

Comunicação Organizacional, Racismo Algorítmico e Inteligência Artificial: Violência nas Imagens da Mulher Negra na Plataforma Dall-E, da OpenAI ¹

Victória Emanuelle Bitencourth FRANÇA ^{1 2}

Rita de Cássia Barbosa LOUBACK ^{2 3}

Polyana Inácio Rezende SILVA ^{3 4}

Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, MG

Resumo

A plataforma DALL-E, da OpenAI, gera imagens digitais por meio de comandos de texto. No ano de 2023, a deputada Renata Souza solicitou que a ferramenta gerasse a imagem de uma mulher negra e periférica, no estilo da Disney Pixar. A resposta mostrou um arquivo estereotipado e preconceituoso, denunciado pela parlamentar como racismo algorítmico (SILVA, 2021;2022). O trabalho investiga de que forma o racismo praticado pela DALL-E se presentifica nas soluções algorítmicas da OpenAI e na discriminação sistêmica da mulher negra (Bento, 2022). A metodologia se fundamentou na Análise de Conteúdo em Bardin (2016) para elaborar um estudo de caso (Duarte, 2006) sobre a repercussão da imagem na imprensa e no X, antigo Twitter. Os resultados analisam os elementos do debate e verifica que o racismo algorítmico se perpetua nas práticas e discursos sobre Inteligência Artificial.

Palavras-chave: Inteligência Artificial; DALL-E; Racismo Algorítmico; Comunicação Organizacional; OpenAI.

Introdução

Em postagem⁵ da OpenAI em 2023, a organização informou que a segunda versão do algoritmo da ferramenta DALL-E 2 estava gerando mais de dois milhões de imagens por dia antes de ser liberada para o público geral. Os usuários da versão beta forneceram *feedbacks* sobre o desempenho da ferramenta. Alguns destes retornos

¹ Trabalho apresentado no Espaço Graduação, atividade integrante do XIX Congresso Brasileiro Científico de Comunicação Organizacional e Relações Públicas.

² Graduada em Relações Públicas e membro do grupo Dialorg pelo PPGCOM da PUC Minas, e-mail: victoria.francca@gmail.com.

³ Doutora em Comunicação Social, Professora na Faculdade de Comunicação e Artes da PUC Minas, e-mail: ritablouback@gmail.com.

⁴ Doutora em Comunicação Social e Sociabilidade, Professora na Faculdade de Comunicação e Artes da PUC Minas, e-mail: polyanainacio@gmail.com.

⁵ Disponível em: <<https://openai.com/about/>>. Acesso em: 15 mar. 2025.

mencionaram correções relativas a resultados preconceituosos e de conteúdo inadequado⁶. A OpenAI destacou uma equipe de moderadores para monitorar os resultados gerados pelo DALL-E com o intuito de ensinar o algoritmo a ignorar certas palavras, não produzir imagens de figuras públicas e evitar conteúdos violentos. Ainda em 2023 foi anunciada a terceira versão DALL-E 3. Ela apresentava maior capacidade de compreender detalhes em comparação com a versão anterior. Isso permitiu à plataforma gerar imagens mais coerentes com as instruções dos moderadores.

Renata Souza, deputada estadual do Rio de Janeiro solicitou ao DALL-E para explorar criar imagens personalizadas de mulheres negras e periféricas, inspiradas nos filmes da Disney Pixar. Contudo, o resultado gerado pela ferramenta foi uma representação de uma mulher negra segurando uma arma, reproduzindo uma imagem estereotipada e que associa esta mulher ao universo criminal. Isto levou Renata a denunciar o resultado como um caso de racismo algorítmico.

Neste sentido a questão central da investigação é: de que forma o racismo algorítmico praticado pela DALL-E no caso apontado por Renata Souza evidencia a discriminação sistêmica da plataforma digital, e como esta experiência pauta os desafios de combate ao racismo e outras violências associadas à mulher negra?

A abordagem metodológica do trabalho consiste no estudo de caso mapeado em uma imagem elaborada pelo DALL-E e uma análise de sua repercussão no Twitter. As novas tecnologias e o papel da organização como agente de mudança e transformação na sociedade também serão mencionados.

