

UNIVERSIDADE FEDERAL DO AMAZONAS - UFAM
FACULDADE DE ESTUDOS SOCIAIS - FES
CURSO DE ADMINISTRAÇÃO

RUTE OLIVEIRA DE ALMEIDA

**INFLUENCIADORES NEURODIVERGENTES E O DISCURSO DE ÓDIO NO
INSTAGRAM**

MANAUS

2025

RUTE OLIVEIRA DE ALMEIDA

**INFLUENCIADORES NEURODIVERGENTES E O DISCURSO DE ÓDIO NO
INSTAGRAM**

Trabalho de Conclusão de Curso apresentado ao Curso de Administração da Universidade Federal do Amazonas (UFAM), como requisito para obtenção do título de Bacharel.

Orientador: Prof(a). MsC. ANA CLÁUDIA DE ARAÚJO MOXOTÓ

MANAUS

2025

RUTE OLIVEIRA DE ALMEIDA

**INFLUENCIADORES NEURODIVERGENTES E O DISCURSO DE ÓDIO NO
INSTAGRAM**

Trabalho de Conclusão de Curso apresentado ao Curso de Administração da Universidade Federal do Amazonas (UFAM), como requisito para obtenção do título de Bacharel.

Este trabalho foi defendido e aprovado pela banca em 11/12/2025.

BANCA EXAMINADORA

Prof.^a MsC.^a Ana Cláudia de Araújo Moxotó - UFAM
Orientadora

Prof.^a Dr. Edjard de Souza Mota - UFAM
Orientador

Prof. Dr. Mauricio Brilhante de Mendonça - UFAM
Avaliador

Prof. Dr. Manoel Carlos de Oliveira Júnior - UFAM
Avaliador

INFLUENCIADORES NEURODIVERGENTES E O DISCURSO DE ÓDIO NO INSTAGRAM

Rute Oliveira de Almeida
rute.almeida@ufam.edu.br

Ana Cláudia de Araújo Moxotó
anaclaudiamoxoto@ufam.edu.br

Edjard de Souza Mota
edjard@icomp.ufam.edu.br

Resumo

A visibilidade digital, embora promova a inclusão, pode expor minorias a vulnerabilidades intensificadas. Este estudo, de perspectiva quali-quantitativa, investiga o fenômeno do "Paradoxo da Visibilidade" ao analisar comparativamente o discurso de ódio direcionado a influenciadores neurodivergentes e neurotípicos no Instagram. O corpus compreende 2.000 comentários coletados nos perfis de Greta Thunberg e Thiago Ávila após o episódio de intercepção humanitária em 08 de junho de 2025. Utilizando uma abordagem mista com Análise de Sentimentos via modelo XLM-RoBERTa e a *Zero-Shot Classification (ZSC)*, os resultados evidenciaram uma assimetria significativa: Greta (neurodivergente) recebeu 88,4% de comentários negativos, contra 73,4% de Thiago (neurotípico). Além disso, a influenciadora foi alvo de 473 insultos pessoais e 465 ocorrências de linguagem obscena, números expressivamente superiores aos 262 e 281 registrados, respectivamente, para o influenciador neurotípico. A análise lexical revelou que os ataques a Greta focam na desumanização, enquanto as críticas a Thiago centram-se na ideologia política. Conclui-se que a neurodivergência atua como um catalisador de violência simbólica e capacitismo, demandando políticas de moderação de conteúdo mais sensíveis à diversidade neurocognitiva.

Palavras-chave: Neurodiversidade; Capacitismo Digital; Discurso de Ódio; Análise de Sentimentos; Paradoxo da Visibilidade.

Abstract

Digital visibility can simultaneously promote inclusion and intensify vulnerabilities for minority groups. This mixed-methods study examines the "Visibility Paradox" by comparing hate speech directed at neurodivergent and neurotypical influencers on Instagram. Based on 2,000 comments collected from the profiles of Greta Thunberg and Thiago Ávila after the humanitarian interception episode on June 8, 2025, sentiment analysis using XLM-RoBERTa and Zero-Shot Classification reveals a significant asymmetry: the neurodivergent influencer received a higher proportion of negative comments and more frequent personal and obscene attacks. Lexical analysis shows that these attacks emphasize dehumanization, whereas criticism of the neurotypical influencer focuses mainly on political ideology. The findings indicate that neurodivergence amplifies exposure to symbolic violence and digital ableism, highlighting the need for more inclusive content moderation policies.

Keywords: Neurodiversity; Digital Ableism; Hate Speech; Sentiment Analysis; Visibility Paradox.

1. INTRODUÇÃO

Nas últimas décadas, profundas transformações sociais e econômicas têm sido impulsionadas pelo avanço das tecnologias digitais e pela centralidade da internet nas interações humanas. Esse ambiente conectado ampliou fluxos comunicacionais e inaugurou novas formas de atuação profissional, especialmente no campo do marketing digital. Esse campo abrange práticas estratégicas realizadas no ambiente online para engajar públicos, enquanto envolve ações publicitárias integralmente apoiadas em plataformas digitais (Dimitrios et al., 2023; Al Adwan et al., 2023). O marketing digital transformou a lógica de promoção e consumo, reforçando o papel das tecnologias na mediação das relações sociais contemporâneas (Desai, 2019). Essa forma de atuação digital tem se mostrado como uma nova oportunidade profissional, e de representatividade, principalmente para pessoas historicamente marginalizadas, como neurodivergentes. O mercado tradicional continua sendo excludente para essa população, devido à rigidez de processos seletivos, a falta de políticas de inclusão efetivas e a persistência de preconceitos, apesar da ampliação do debate a respeito da neurodiversidade e do fortalecimento do ativismo neurodivergente (Leopoldino, 2015).

Muitos neurodivergentes encontram na atuação como influenciadores digitais uma alternativa para superar as barreiras empregatícias. O mercado tradicional de trabalho apresenta desafios para a maioria dos candidatos; contudo, quando se trata de pessoas neurodivergentes, as dificuldades se tornam ainda maiores, com uma taxa de desemprego que chega a ser três vezes superior à de pessoas com deficiências físicas e seis vezes maior do que a de indivíduos neurotípicos (Branton et al., 2023). Por outro lado, o ambiente digital também expõe a novas formas de vulnerabilidade, especialmente na discriminação e no discurso de ódio. Cria-se, um "paradoxo da visibilidade": o mesmo elemento essencial para o sucesso do influenciador é o que atrai violência simbólica.

Este estudo visa compreender se a visibilidade digital expõe pessoas neurodivergentes a uma carga de toxicidade e hostilidade superior à enfrentada por seus pares neurotípicos, evidenciando as barreiras da inclusão nas plataformas digitais. Para tanto, investiga as manifestações de discurso de ódio e capacitismo digital direcionadas a influenciadores neurodivergentes, por meio de análise comparativa entre os perfis de Greta Thunberg (neurodivergente) e Thiago Ávila (neurotípico) no Instagram, durante o episódio de interceptação humanitária de junho de 2025.

A relevância deste estudo é tripla. Socialmente, busca compreender se o Instagram atua como espaço de inclusão ou se reproduz a exclusão via discurso de ódio. Academicamente, contribui para a intersecção entre estudos de neurodiversidade e comunicação digital. Politicamente, investiga os limites da inclusão online e o papel das plataformas no combate ao *hate speech*. Espera-se, por fim, oferecer subsídios para estratégias comunicacionais sensíveis à diversidade e fomentar políticas de moderação que valorizem a pluralidade cognitiva nas mídias digitais.

