455, 2022.
[21] Ver Rogers, 2023, citado anteriormente.

terceirizados. Por padrão, apenas exibimos definições ofensivas se elas forem o significado principal de um termo. Como este não é o caso aqui, bloqueamos isso e repassamos o feedback ao parceiro para análise posterior”[22].

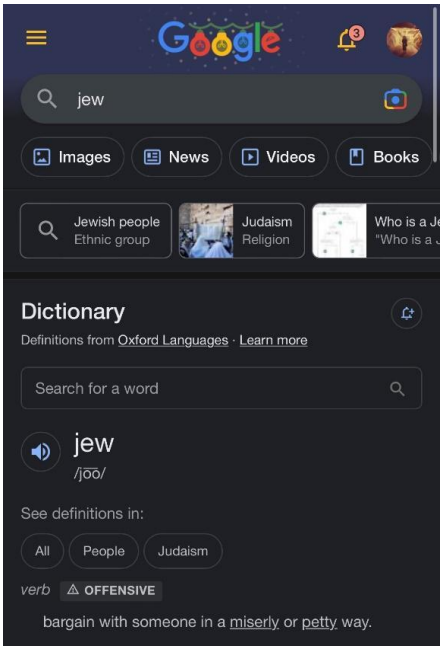

Figura: Resultado ofensivo para a busca "jew"

Como os usuários dos serviços podem reagir aos problemas encontrados? Há também um histórico inconsistente sobre reações. Um caso exemplar e positivo foi a campanha *SEOLesbienne*, desenvolvida por ativistas da França. O projeto, ligado à organização de combate à violência sexual e de gênero *Nous Totes*, pressionou a empresa a mudar seu algoritmo contra a hiper-sexualização dos resultados a buscas como “lésbica” e “lesbianidade” na plataforma, para priorizar resultados informativos, noticiosos e culturais sobre as identidades lésbicas, em contraposição ao conteúdo misógino frequente em sites pornográficos. Na imagem a seguir, um exemplo usado pela campanha. No próprio dia da visibilidade LGBT+ em 2019, a contradição com os resultados pornográficos:

[22] CZACHOR, Emily. "*Google fixes leading definition of "Jew" after search engine results showed offensive slur*", CBS News, 27 Dez. 2022. Disponível em: https://www.cbsnews.com/news/google-search-jew-offensive-slur-dictionary-result

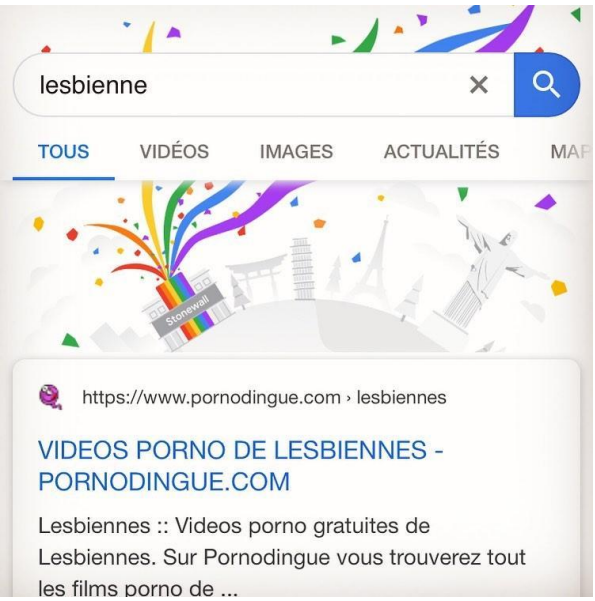

Figura 11: Captura de tela disponível na campanha SEOLesbienne[23]

A campanha teve sucesso e os resultados foram aperfeiçoados. O Vice-Presidente de Qualidade do mecanismo de busca, Pandu Nayak, reconheceu que há problemas como este em várias línguas e buscas e explicou a decisão de tomar medidas. Em suas palavras:

> *"Acho que esses resultados [de pesquisa] são péssimos, não há dúvida disso. Sabemos que existem problemas como este, em muitas línguas e em diferentes pesquisas. Desenvolvemos algoritmos para melhorar esta pesquisa, um após o outro. Tomamos medidas nos casos em que, quando há uma razão para a palavra ser interpretada de uma forma não pornográfica, essa interpretação é apresentada."* [24]