Dall-e: Racismo Algorítmico, Comunicação Organizacional e Práticas Discursivas

O termo "racismo algorítmico", de acordo com Silva (2022), descreve tecnologias desenvolvidas em contextos de privilégio branco que perpetuam estruturas racializadas e reforçam uma falsa ideia de superioridade. Os recursos de AI podem ser vistos como pequenos avanços tecnológicos, mas, na verdade, representam a

⁶ Disponível em:

<<https://olhardigital.com.br/2023/07/07/internet-e-redes-sociais/ia-usa-base-de-dados-que-perpetua-preconceitos-e-desigualdades/>>. Acesso em: 25 jan 2025.

concentração de poder de representação e financeiro, e servem, em determinados casos, como instrumentos de violência no âmbito mundial, reforçando atitudes de controle (Silva, 2022).

A plataforma de Inteligência Artificial generativa de imagens DALL-E é um tipo de rede neural artificial que utiliza modelos matemáticos de Aprendizado de Máquina (Machine Learning), uma das subáreas mais importantes dos estudos de IA, e é inspirado no funcionamento do cérebro humano. A ferramenta funciona por meio de um conjunto de algoritmos programados para criar imagens a partir de uma descrição textual também chamada de prompt.

Por meio desse método de Aprendizado de Máquina, a IA consegue aproximar-se do desempenho humano em diversas tarefas. No entanto, é igualmente possível realizar correções e evitar que novos casos de violência e discriminação sejam gerados. Para corrigir o funcionamento das ferramentas de IA, é necessário que a máquina aprenda por meio de exemplos repetidos, conforme proposto por Mitchell (1997 apud Ludermir, 2021). No caso do DALL-E, os algoritmos permitem que a ferramenta gere imagens conforme os comandos informados durante o processo de aprendizagem.

A maneira como esse fator impacta a experiência digital das pessoas negras, por meio do aprendizado de máquina invejada, é um elemento determinante na produção de casos discriminatórios, existentes desde o período colonial do século XVI e atualizados para a era das tecnologias. A discriminação de grupos sociais marginalizados está presente nos códigos computacionais e, consequentemente, na inteligência artificial. Além da invisibilidade desses grupos, há uma hipervisibilidade negativa, reforçando estereótipos (Gomes, 2023, p. 11).

Reproduções racistas também resultam de uma grande falta de representatividade de grupos minoritários no desenvolvimento de algoritmos e da ausência de uma análise crítica por parte de engenheiros e desenvolvedores das plataformas. De acordo com Lopes (2023), pesquisadores da startup de IA Hugging Face e da Universidade de Leipzig desenvolveram ferramentas para avaliar preconceitos em plataformas de IA que geram imagens, como o DALL-E 2 e as versões recentes do Stable Diffusion.

Plataformas como a DALL-E já passaram por testes de correção após resultados preconceituosos que reforçavam estereótipos. Em um estudo, 96 mil imagens foram geradas a partir da descrição de pessoas de diferentes etnias, gêneros e profissões. O estudo revelou que os modelos tendem a produzir imagens de pessoas brancas e masculinas, especialmente ao retratar indivíduos em posições de corpo.

No contexto do campo de Relações Públicas e Comunicação Social, a pesquisa está diretamente associada às práticas discursivas da organização OpenAI e a construção de sentido. E, para análise, o conceito de comunicação no contexto organizacional, segundo Mattos (2008), se define como:

um campo de estudos quanto a um conjunto de preferências empíricas, sendo o primeiro considerado subdisciplina ou subárea da Comunicação e o segundo, práticas de comunicação complexas e variadas envolvidas no âmbito das organizações, sejam elas públicas, privadas ou do terceiro setor (Mattos, 2008, p. 18).