2. REVISÃO BIBLIOGRÁFICA

2.1. Marketing, Visibilidade e a Neurodiversidade

As novas ferramentas do marketing digital permitem a personalização, interação e a otimização da experiência do cliente através de estratégias coordenadas envolvidas no marketing digital (Nuseir et al., 2023). Os autores afirmam que essa abordagem permite o monitoramento em tempo real do comportamento dos consumidores e possibilita a melhoria da eficácia das campanhas. O desenvolvimento do marketing digital está diretamente relacionado ao avanço das tecnologias e à crescente utilização de dispositivos digitais no dia a dia dos consumidores (Desai, 2019). Neste sentido, Al Adwan et al. (2023) afirmam que as plataformas digitais, como Facebook e Instagram são ambientes de uma constante interação, tornando o engajamento e a experiência do usuário fatores-chave para o sucesso das marcas dentro destas plataformas. O marketing digital evoluiu de forma, para não ser apenas uma exposição online, mas sim uma complexa estratégia de construção de relacionamento e personalização massiva (Desai, 2019), transformando a forma como as empresas e consumidores interagem (Dimitrios et al., 2023).

Influenciadores digitais são “microcelebridades” que acumulam seguidores ao monetizar narrativas cotidianas via publicidade, (Abidin, 2016), categorizados por alcance e especialidade (ex: ativistas, *lifestylers*), destacando seu papel na criação de conteúdo autoral para promover ideias e produtos (Wielki, 2020). Ao projetarem uma imagem de acessibilidade e confiança, esses influenciadores garantem alto engajamento e eficácia na transmissão de mensagens. A forma de trabalho dos influenciadores, para Wielki (2020), é através da criação e compartilhamento de conteúdo próprio nas redes sociais, que podem incluir opiniões, dicas, experiências e análises relacionadas a produtos, serviços ou temas de interesse do seu público, o que, para Abidin (2016), se trata da “visibilidade laboral” com o objetivo de fortalecer sua marca pessoal para atrair parcerias comerciais.

O marketing tem se tornado um vetor de inclusão para a neurodiversidade, termo cunhado por Judy Singer (1999) para equiparar variações neurológicas (ex: autismo, TDAH). Fundamentado no ativismo autista, o conceito rejeita a patologização, propondo que tais divergências sejam vistas não como doenças, mas como identidades a serem valorizadas. Essa perspectiva é politicamente significativa, pois desafia estruturas dominantes de "normalidade" e produtividade, reivindicando reconhecimento social (Branton et al., 2023).

Kenny et al. (2016) realizaram um estudo com mais de 3.400 participantes e observaram que a grande maioria das pessoas autistas preferem o termo “autista” ao invés de “pessoa com autismo”, por reconhecerem essa característica como parte integral de sua identidade, o que reflete o esforço de comunidades neurodivergentes em reivindicar autonomia linguística e social. Esse entendimento é reforçado ao se afirmar que o uso do termo neurodiversidade não busca evitar o rótulo de deficiência, mas garantir que o indivíduo neurodivergente tenha oportunidades de realização pessoal (Mousa et al., 2024).

As marcas têm realizado parcerias com influenciadores digitais em razão do crescimento exponencial de sua popularidade, o que possibilita alcançar públicos engajados, uma vez que, em parte, essas parcerias tendem a não ser percebidas como publicidade, além de apresentarem risco e custo relativamente menores (Ferraro et al., 2024). Essa forma de trabalho tem sido considerada como uma aspiração profissional, em que muitas pessoas confiam na promessa de crescimento do capital social e econômico (Duffy, 2016).

No contexto neurodivergente, injustiças epistêmicas, estigma e marginalização social afetam as possibilidades de reconhecimento e valorização das diversas formas de prosperidade e realização (Chapman & Carel, 2022). Pessoas dentro do Transtorno do Espectro Autista (TEA) relatam dificuldades tanto para conseguir quanto para manter um emprego, além dos desafios de encontrar posições que correspondam à sua formação e ambições, para muitos autistas, a omissão do diagnóstico torna-se a única forma de acessar oportunidades de trabalho (Leopoldino, 2015).

Atualmente, observa-se uma maior representatividade online no movimento de neurodiversidade e no ativismo autista no Brasil, uma vez que a presença digital permite que pessoas com Transtorno do Espectro Autista (TEA) expressem suas perspectivas, reivindiquem direitos e construam redes de apoio, contribuindo para a visibilidade e o reconhecimento social da neurodiversidade (da Nóbrega et al., 2024).

No contexto das minorias o paradoxo da visibilidade adquire contornos de segurança e direitos humanos. Para grupos minoritários, sair da invisibilidade é crucial para a reivindicação de direitos e formação de comunidade. No entanto, como sugere Barnhurst

(2007) ao analisar a mídia LGBTQ+, essa exposição resulta frequentemente em uma "hipervisibilidade" que atrai violência simbólica e física. Essa questão é aprofundada ao se demonstrar como os algoritmos perpetuam um “paradoxo da representação”, no qual minorias são tornadas visíveis de maneiras estereotipadas e prejudiciais, reforçando desigualdades estruturais sob a fachada de inclusão digital (Noble, 2018). Configura-se a 'nova visibilidade' mediada pela tecnologia, onde o controle da própria imagem escapa das mãos do indivíduo, tornando-o vulnerável a reações imprevisíveis do público (Thompson, 2005). Assim, a literatura contemporânea demonstra que a visibilidade nas mídias digitais opera como uma arma de dois gumes nas lutas identitárias de grupos minoritários.

2.2. Discurso de Ódio e Capacitismo no Ambiente Digital

O discurso de ódio envolve a vilificação de indivíduos por características como etnia, religião ou gênero e não apenas reforça a intolerância, mas também compromete a dignidade das pessoas e limita sua participação igualitária no espaço público, ao silenciar ou marginalizar aqueles que são alvo desse tipo de manifestação (Lang, 2023). No contexto das redes sociais, o discurso de ódio pode se manifestar de diversas formas, desde práticas sutis que utilizam ironia ou humor para disfarçar conteúdos racistas e discriminatórios, como *memes*, *GIFs* (Graphics Interchange Format) e emojis, até discursos abertamente ofensivos (Matamoros-Fernández & Farkas, 2021).

O discurso de ódio adaptou-se ao ambiente digital por meio de estratégias linguísticas e simbólicas que dificultam sua detecção, inclusive por sistemas automatizados. Com o avanço de modelos baseados em transformadores, como *BERT* e *GPT*, tornou-se possível capturar nuances mais complexas da linguagem; ainda assim, o *hate speech* continua a se manifestar de forma ambígua, exigindo análises mais profundas e sensíveis ao contexto (Ramos et al., 2024).

O capacitismo, por sua vez, pode ser compreendido como uma vertente do discurso de ódio, este acontece por meio de crenças, práticas e processos que discriminam pessoas com deficiências ou diferenças neurológicas (Nóbrega et al., 2024). No contexto brasileiro, o capacitismo é manifesto por meio da recusa em reconhecer as diversidades neurológicas e na permanência de preconceitos que privam a inclusão plena destes indivíduos em espaços sociais. Normas sociais e culturais reforçam a visão do que é considerado “normal” e excluem aqueles que dela divergem, contribuindo para processos de desumanização e desrespeito às diferenças (Legault et al., 2021).