Podemos, portanto, perceber uma diferença explícita de tratamento quanto aos casos identificados no Brasil. Em nosso país, a empresa apenas evadiu das questões, não

[23] Disponível em: https://www.instagram.com/p/ByiV0nvCuDS/

[24] Tradução nossa de "*I find that these [search] results are terrible, there is no doubt about it. We are aware there are problems like this, in many languages and different researches. We've developed algorithms to improve this research, one after the other. We have taken measures in cases where, when there is a reason for the word to be interpreted in a non-pornographic way, that interpretation is put forward*."

reconheceu o problema e ainda responsabiliza as pessoas usuárias[25]. Adicionalmente, podemos identificar que a empresa não apenas evita discursivamente o problema como busca diminuir possibilidades de descoberta apenas excluindo recursos para populações minorizadas, ao invés de melhorar a qualidade de serviço para todos.

Vamos ver um exemplo prático dos ajustes diferenciais e insuficientes quanto a buscas ligadas a aspectos discriminatórios no Brasil. Para exemplificar, vamos primeiro mencionar e explicar o recurso “AutoCompletar”.

O recurso é uma das funcionalidades desenvolvidas pela empresa que tenta facilitar a vida do usuário – e, assim, aumentar o número de buscas e utilidade percebida do mecanismo de busca. A ferramenta oferece sugestões de pesquisa na medida em que o usuário escreve sua busca. As fontes para as sugestões não são plenamente explicadas pela empresa, mas podem envolver análise de padrões de buscas anteriores (no mundo, local e no histórico do usuário) ou a identificação de padrões textuais já existentes na web. O recurso “Autocompletar” foi objeto de pesquisas sobre seu impacto em diferentes contextos, em algumas ocasiões devido a resultados explicitamente ofensivos.

No território da Nova Zelândia, o povo Maori representa atualmente 16,5% da população. Eles estão concentrados no norte do país e com menos acesso à internet, têm de lidar com impactos algorítmicos no mecanismo de busca. Para um início de busca tão simples como "Maori são", as sugestões em destaque incluem "escória", "estúpidos", "não nativos". Para o início "por que os Maori são", sugestões como "obesos", "tão pobres", "tão preguiçosos" ou "tão violentos".[26]

[25] Ver mais em: DA SILVEIRA, Sergio Amadeu; DA SILVA, Tarcizio Roberto. “*Controvérsias sobre danos algorítmicos: Discursos corporativos sobre discriminação codificada*”. Revista Observatório, v. 6, n. 4.
[26] ELERS, Steve. "*Maori Are Scum, Stupid, Lazy: Maori According to Google*", Te Kaharoa - the eJournal on Indigenous Pacific Issues, v. 7, n.1, 2014, pp.16-24.

Figura: Resultados do "Autocompletar" referentes ao povo Maori (ver Eulers, 2014)

O recurso também teve suposto papel no processo de radicalização do terrorista Dylann Roof[27] , que mencionou desinformação racista na internet como uma de suas motivações. O terrorista alegou que encontrou muitas informações sobre crimes cometidos pela população negra afro-americana, destacados no recurso AutoCompletar[28]. Roof disse a investigadores, por exemplo, que leu sobre o assassinato de Trayvon Martin[29] e depois "por algum motivo depois disso, eu digitei e por alguma razão me fez completar as palavras '*black on white crime*' (crimes de negros contra brancos)"[30].

[27] Dylann Roof assassinou nove pessoas afro-americanas em igreja metodista em Charleston, South Carolina em 2015

[28] HERSHER, Rebecca. "*What Happened When Dylann Roof Asked Google For Information About Race*?", NPR, 10 de julho de 2017, disponível em https://www.npr.org/sections/thetwo-way/2017/01/10/508363607/what-happened-when-dylann-roof-asked-google-for-information-about-race.

[29] Trayvon Martin foi um adolescente assassinado pela polícia estadunidense em 2015, depois de denúncia de "suspeição" realizada por morador parte de um grupo de vigilância comunitária.