A partir deste conceito, a comunicação organizacional compreende as práticas das organizações, que são complexas e precisam estar alinhadas às narrativas comunicacionais. Além disso, Oliveira e Amaral (2021), na publicação sobre Discurso organizacional: presença de poder e sentidos na busca de legitimidade. Desta forma, a discussão em torno do tema da IA é preciso ir além dos apelos mercadológicos para se pensar impactos, sistemas generativos e uso de dados nas organizações.

Afinal, em uma sociedade encantada pela Inteligência Artificial, a IA não está limitada às plataformas de inteligência generativa, como as lançadas pela OpenAI, e tem sido utilizada em diversos sistemas há décadas (Terra, Saad, 2024). Assim, o impacto desses sistemas na comunicação corporativa é significativo e requer uma reflexão constante.

Percurso Metodológico

Para entender de que forma o racismo praticado pela DALL-E se presentifica nas soluções algorítmicas da OpenAI e na discriminação sistêmica da mulher negra (Bento, 2022), este resumo apresenta a análise de uma das postagens que gerou repercussão em torno do caso de Renata Silva, conforme imagem (1):

Imagem 1 - Publicação do post



Fonte: (Souza 2023)

Segundo Bardin (1977), a análise de conteúdo é um conjunto de técnicas para estudo das comunicações, com procedimentos sistemáticos e objetivos para descrever o conteúdo das mensagens. Essa técnica envolve a coleta de dados e análise de indicadores de engajamento, como curtidas e comentários, permitindo a interpretação dos resultados para identificar padrões significativos de discurso.

Diante disso, questiona por que uma mulher negra, originária da favela, é associada à violência, evidenciando como a "desinteligência artificial" reproduz estereótipos racistas. A publicação foi feita no X no dia 26 de outubro de 2023. Desde o período estabelecido para a coleta das informações, passado 1 ano, já são somados 8.3M de visualizações, 5 mil comentários, 22 mil likes e mais 4 mil compartilhamentos (Souza, 2023). Para a análise, foi estabelecido um critério temporal de coleta de dados e optamos por considerar uma base de 172 comentários em ordem cronológica, para capturar a espontaneidade dos primeiros comentários feitos.

Os resultados preliminares deste trabalho revelam alguns aspectos: a prevalência de comentários negativos e a recorrência da dúvida sobre o que foi relatado pela deputada. Isso indica que a deputada atrai para o seu círculo pessoas que colocam em dúvida sua credibilidade e abordam a questão do racismo algorítmico de maneira politizada.

Na perspectiva das práticas discursivas na Comunicação Organizacional e

Relações Públicas, o caso vivenciado pela Renata Souza se apresenta como uma divergência da narrativa da Open IA sobre seu propósito. Uma vez que, a organização afirma ter como missão, garantir que a inteligência artificial geral beneficie a sociedade.

Conforme afirma Vieira (2023), essa situação levanta importantes questões sobre os vieses preconceituosos perpetuados pela IA acerca dos corpos negros, periféricos e sobretudo das mulheres negras⁷. Deste modo é que identificamos casos de racismo e preconceito nas imagens geradas pelo DALL-E, apresentando a hipótese de que os métodos de Aprendizado de Máquina, alimentados pela inteligência humana, são uma das causas desses problemas. Essas estruturas, que envolvem o funcionamento dos algoritmos, revelam as lógicas discriminatórias de empresas e dos indivíduos, que os desenvolvem e os alimentam por meio dos bancos de dados.

Referências

ANGELONI, Maria Terezinha. Comunicação nas organizações da era do conhecimento. São Paulo: Atlas, 2010.

BALDISSERA, Rudimar. Comunicação organizacional na perspectiva da complexidade. *Organicom*, São Paulo, v. 6, n. 10-11, 2009. Disponível em: <<https://revistas.usp.br/organicom/article/view/139013>>. Acesso em: 15 nov. 2024.

BARDIN, Laurence. *Análise de Conteúdo*. São Paulo: Edições 70, 2016.