Apesar do crescimento da presença do ativismo neurodivergente nas redes sociais, o capacitismo e preconceitos continuam presentes, demonstrando, desta forma, como as redes sociais refletem o que acontece no mundo fora delas (Nóbrega et al., 2024). Observa-se uma recorrente exposição a comentários ofensivos e capacitistas, sem que as plataformas adotem mecanismos eficazes de moderação ou proteção individualizada, o que evidencia falhas nas políticas de conteúdo e na forma como o discurso de ódio direcionado a pessoas com deficiência é negligenciado (Heung et al., 2025). Obermaier et al. (2023) evidenciam que muitas vezes o público assume uma postura de indiferença ou omissão frente a discursos de ódio, especialmente quando esses ataques se dirigem a grupos menos visibilizados, como os neurodivergentes. Essa passividade contribui para a manutenção de ambientes hostis e reforça práticas discriminatórias já enraizadas na sociedade.

2.3. Violência Online: Ameaças, Insulto Pessoal, Críticas Agressivas no Ambiente Digital

A violência online, também conhecida como cyberviolência, tem se intensificado com o crescimento das redes sociais e da exposição digital, uma vez que esse tipo de agressão transcende fronteiras físicas e temporais e afeta profundamente o bem-estar emocional (Cowie & Myers, 2023). Vítimas de violência digital frequentemente apresentam sintomas como depressão, raiva e retraimento social (Espino et al., 2025). A ausência de uma definição consensual sobre o *cyberbullying* dificulta a formulação de políticas públicas eficazes (Ray et al., 2024). A falta de clareza conceitual compromete a capacidade de intervenção e prevenção em ambientes digitais cada vez mais complexos, o que reforça a necessidade de abordagens mais integradas.

A violência digital pode se manifestar de diversas formas, tais como ameaças, críticas agressivas e insulto pessoal e o tema tem sido objeto de estudo em diversas áreas, especialmente diante do crescimento das interações sociais mediadas por tecnologia. A agressão/violência online é resultado de uma interação entre intenções individuais, oportunidades tecnológicas e ausência de controle social eficaz, o que exige ações coordenadas entre plataformas digitais, educadores e legisladores (Shaikh et al., 2025). A compreensão desses elementos é essencial para o desenvolvimento de intervenções eficazes que reduzam os danos psicossociais da violência digital (Tennakoon et al., 2024).

Dentre as formas de violências no ambiente digital destaca-se o insulto pessoal online, que é uma forma de agressão verbal que ocorre em ambientes digitais, geralmente

direcionada à aparência, identidade, crenças ou posição social de um indivíduo. Dreißigacker et al. (2024) destacam que vítimas deste discurso ofensivo digital relatam sentimento de insegurança e vulnerabilidade. Os autores afirmam que a experiência com este tipo de discurso de ódio online está associada a uma percepção ampliada de ameaça, especialmente entre grupos minoritários e profissionais expostos publicamente. A experiência de assédio online é marcada por danos psicológicos, sociais e até físicos, dependendo da intensidade e da frequência dos ataques (Schoenebeck et al., 2023). Além disso, a exposição repetida a ameaças e mensagens ofensivas pode comprometer o desenvolvimento emocional e social dos jovens (Obermaier et al., 2023).

A propagação dos insultos pessoais e ameaças online é favorecida por fatores como anonimato, desinibição social e polarização ideológica, que tornam os ambientes digitais propícios à agressividade verbal. Conforme apontado, a persistência do discurso hostil online afeta negativamente a qualidade das interações e o bem-estar dos usuários, especialmente quando não há mecanismos eficazes de moderação ou responsabilização. Esses elementos contribuem para a normalização de comportamentos ofensivos, dificultando a identificação e o enfrentamento dos danos causados por esse tipo de violência simbólica (Garland et al., 2022).

De acordo com Oliveira *et al.* (2023) a crítica agressiva nas redes sociais pode ser considerada uma forma *deciberagressão*. Conteúdos agressivos, mesmo quando disfarçados de opinião, geram repercussões negativas imensuráveis para os envolvidos, afetando tanto o bem-estar quanto a reputação das vítimas. A crítica agressiva online frequentemente ultrapassa os limites do debate saudável e se configura como uma forma de assédio, especialmente quando envolve repetição, exposição pública e intenção de humilhar, o que exige maior atenção por parte das plataformas e dos usuários (Schoenebeck et al., 2023). Além disso, a maioria das críticas agressivas é direcionada a mulheres, minorias e figuras públicas, além de funcionar como uma ferramenta de silenciamento utilizada para intimidar e excluir vozes dissidentes, revelando seu papel estrutural na manutenção de desigualdades (Jane, 2020).

Baseado na revisão da literatura propõe-se a seguinte hipótese de pesquisa:

H1: Influenciadores digitais neurodivergentes são mais suscetíveis a receber comentários com conteúdo de ódio do que influenciadores neurotípicos.

Reconhecemos que a experiência de hostilidade online possui uma natureza interseccional, sendo o ódio frequentemente amplificado e moldado pela combinação de múltiplas identidades, como neurodiversidade, gênero e ativismo. O fato de uma influenciadora neurodivergente ser mulher, por exemplo, introduz um componente significativo de misoginia online, funcionando como um multiplicador de risco que direciona ataques mais sexualizados e pessoalizados. No entanto, a complexidade inerente à análise rigorosa de todas essas variáveis em conjunto demandaria um escopo de pesquisa extenso, o que inviabilizaria os objetivos centrais deste trabalho. Dessa forma, e visando garantir a viabilidade metodológica e a focalização da pesquisa, a análise se restringirá estritamente à verificação da Hipótese H1, que compara a suscetibilidade ao discurso de ódio apenas em função da neurodiversidade (comparação entre influenciadores neurodivergentes versus neurotípicos). A variável gênero, embora reconhecida como crucial para a intensidade do fenômeno, será tratada como uma limitação do estudo, abrindo caminho para futuras pesquisas que possam incorporá-la como fator moderador ou multiplicador do risco.

3. METODOLOGIA

3.1 Contexto da Pesquisa: O conflito em Gaza

A intensificação do conflito entre Israel e Gaza após outubro de 2023 resultou em uma crise humanitária severa, contabilizando mais de 55.000 mortos e o colapso da infraestrutura civil até junho de 2025. Além da destruição física, observam-se níveis críticos de insegurança alimentar, sugerindo o uso da fome como tática de guerra em violação ao direito humanitário (Van Steenberghe, 2024; Di Maio & Sciabolazza, 2021). Nesse cenário, a mídia e as redes sociais atuam como campos de disputa narrativa, embora Falah (2021) e Wolfsfeld (2018) alertem para a parcialidade da cobertura e o uso político da informação.

É neste contexto de tensão e disputa discursiva que se insere o episódio envolvendo os ativistas Greta Thunberg e Thiago Ávila. Em 08 de junho de 2025, durante uma missão humanitária destinada a levar auxílio à Faixa de Gaza, ambos foram interceptados e detidos pela Marinha Israelense. O evento ganhou repercussão imediata após os influenciadores publicarem *reels* no Instagram relatando o ocorrido. A visibilidade desse episódio serve como base para este estudo, permitindo comparar como a audiência reagiu à narrativa de uma ativista neurodivergente (Greta) em contraste com um ativista neurotípico (Thiago) diante do mesmo fato político e humanitário. A proposta desta pesquisa é observar se a condição

neurodivergente influencia as percepções e relações sociais expressas nos comentários, especialmente em situações de alta visibilidade pública e polarização política.