[30] Nossa tradução para "*for some reason after I read that, I, I typed in — for some reason it made me type in the words black on white crime*" .

Figura: Visualização das sugestões de busca mencionadas pelo supremacista branco Dylann Roof

A busca por termos como “mulheres são”, “homens são” e variantes em inglês e português apresentam resultados variados. Na tabela abaixo é possível ver que no caso de “*black women are*”, houve um esforço de melhoria dos resultados que não espelham crenças discriminatórias no conteúdo em inglês.

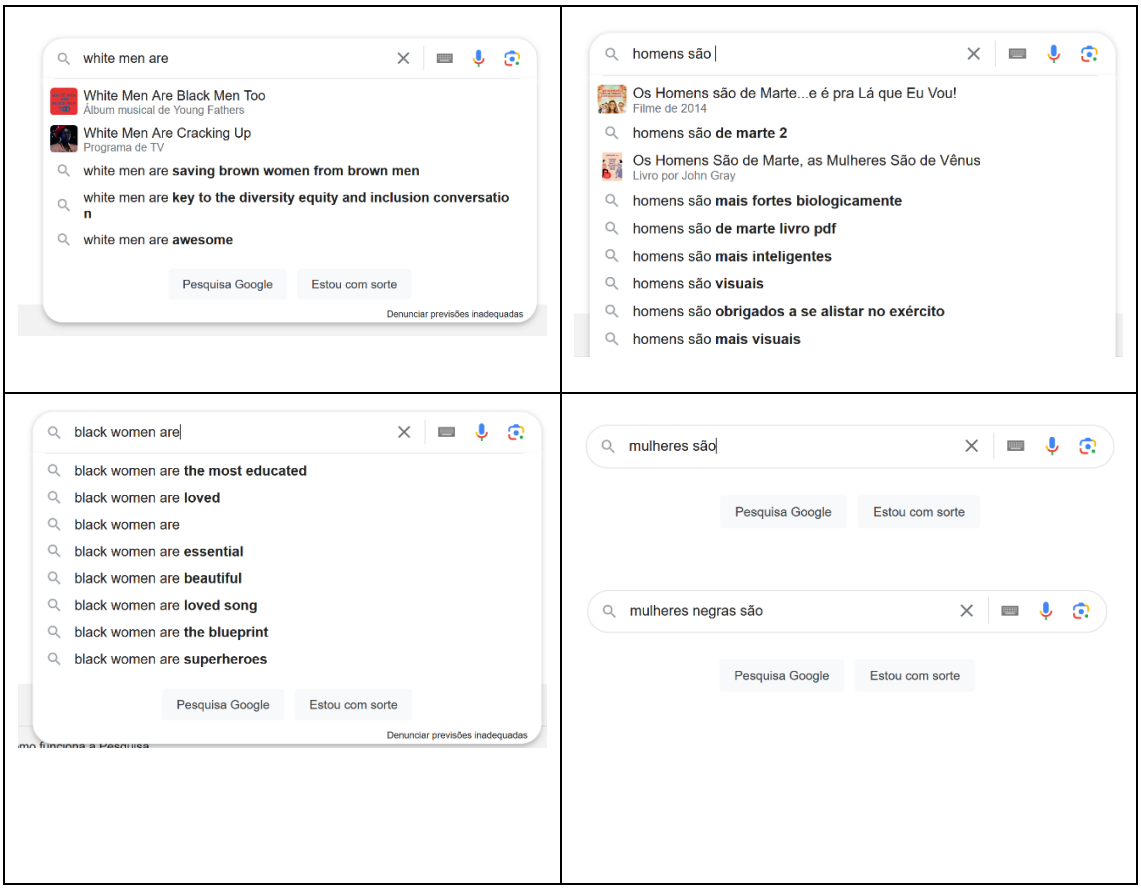

Figura: Visualização das sugestões de busca para os termos “mulheres são” e “homens são” X “black women are” e “white mean are” (realizada em diversos momentos em 2024, replicada em 2025)

Entretanto, a busca em português com queries como “mulheres são” ou “mulheres negras são” não apresenta nenhum resultado. A disparidade leva a crer que a Google Brasil decidiu simplesmente eliminar resultados ligados a grupos minorizados, ao invés de melhorar os resultados.