DEPUTADA acusa robô que faz desenhos ao estilo Pixar de racismo após tecnologia gerar imagem de mulher negra armada. São Paulo: G1, 26 out. 2023. Disponível em: <<https://g1.globo.com/tecnologia/noticia/2023/10/26/deputada-do-rj-diz-que-robo-que-faz-desenhos-ao-estilo-pixar-entregou-imagem-de-mulher-negra-com-arma-na-mao-ao-receber-pedido-de-personagem-em-favela.ghml>>. Acesso em: 15 nov. 2024.

DOSSIÊ: o dilema da inteligência artificial: algoritmos não têm a liberdade de ser morais ou imorais, somente o ser humano. *Revista Cult*, São Paulo, n. 297, 2023. *E-book*. Disponível em: <https://plataforma.bvirtual.com.br>. Acesso em: 1 jul. 2024.

GOMES, Amanda Maria de Sobral. Sexismo e racismo algorítmico: um olhar interseccional sobre o neomaterialismo através do Midjourney. *Revista da UFMG*, v. 30, dez. 2023. Disponível em: <<https://periodicos.ufmg.br/index.php/revistadaufmg/issue/view/2122>>. Acesso em: 29 maio 2024.

⁷ Disponível em:

<<https://g1.globo.com/tecnologia/noticia/2023/10/26/deputada-do-rj-diz-que-robo-que-faz-desenhos-ao-estilo-pixar-entregou-imagem-de-mulher-negra-com-arma-na-mao-ao-receber-pedido-de-personagem-em-favela.ghml>>. Acesso em: 25 nov. 2024.

MATTOS, Maria Ângela. Interfaces do saber comunicacional e da comunicação organizacional com outras áreas de conhecimento. *In*: OLIVEIRA, Ivone de Lourdes; SOARES, Ana Thereza Nogueira (org.) Interfaces e tendências da comunicação no contexto das organizações. São Caetano do Sul, SP: Difusão Editora, 2008. p.17-37.

MÜLLER, Léo. Tay: Twitter conseguiu corromper a IA da Microsoft em menos de 24 horas. [S. l.]: TecMundo, 24 mar. 2016. Disponível em: <<https://www.tecmundo.com.br/inteligencia-artificial/102782-tay-twitter-conseguiu-corromper-ia-microsoft-24-horas.htm>>. Acesso em: 2 jul. 2024.

OLIVEIRA, Ivone de Lourdes. Objetos de estudo da Comunicação Organizacional e das Relações Públicas: um quadro conceitual. *Organicom: Revista Brasileira de Comunicação Organizacional e Relações Públicas*, São Paulo, v. 6, n. 10-11, p. 57-63, 2009. Disponível em: <<https://revistas.usp.br/organicom/article/view/139004>>. Acesso em: 18 dez. 2024.

SILVA, Tarcízio. Racismo algorítmico: inteligência artificial e discriminação nas redes digitais. São Paulo: SESC, 21 fev. 2022.

SILVA, Tarcízio. Racismo algorítmico em plataformas digitais. *In*: SILVA, Tarcízio (org.). Comunidades, algoritmos e ativismos digitais: olhares afrodiaspóricos. São Paulo: Editora LiteraRUA, 2021a. p 127-145. Disponível em: <https://www.researchgate.net/publication/339954112_Comunidades_Algoritmos_e_Ativismos_Digitais_olhares_afrodiasporicos>. Acesso em: 30 jun. 2023.

SILVA, Tarcízio. Racismo algorítmico em plataformas digitais: microagressões e discriminação em código. *In*: SILVA, Tarcízio (org.). Comunidades, algoritmos e ativismos digitais: olhares afrodiaspóricos. São Paulo: Editora LiteraRUA, 2021b. p 127-145. Disponível em: <https://www.researchgate.net/publication/339954112_Comunidades_Algoritmos_e_Ativismos_Digitais_olhares_afrodiasporicos>. Acesso em: 30 jun. 2023.

VIEIRA, Carla. Como os vieses entram na IA generativa? *In*: Projeto Desvelar – justiça racial na inteligência artificial e TICs. Disponível em: <<https://desvelar.org/2023/10/25/como-vieses-entram-na-ia-generativa/>>. Acesso em: 15 mar. 2025.