3.2 Caracterização da pesquisa

Esta pesquisa adota a perspectiva quali-quantitativa, por permitir analisar os sentidos simbólicos e as construções sociais presentes nas falas digitais. Segundo Minayo (1992) essa linha de investigação possibilita compreender como os sujeitos atribuem significado às experiências, levando em conta os contextos em que estão inseridos. A pesquisa qualitativa oferece caminhos para investigar fenômenos que envolvem relações, valores e tensões sociais (Denzin & Lincoln, 2006).

Trata-se de um estudo de natureza aplicada, uma vez que pretende contribuir para reflexões concretas sobre o capacitismo digital e seus desdobramentos no cotidiano virtual, pois visa gerar conhecimento com utilidade imediata, sendo fortemente orientada para a prática e a transformação social (Tisdell et al., 2025), além de contribuir para a compreensão e enfrentamento do discurso de ódio em ambientes digitais (Marconi & Lakatos, 2019). Gil (2008) destaca que esse tipo de pesquisa é voltado à solução de problemas específicos, conectando teoria e prática, através da ferramenta análise de sentimentos, que se trata da mineração de informações relacionadas a sentimentos ou opiniões de um grupo sobre determinado tema (Saad & Saberi, 2017).

3.3 Coleta de dados

Os dados foram reunidos a partir de comentários públicos realizados em publicações do dia 08 de junho de 2025 no Instagram, em que os influenciadores ativistas Greta Thunberg, que à época contava com aproximadamente 16 milhões de seguidores, e Thiago Ávila, com cerca de 1,2 milhão de seguidores, relatam terem sido sequestrados por forças marítimas israelenses. A escolha por esse recorte temporal busca concentrar a análise nos momentos de maior visibilidade pública. Foram coletados manualmente por meio de cópia direta mais de 12 mil comentários de ambos os influenciadores, respeitando os princípios éticos da pesquisa e sem identificação dos autores. Para o estudo foram incluídos apenas comentários que mencionaram diretamente o episódio de sequestro, e foram excluídas interações irrelevantes, como *memes*, emojis soltos ou assuntos alheios. Após os procedimentos de limpeza e filtragem, foram selecionados os 1.000 primeiros comentários

resultantes desse processo, contemplando manifestações tanto positivas quanto negativas, os quais constituíram o corpus final de análise. A seleção dos influenciadores se justifica pela intenção de comparar a recepção pública direcionada a uma figura neurodivergente (Greta) e outra neurotípica (Thiago), em um mesmo contexto de exposição midiática e ativismo político.

3.4 Análise dos Dados

A metodologia usada para análise de dados do estudo é Análise de Sentimentos. A Análise de Sentimentos (AS) é uma área de pesquisa em constante desenvolvimento no campo da mineração de texto/dados. Esta metodologia consiste no tratamento computacional de opiniões, sentimentos e subjetividade presentes em textos (Medhat et al., 2014). A análise de sentimento é processo de coletar e analisar as opiniões, pensamentos e impressões das pessoas sobre diversos tópicos, produtos, assuntos e serviços (Wankhade et al., 2022). A análise de sentimentos e a mineração de opiniões em redes sociais tornam-se cada vez mais relevantes em plataformas como o Instagram, uma das principais comunidades digitais onde usuários geram diariamente um volume substancial de opiniões sobre os mais variados temas. No entanto, esse ambiente também se caracteriza pela disseminação de comentários depreciativos e prejudiciais, evidenciando a necessidade de moderação e análise crítica desse conteúdo (Eyvazi-Abdoljabbar et al., 2024).

A análise de sentimentos foi conduzida no ambiente Google Colab, utilizando a linguagem Python e a biblioteca *Transformers* da *Hugging Face*. O modelo selecionado para a tarefa de *Zero-Shot Classification* (ZSC) foi o XLM-RoBERTa Large XNLI (*joeddav/xlm-roberta-large-xnli*). Esta arquitetura é uma variação multilíngue do RoBERTa (*RobustlyOptimized BERT Approach*), um modelo desenvolvido pelo Facebook AI que otimiza o desempenho do BERT tradicional (Liu et al., 2019; Conneau et al., 2020).

A escolha do XLM-RoBERTa fundamenta-se em sua robustez para processar *corpora* mistos e em sua superioridade de desempenho: o modelo atinge 78,37% de acurácia em análises multilíngues, superando métodos clássicos como SVM (64,2%) e *XGBoost* (62%) (Prytula, 2024). No presente estudo, utilizou-se o modelo em tarefas de *Natural LanguageInference* (NLI), técnica que avalia a probabilidade de uma premissa (o comentário) corresponder a uma hipótese (a categoria de toxicidade), gerando um *score* de confiança entre 0 e 1.

A técnica de *Zero-Shot Classification (ZSC)* foi adotada conforme a linha de trabalhos de zero-shot learning em modelos pré-treinados (Xian et al., 2019; Conneau et al., 2018), permitindo classificar os comentários em categorias de toxicidade sem necessidade de um conjunto de treino rotulado especificamente para esta pesquisa.

Bardin (2016), aborda a importância das etapas de pré-análise, exploração do material e tratamento dos resultados. Essa abordagem mostra-se especialmente adequada para textos curtos, como comentários em redes sociais, por permitir a análise e mapeamento de sentimentos, padrões discursivos e manifestações de preconceito. Como ressalta Flick (2008), a análise de conteúdo possibilita apreender tanto o que é dito explicitamente quanto o que se manifesta de forma indireta, aspecto fundamental para identificar discursos de ódio disfarçados ou normalizados. Portanto, antes da etapa de análise, o *corpus* do texto passou por um processo de limpeza automática implementado no *script*, que incluiu a remoção de marcas, caracteres especiais e numeração, a unificação de espaços, a exclusão de linhas vazias ou com menos de cinco caracteres e a padronização do texto para reduzir ruídos. Além disso, foram carregadas *stopwords* multilíngues, a fim de gerar análises mais limpas e concentradas nos termos mais relevantes.

Combinando a análise de sentimentos com a ZSC, foram aplicados três procedimentos principais de forma articulada. Em primeiro lugar, realizou-se a análise da polaridade dos sentimentos, classificando os comentários como positivos ou negativos, a fim de identificar padrões de aceitação ou rejeição em relação a Greta Thunberg e Thiago Ávila. Essa técnica é amplamente utilizada em estudos de mídias sociais, pois permite captar percepções coletivas e mapear a recepção discursiva em contextos de alta visibilidade (Mohammad, 2016).

Em seguida, procedeu-se à análise da frequência das categorias negativas detectadas, com a contabilização de insultos pessoais, linguagem obscena, discurso de ódio, crítica agressiva, ameaça/violência e toxicidade geral, o que possibilitou comparar quantitativamente a intensidade dos ataques entre os dois perfis e identificar quais tipos de ataques foram mais frequentes contra cada influenciador, oferecendo uma visão comparativa clara (Mollas et al., 2022).

Por fim, foi realizada uma análise lexical, destacando os termos mais recorrentes e evidenciando como determinadas expressões funcionam como marcadores de violência simbólica. Essa técnica, como destaca Bardin (2016), facilita a identificação de vocábulos recorrentes e do tom predominante dos ataques, permitindo observar como certas expressões se tornam marcadores de hostilidade e mostrando que a análise lexical é fundamental para

identificar regularidades e singularidades no uso das palavras, revelando sentidos sociais implícitos.