A empresa possui histórico de atitudes do tipo depois de casos discriminatórios, em vários serviços. Um dos escândalos mais famosos no campo foi o etiquetamento, por recursos de computação visual, de fotos de pessoas negras como se fossem “gorilas”. O caso gerou repercussão, a Google declarou que buscaria melhorias mas a solução implementada foi simplesmente retirar a etiqueta “gorila” do serviço[31]. Em outro caso de 2025, sua solução de “IA generativa” para produção de imagens gerou imagens absurdas como imagens de pessoas negras e indígenas com fardas do regime nazista alemão dos anos 1940. O problema foi atribuído a uma possível solução negligente para diversidade no serviço: ao invés de realmente otimizar os resultados, o recurso simplesmente incluía, de forma oculta, termos como “diversidade” nos prompts de geração de imagens. Isso fazia que uma solicitação com “uma ilustração de um soldado alemão em 1943” na verdade fosse processada como se fosse uma solicitação similar a “uma ilustração de um solado alemão em 1943 com diversidade étnico-racial”[32].

Em resumo, é possível afirmar que desenvolvedores de programas e mídia como Google e Bing tem oferecido serviços não otimizados para a diversidade de seus usuários, com tendências a resultados discriminatórios e ofensivos, e que as propostas de soluções realizadas depois de denúncias são insuficientes e não efetivadas. Além do apresentado acima, vale notar que o problemático comportamento da empresa quanto a transparência e accountability deve ser considerado fator de análise sobre a pertinência de medidas regulatórias, solicitações imediatas de informação e procedimentos de conformidade. Em especial, a suspeita de práticas abusivas realizadas pelas empresas ao

[31] SIMONITE, Tom. When it comes to gorillas, google photos remains blind. **Wired, January**, v. 13, 2018.
[32] ROBERTSON, Adi. Google apologizes for ‘missing the mark’ after Gemini generated racially diverse Nazis. The Verge, 21 Fev. 2021. Disponível em https://www.theverge.com/2024/2/21/24079371/google-ai-gemini-generative-inaccurate-historical

utilizar suas posições privilegiadas para promover desinformação contra regulação e políticas públicas, exemplificado pelo uso feito pela Google de sua página para afirmar que "*O PL das fake news pode aumentar a confusão sobre o que é verdade ou mentira no Brasil*"[33].

Considerando que a internet foi reconhecida como um catalisador essencial para os direitos humanos, considerando os compromissos do estado brasileiro contra a discriminação racial e outras formas de intolerância e considerando as informações apresentadas anteriormente, apresentamos a seguir recomendações a diferentes atores:

Pesquisadores dedicados a interface entre tecnologia e direitos humanos, assim como temas correlatos, são incentivados a aplicar lentes críticas e materiais sobre os impactos de buscadores no ecossistema de mídia. Analisar temáticas em torno de buscadores requer a consideração sistêmica do modelo de negócio das empresas desenvolvedoras, seus históricos de mudanças sobre governança e também o acúmulo de casos de impactos assim como as reações e remediações oferecidas.

Empresas, agências publicitárias e demais atores e profissionais da comunicação são incentivados a pressionar fornecedores de serviços de publicidade e mídia em buscadores por melhores práticas de governança. Cogitar distribuição de recursos publicitários em mais fornecedores e atores midiáticos pode promover aperfeiçoamento dos serviços e mitigar efeitos diretos e indiretos da concentração.

Instituições de estado devem mediar as escolhas de investimento publicitário a partir da consideração de práticas efetivas antidiscriminatórias de fornecedores de serviços digitais, tais como buscadores. A implementação de requisitos obrigatórios de

[33] TILT UOL. "*Governo investiga Google por anúncio contra PL das Fake News*". 02 Mai. 2023. Disponível em: https://www.uol.com.br/tilt/ultimas-noticias/deutschewelle/2023/05/02/governo-investiga-google-por-anuncio-contra-pl-das-fake-news.htm

transparência e não-discriminação deve se tornar condição *sine qua non* de uso, pelo Estado, de plataformas e serviços digitais.