A aplicação da metodologia ZSC em conjunto com essas técnicas representa um avanço significativo na investigação de discursos de ódio em ambientes digitais. Como ressaltam Anjum e Katarya (2024), a flexibilidade dos modelos *zero-shot* permite sua aplicação direta em novos domínios, sem necessidade de reconfiguração. Esses modelos são capazes de identificar múltiplas camadas de discurso ofensivo, mesmo sem treinamento específico (Mollas et al., 2022). O aprendizado zero-shot oferece alternativas promissoras para superar os desafios de generalização em tarefas de detecção de ódio (Yin & Zubiaga, 2021). Dessa forma, a análise comparativa entre os influenciadores neurodivergente e neurotípico pôde ser conduzida de maneira robusta, revelando padrões distintos de vulnerabilidade discursiva em contextos de alta exposição pública.

4 RESULTADOS E DISCUSSÕES

Nesta seção são apresentados e discutidos os resultados da pesquisa, organizados em torno da análise comparativa de sentimentos, que constitui o eixo central deste estudo. A comparação entre os comentários direcionados a Greta Thunberg (neurodivergente) e Thiago Ávila (neurotípico) permite compreender como diferentes perfis de influenciadores são percebidos em contextos de alta visibilidade digital.

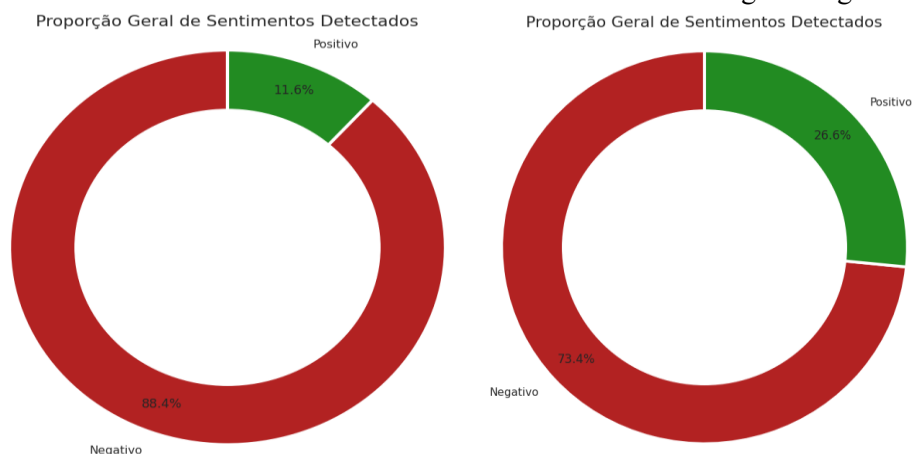
4.1 Análise Comparativa de Sentimentos

A análise comparativa de sentimentos foi aplicada em todo o *corpus* de comentários, constituindo a base para os diferentes procedimentos analíticos apresentados a seguir. Essa abordagem permitiu identificar não apenas a polaridade dos comentários (positivos e negativos), mas também a frequência das categorias negativas e os padrões lexicais associados à linguagem obscena.

4.1.1 Análise da Polaridade dos sentimentos

Na Figura 1, o gráfico à esquerda, apresentam-se os resultados referentes à Greta Thunberg; à direita, os dados de Thiago Ávila. Esta primeira etapa da análise consistiu na identificação da polaridade dos comentários, classificados como positivos ou negativos. A análise de sentimentos, conforme definido por Mohammad (2016), oferece uma abordagem sistemática para quantificar emoções e opiniões expressas em textos, permitindo a síntese do panorama discursivo em torno dos *reels* publicados em 08 de junho de 2025.

Figura 1. Polaridade dos Sentimentos Detectados - Greta Thunberg e Thiago Ávila



Fonte: Dados da Pesquisa

Os resultados apresentados na Figura 1 revelam uma diferença significativa na polaridade da recepção dos dois influenciadores. A Greta Thunberg recebeu 88,4% de comentários negativos e apenas 11,6% positivos. Já o Thiago Ávila apresentou uma distribuição menos hostil, com 73,4% de comentários negativos e 26,6% positivos. Essa diferença de mais de 15 pontos percentuais na taxa de polaridade negativa sugere uma hostilidade desproporcional direcionada à figura de Greta, mesmo em um contexto idêntico de atuação humanitária.

Do ponto de vista teórico, essa disparidade pode ser interpretada à luz dos estudos sobre estigma social e neurodivergência. Segundo Goffman (1988), indivíduos que carregam atributos considerados desviantes pela sociedade, como a neurodivergência, tendem a ser alvo de processos de desvalorização simbólica e exclusão discursiva. Greta, enquanto pessoa autista e ativista global, representa uma figura que desafia normas sociais e cognitivas estabelecidas, o que pode intensificar reações negativas em ambientes digitais, marcados pela polarização e anonimato.

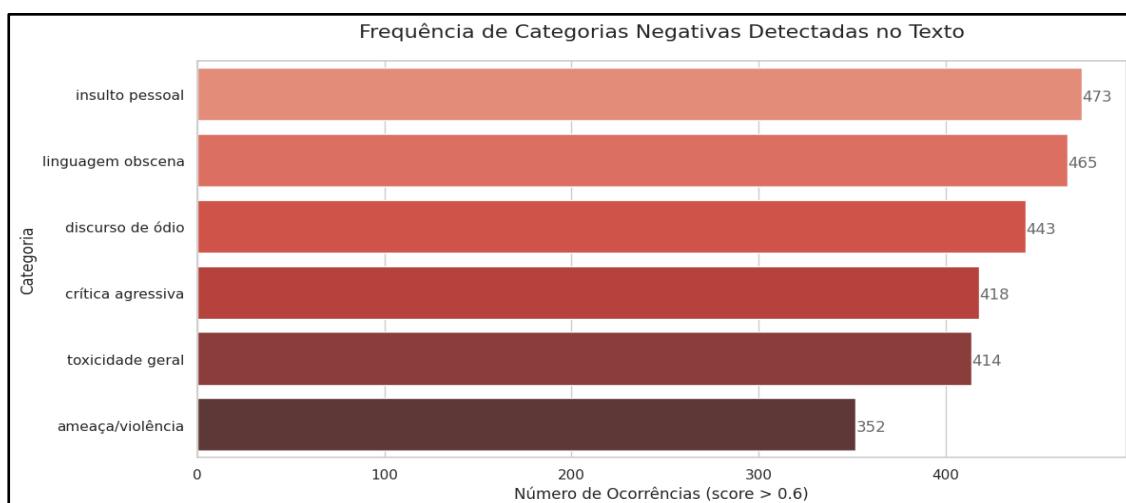
Além disso, autores como Walker e Raymaker (2021) defendem que a neurodivergência não se limita a uma diferença neurológica, mas constitui também uma

identidade política e cultural sistematicamente marginalizada. A predominância de sentimentos negativos nos comentários direcionados a Greta pode refletir não apenas discordância ideológica, mas também preconceito estrutural contra pessoas neurodivergentes, especialmente quando ocupam espaços de visibilidade e poder. Por outro lado, Thiago Ávila, embora também ativista e envolvido na mesma ação, não carrega o marcador da neurodivergência em sua identidade pública, o que pode ter contribuído para uma recepção menos hostil e mais equilibrada. Essa diferença reforça a hipótese central deste trabalho: personalidades neurodivergentes são mais suscetíveis ao discurso de ódio e à rejeição emocional nas redes sociais, mesmo quando atuam em contextos semelhantes aos de seus pares neurotípicos.

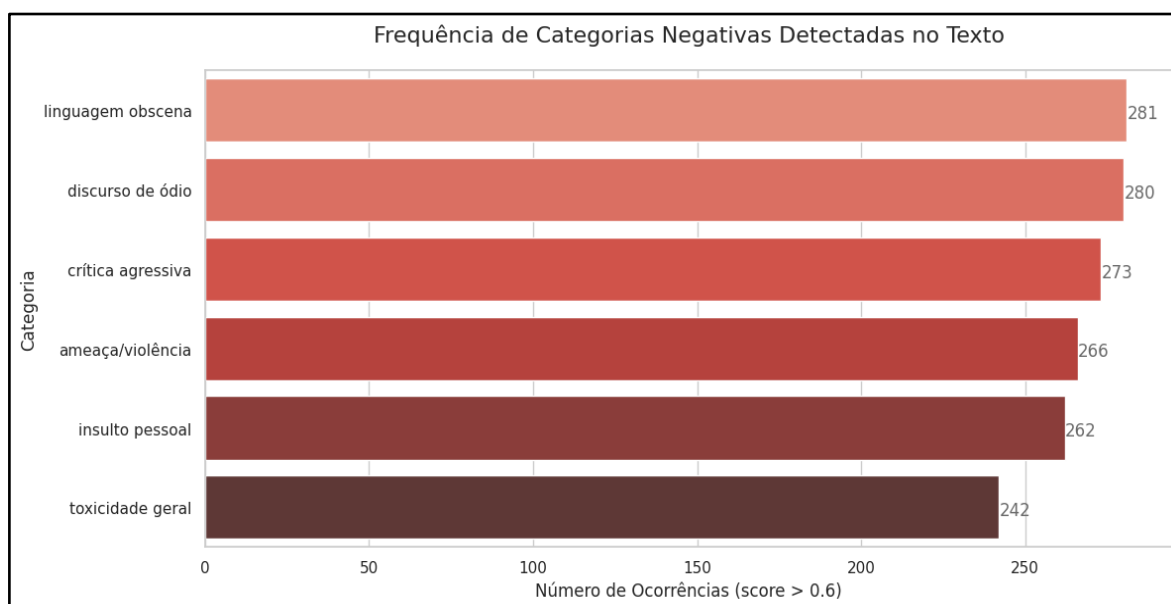
4.1.2 Análise da Frequência das Categorias Negativas Detectadas

A segunda etapa, Análise da Frequência das Categorias Negativas Detectadas, consistiu em contabilizar a frequência das categorias de polaridade negativa de toxicidade identificadas nos comentários. Esse procedimento é essencial para compreender não apenas a presença de discursos hostis, mas também sua intensidade e distribuição entre os dois influenciadores, conforme representado nas Figuras 2 e 3 a seguir:

Figura 2. Frequência de Categorias Negativas Detectadas – Greta Thunberg



Fonte: Dados da pesquisa

Figura 3. Análise das Categorias Negativas Detectadas – Thiago Ávila

Fonte: Dados da Pesquisa

Para esta análise, foram consideradas as ocorrências com score superior a 0,6, distribuídas entre seis categorias de toxicidade: insulto pessoal, linguagem obscena, discurso de ódio, crítica agressiva, ameaça/violência e toxicidade geral. A Figura 2 apresenta os resultados obtidos nos comentários direcionados a Greta Thunberg, enquanto a Figura 3 mostra os resultados referentes a Thiago Ávila, permitindo uma comparação direta entre os dois perfis.

A contagem das ocorrências permite verificar quais formas de toxicidade foram mais recorrentes e como se diferenciaram entre Greta Thunberg e Thiago Ávila. Greta Thunberg recebeu 473 ocorrências de insulto pessoal, enquanto Thiago Ávila teve 262, uma diferença de 211 casos. A linguagem obscena também foi significativamente mais presente nos comentários contra Greta (465 ocorrências) do que contra Thiago (281), indicando um padrão de agressividade verbal mais intenso. O discurso de ódio aparece com 443 ocorrências para Greta e 280 para Thiago, reforçando a disparidade. Já a categoria “ameaça/violência” teve 352 registros no perfil de Greta, contra 266 em Thiago. No caso das críticas agressivas, Greta também apresentou maior número de ocorrências (418) em comparação a Thiago Ávila, embora também tenha recebido críticas agressivas (273 ocorrências). Esses dados revelam que os ataques contra Greta não apenas foram mais frequentes, mas também mais direcionados à sua identidade pessoal, com maior presença de insultos e linguagem obscena. A toxicidade nos comentários de Thiago, embora presente se manifesta de forma mais branda e genérica. Essa diferença sugere que Greta, enquanto figura pública neurodivergente, é alvo

de um tipo de toxicidade mais personalizada e violenta, o que pode ser interpretado como reflexo de processos de desumanização discursiva. A Tabela 1 apresenta o sumário das diferenças entre as categorias analisadas.

Tabela 1: Comparação por categoria de Toxicidade

Categoria de Toxicidade	Greta Thunberg (Neurodivergente)	Thiago Ávila (Neurotípico)	Diferença (Volume de Ataques)
Insulto Pessoal	473	262	211
Linguagem Obscena	465	281	184
Discurso de Ódio	443	280	163
Crítica Agressiva	418	273	145
Ameaça/Violência	352	266	86
Total (Categorias analisadas)	2.151	1.362	789

Fonte: Dados da pesquisa

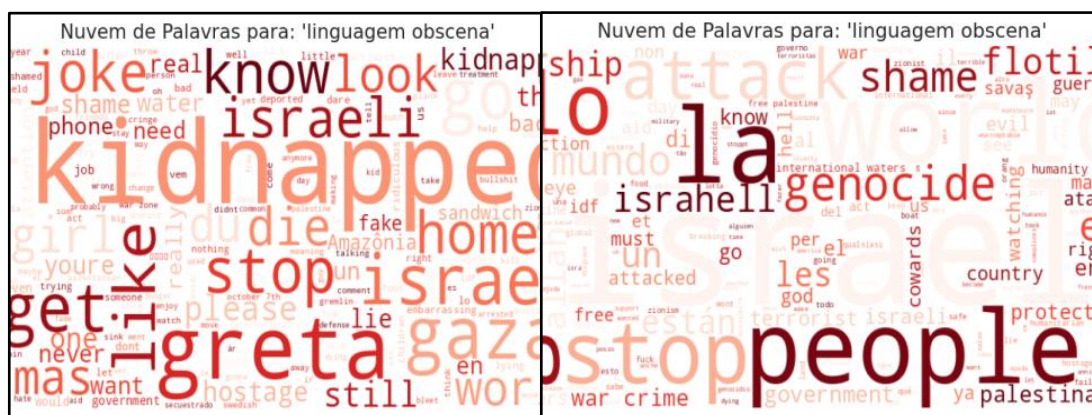
Do ponto de vista teórico, essa assimetria pode ser interpretada à luz dos estudos sobre estigma e exclusão discursiva. Indivíduos com identidades socialmente marcadas são mais suscetíveis à deterioração simbólica, sendo tratados como “menos humanos” em espaços públicos (Goffman, 1988). Essa lógica é intensificada nas redes sociais, onde o anonimato e a polarização favorecem a propagação de discursos violentos. Butler (2021) argumenta que certos corpos são considerados mais "choráveis" que outros, ou seja, mais dignos de empatia e proteção. A intensidade dos ataques contra Greta pode indicar que sua condição neurodivergente a posiciona fora do espectro de reconhecimento social pleno.

Além disso, Ahmed (2013) discute como figuras que desafiam normas sociais, são frequentemente alvos de rejeição emocional e discursiva. A presença de insultos e obscenidades nos comentários contra Greta reforça essa lógica, evidenciando um padrão de exclusão que ultrapassa o debate político e se inscreve na esfera da identidade. Por outro lado, Thiago Ávila, embora também envolvido em uma ação politicamente sensível, não apresenta marcadores identitários que o coloquem em posição de vulnerabilidade simbólica semelhante. Isso pode explicar a menor incidência de ataques pessoais e a predominância de críticas voltadas à ação, e não ao indivíduo.