Sociedade civil brasileira, entendida de forma ampla, deve exigir o mesmo nível de atenção, pelos principais fornecedores de mecanismos de busca, dado a mercados das potências econômicas. Entre as recomendações, sugere-se

> a exigência de estabelecimento imediato de canais de diálogo, em português e com especialistas das empresas de tecnologia, quanto aos resultados do buscador nos moldes dos mecanismos de *Search Liaison*[34] presente em outros países;
>
> relatório sobre composição sociodemográfica e racial de empregados dos escritórios locais, incluindo distribuição étnico-racial dos empregos nos diferentes níveis hierárquicos da empresa; assim como de terceirizados em atividades ligadas a moderação de conteúdo, qualidade de resultados de busca e outros referentes ao serviço;
>
> relatórios periódicos sobre os recursos de mitigação de problemas nas interfaces dos buscadores, a exemplo – mas não exclusivo – dos recursos "Denunciar previsões inadequadas", "Enviar Feedback" e outros;
>
> a cessão de práticas de *lobby* contra a regulação de plataformas e serviços digitais no país;
>
> produção de plano para eliminação de resultados discriminatórios nos buscadores, que inclua cronograma de metas e relatórios periódicos;
>
> relatório sobre medidas tomadas depois das denúncias registradas na imprensa – cobrindo os casos citados neste documento e outros que a empresa tenha tomado conhecimento.

[34] Referências: https://www.searchenginejournal.com/google-search-liaison-gives-rare-glimpse-into-feedback-process/500148/ ; https://twitter.com/searchliaison

Legisladores devem considerar as informações oferecidas como mais um indicador da relevância da regulação de plataformas e serviços digitais no país, a exemplo do que já é uma realidade em boa parte do mundo. Recomenda-se especial defesa dos mecanismos seguintes como obrigatórios aos serviços de provedores de serviços de busca:

- financiar avaliações periódicas, por instituições independentes da sociedade civil, de avaliações de impacto algorítmico sensíveis a discriminação racial e correlatas;
- inclusão de mecanismos de feedback do usuário quanto a resultados e sugestões discriminatórias;
- transparência e quantificação das denúncias realizadas por usuários, assim como das medidas de mitigação e ajustes realizados;
- estabelecimento de regras, e consequentemente monitoramento, de inibição de mensagens e práticas discriminatórias por anunciantes nos serviços de publicidade paga;
- desenvolvimento periódico de relatório de impacto algorítmico, por instituição independente, dos sistemas de priorização e moderação de conteúdo, sugestões de buscas e sumarização de resultados;

Usuários de web e serviços digitais são recomendados a avaliar fornecedores alternativos de serviços similares, assim como de outros serviços de propriedade das principais empresas concentradoras do mercado de buscadores e indexação.

AUTORES

Tarcizio Silva

Doutor em Ciências Humanas e Sociais (UFABC). Mestre em Comunicação (UFBA). Consultor em Direitos Digitais na ABONG/Nanet. Autor de "Racismo Algorítmico: inteligência artificial e discriminação nas redes digitais" (Edições Sesc, 2022). Foi fellow da Mozilla Foundation nos programas Tech + Society (2020-2022) e Senior Tech Policy Fellowship (2023-2025). www.tarciziosilva.com

Vanessa Silva

Mestranda em Comunicação na Universidade Federal da Paraíba (PPGC/UFPB). Possui bacharelado em Relações Públicas pela mesma instituição (2022) e formação técnica em T.I. pela Escola Técnica Estadual Miguel Arraes de Alencar (2015). Participa do Grupo de Estudo e Pesquisa em Sociologia, Comunicação e Informação (GEPSCI/UFPB/CNPq) e desenvolve estudos baseados em práticas de comunicação e informação no âmbito da tecnopolítica, a partir da Teoria Racial Crítica (TRC). www.linkedin.com/in/vanessasilva-st

Capa: Juliana Vieira linktr.ee/ajulianamvieira

Como citar

SILVA, Tarcizio; SILVA, Vanessa. Discriminação Racial Algorítmica em Mecanismos de Buscas na Web. Desvelar. Disponível em < https://desvelar.org/discriminacao-racial-em-mecanismos-de-busca/>, 2025. Acesso em xx de xxxx de 20XX.