4.1.4 Análise Lexical da Nuvem de Palavras

A quarta etapa da análise consistiu na utilização da nuvem de palavras para identificar os termos mais recorrentes associados à violência digital, linguagem obscena, discurso de ódio nos comentários.

Figura 4. Análise Lexical para: Greta Thunberg e Thiago Ávilla



Fonte: Dados da Pesquisa

Essa abordagem permite visualizar com mais clareza os vocábulos que se destacam em volume e frequência, revelando padrões lexicais que expressam formas específicas de hostilidade. A nuvem de palavras da Figura 4 evidenciou que os ataques contra Greta Thunberg se concentraram em expressões de ridicularização e desumanização, enquanto os contra Thiago Ávila estavam mais relacionados ao contexto político e ideológico.

Na nuvem de palavras referente à Greta Thunberg, ao lado esquerdo, observam-se termos como “*joke*”, “*fake*”, “*lie*” e “*die*”, que indicam um padrão de ridicularização e desumanização. Palavras como “*joke*” e “*fake*” sugerem que a figura pública de Greta é percebida como falsa ou manipuladora, enquanto “*lie*” reforça a ideia de descredibilização. O termo “*die*”, por sua vez, ultrapassa o campo da crítica e adentra o território da violência simbólica explícita. Esses vocábulos revelam que os ataques contra Greta não se limitaram a discordâncias ideológicas, mas assumiram caráter pessoal, com foco na desqualificação de sua identidade e atuação pública. Em contraste, a nuvem de palavras de Thiago Ávila apresenta vocabulário predominantemente político e contextual, com destaque para termos como “*gaza*”, “*genocide*”, “*israhell*”, “*war*”, “*crime*”, “*massacre*”, “*terrorist*”, “*cowards*” e “*media*”. Esses termos indicam que a linguagem obscena nos comentários direcionados a Thiago está fortemente vinculada ao debate geopolítico e à sua atuação como ativista. Embora haja expressões ofensivas, como “*cowards*” e “*terrorist*”, elas não se dirigem

diretamente à sua pessoa, mas sim ao conteúdo político que ele representa. Isso sugere que os ataques a Thiago, quando presentes, são mais ideológicos e menos personalizados.

Essa diferença lexical é significativa porque mostra que a negatividade direcionada a Greta assumiu caráter mais pessoal e ofensivo, enquanto em Thiago os ataques se voltaram para sua atuação política. A análise lexical é fundamental para identificar regularidades e singularidades no uso das palavras, revelando sentidos sociais implícitos (Bardin, 2016).

Do ponto de vista metodológico, a nuvem de palavras complementa as análises de sentimentos e toxicidade ao oferecer uma dimensão qualitativa dos ataques. Enquanto os gráficos anteriores mostraram polaridade e categorias, a análise lexical revelou o conteúdo concreto das expressões utilizadas. Essa etapa é essencial para compreender como os discursos de ódio se materializam linguisticamente, permitindo identificar não apenas a intensidade dos ataques, mas também os significados sociais que eles carregam. A presença de termos de ridicularização em Greta confirma que a neurodivergência pode ser alvo de formas específicas de violência simbólica, que se diferenciam das críticas políticas recebidas por Thiago.

Os resultados empíricos apontam que hostilidade não é acidental, mas estrutural, corroborando os estudos de Haimson et al. (2021), que, embora focados em identidades de gênero, oferecem um paralelo crucial para compreender a experiência neurodivergente. Os autores demonstram que usuários que desafiam as normas de comportamento "típico", seja por expressão de gênero ou, por analogia, por comunicação neurodivergente, sofrem punições desproporcionais. Isso ocorre tanto pela via algorítmica (moderação de conteúdo enviesada) (Noble, 2018), quanto pela via social (ataques de usuários que policiam a "anormalidade") (Legault et al., 2021).

5 CONCLUSÃO

Este estudo teve como objetivo analisar comparativamente os discursos de ódio direcionados a um influenciador neurodivergente e um influenciador neurotípico no Instagram, buscando compreender se a neurodivergência atua como marcador de vulnerabilidade em contextos de debate público.

Por meio da combinação entre análise de sentimentos, usando a modelagem XLM-RoBERTa Large XNLI, com técnicas de *zero-shot* multilíngue, foi possível examinar sentimentos, padrões de toxicidade e manifestações de discurso ofensivo dirigidos a Greta Thunberg e Thiago Ávila.

Os resultados delinearam uma clara assimetria: Greta Thunberg foi alvo de uma hostilidade personalizada e desumanizadora, com alta incidência de insultos e obscenidades, enquanto Thiago Ávila enfrentou críticas voltadas majoritariamente à sua posição política. A análise lexical confirmou que, para a influenciadora neurodivergente, o ataque foca na identidade; para o neurotípico, na ideologia. Essa dinâmica confirma a existência de um 'Paradoxo da Visibilidade', onde a oportunidade de inclusão digital se converte em exposição ampliada à violência simbólica e ao capacitismo. A pesquisa demonstra que a neurodivergência intensifica a vulnerabilidade a discursos de ódio, ultrapassando os limites da crítica racional. Conclui-se, assim, ser imperativo o desenvolvimento de políticas de moderação e educação digital que protejam a pluralidade cognitiva e combatam a normalização do ódio contra minorias neurodivergentes.

Os resultados empíricos corroboram a hipótese da pesquisa de que a influenciadora neurodivergente enfrenta maior vulnerabilidade discursiva em ambientes digitais, sendo alvo de formas específicas de hostilidade que ultrapassam o campo da crítica racional. A pesquisa também contribui para o debate sobre capacitismo digital, ao demonstrar que a neurodivergência não apenas afeta a percepção pública, mas também intensifica a exposição a discursos de ódio, mesmo em contextos de empatia e solidariedade. Nesse sentido, o estudo reforça a necessidade de políticas de moderação mais sensíveis à diversidade neurocognitiva, bem como de estratégias educativas que promovam respeito e inclusão nas plataformas digitais.

Como limitações desta pesquisa, destacam-se o recorte temporal e temático, restrito a um único episódio e a dois influenciadores específicos. É fundamental reconhecer que, embora o foco da análise tenha sido a neurodivergência, outros marcadores sociais interseccionais podem ter influenciado a disparidade nos ataques. O viés de gênero e as diferenças de nacionalidade (Norte Global versus Sul Global) constituem variáveis que não puderam ser totalmente isoladas, podendo a hostilidade contra Greta Thunberg ter sido exacerbada também por componentes misóginos, além dos capacitistas. Sugere-se que estudos futuros ampliem o escopo para diferentes contextos e perfis, controlando variáveis como gênero e etnia para uma compreensão mais precisa dessas dinâmicas. Além disso, seria relevante investigar os efeitos psicológicos da exposição a esse ódio e os mecanismos de resistência que emergem nesses espaços.

REFERÊNCIAS

- Abidin, C. (2016). Visibilitylabour: Engagingwith influencers' fashion brands and #OOTD advertorialcampaignson Instagram. *Media International Australia*, 161(1), 86–100. <https://doi.org/10.1177/1329878X16665177>
- Ahmed, S. (2013). *The cultural politicsofemotion*. Routledge. <https://doi.org/10.4324/9780203700372>
- Al Adwan, A., Kokash, H., Al Adwan, R., &Khattak, A. (2023). Data analytics in digital marketing for tracking theeffectivenessofcampaignsandinformstrategy. *InternationalJournalof Data and Network Science*, 7(2), 563–574. <https://doi.org/10.5267/j.ijdns.2023.3.015>
- Anjum, &Katarya, R. (2024). Hate speech, toxicitydetection in online social media: A recentsofstateoftheartandopportunities. *InternationalJournalofInformation Security*, 23(1), 577–608. <https://doi.org/10.1007/s10207-023-00755-2>
- Barnhurst, K. G. (2007). Visibility as paradox: Representationandsimultaneouscontrast.
- Bardin, L. (2016). *Análise de conteúdo*. Edições 70.
- Branton, S. E., Villamil, A. M., & Reed, J. L. (2023). Branding neurodiversity: A criticaldiscourseanalysisofcommunicativecapitalismandchange empowerment amongneurodiversityworkforceintermediaries. *JournalofPublicRelationsResearch*, 35(5–6), 357–374. <https://doi.org/10.1080/1062726X.2023.2244619>
- Butler, J. (2021). *Discurso de ódio: Uma política do performativo*. Editora Unesp.
- Chapman, R., &Carel, H. (2022). Neurodiversity, epistemicinjustice, andthegoodhumanlife. *Journalof Social Philosophy*, 53(4), 614–631. <https://doi.org/10.1111/josp.12456>
- Conneau, A., Baeovski, A., Collobert, R., Mohamed, A., & Auli, M. (2020). *Unsupervisedcross-lingual representation learning for speech recognition*. arXiv. <https://doi.org/10.48550/arXiv.2006.13979>
- Conneau, A., Rinott, R., Lample, G., Williams, A., Bowman, S. R., Schwenk, H., &Stoyanov, V. (2018). XNLI: Evaluatingcross-lingual sentencerepresentations. In *Proceedingsofthe 2018 ConferenceonEmpiricalMethods in Natural LanguageProcessing* (pp. 2475–2485). Association for ComputationalLinguistics. <https://doi.org/10.18653/v1/D18-1269>
- Cowie, H., & Myers, C. A. (Eds.). (2023). *Cyberbullying and online harms: Preventionsandinterventionsfromcommunityto campus*. Routledge. <https://doi.org/10.4324/9781003258605>
- da Nóbrega Rogoski, B., Caldas, R. C. S., Souza Guevara, V. L. D., Louzeiro, R., Vasconcelos, L. A., & Flores, E. P. (2024). Autismactivismmovement in Brazil: Contingencyanalysisandthe Pacto pela Neurodiversidade (Pledge for Neurodiversity). *Trends in Psychology*. <https://doi.org/10.1007/s43076-024-00398-9>

Denzin, N. K., & Lincoln, Y. S. (2006). *O planejamento da pesquisa qualitativa: Teorias e abordagens*. Artmed.

Desai, V. (2019). Digital marketing: A review. *International Journal of Trend in Scientific Research and Development*, 5(5), 196–200. <https://doi.org/10.31142/ijtsrd23100>

Di Maio, M., & Leone Sciabolazza, V. (2021). Conflict exposure and health: Evidence from the Gaza Strip. *Health Economics*, 30(9), 2287–2295. <https://doi.org/10.1002/hec.4270>

Dimitrios, B., Ioannis, R., Angelos, N., & Nikolaos, T. (2023). Digital marketing: The case of digital marketing strategies on luxurious hotels. *Procedia Computer Science*, 219, 688–696. <https://doi.org/10.1016/j.procs.2023.01.345>

Dreiβigacker, A., Müller, P., Isenhardt, A., & Schemmel, J. (2024). Online hate speech victimization: Consequences for victims' feelings of insecurity. *Crime Science*, 13(1), 4. <https://doi.org/10.1186/s40163-024-00204-y>

Duffy, B. E. (2016). The romance of work: Gender and aspirational labour in the digital culture industries. *International Journal of Cultural Studies*, 19(4), 441–457. <https://doi.org/10.1177/1367877915572186>

Espino, E., Monks, C. P., Casas, J. A., & Del Rey, R. (2025). The role of bullying victimisation, online activity, and parental supervision on the emotional impact of cyberbullying. *Journal of Child and Family Studies*, 34(3), 706–723. <https://doi.org/10.1007/s10826-024-02982-0>

Eyvazi-Abdoljabbar, S., Kim, S., Feizi-Derakhshi, M. R., Farhadi, Z., & Mohammed, D. A. (2024). An ensemble-based model for sentiment analysis of Persian comments on Instagram using deep learning algorithms. *IEEE Access*.

Falah, G. W. (2021). The (im)possibility of achieving a peaceful solution to the Israeli-Palestinian conflict. *Human Geography*, 14(3), 333–345. <https://doi.org/10.1177/19427786211019014>

Ferraro, C., Sands, S., Zubcevic-Basic, N., & Campbell, C. (2024). Diversity in the digital age: How consumers respond to diverse virtual influencers. *International Journal of Advertising*, 43(8), 1342–1365. <https://doi.org/10.1080/02650487.2023.2300927>

Flick, U. (2008). *Introdução à pesquisa qualitativa* (3ª ed.). Artmed.

Garland, J., Ghazi-Zahedi, K., Young, J. G., Hébert-Dufresne, L., & Galesic, M. (2022). Impact and dynamics of hate and counter speech online. *EPJ Data Science*, 11(1), 3. <https://doi.org/10.1140/epjds/s13688-021-00314-6>

Gil, A. C. (2008). *Métodos e técnicas de pesquisa social* (6ª ed.). Atlas.

Goffman, E. (1988). *Estigma: Notas sobre a manipulação da identidade*. LTC.

Haimson, O. L., Delmonaco, D., Nie, P., & Wegner, A. (2021). Disproportionate removals and differing content moderation experiences for conservative,

transgender, and black social media users: Marginalization and moderation gray areas. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–35. <https://doi.org/10.1145/3476077>

Heung, S., Jiang, L., Azenkot, S., & Vashistha, A. (2025, April). “Ignorance is not bliss”: Designing personalized moderation to address ableist hate on social media. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1–18). ACM. <https://doi.org/10.1145/3613904.3642391>

Jane, E. A. (2020). Online abuse and harassment. In K. Ross & C. Padovani (Eds.), *The international encyclopedia of gender, media, and communication* (pp. 116–122). Wiley. <https://doi.org/10.1002/9781119429128.iegmc052>

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which term should be used to describe autism? Perspectives from the UK autism community. *Autism*, 20(4), 442–462. <https://doi.org/10.1177/1362361315588200>

Lang, G. (2023). Hate speech and the limit of free speech. In M. T. Pérez-Escobar & J. M. Noguera-Vivo (Eds.), *The Routledge handbook of philosophy and media ethics* (pp. 11–25). Routledge. <https://doi.org/10.4324/9781003134749-3>

Legault, M., Bourdon, J.-N., & Poirier, P. (2021). From neurodiversity to neurodivergence: The role of epistemic and cognitive marginalization. *Synthese*, 199(5–6), 12843–12868. <https://doi.org/10.1007/s11229-021-03356-5>

Leopoldino, C. B. (2015). Inclusão de autistas no mercado de trabalho: uma nova questão de pesquisa. *Gestão & Sociedade*, 9(22), 853–868.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). *RoBERTa: A robustly optimized BERT pretraining approach*. arXiv. <https://doi.org/10.48550/arXiv.1907.11692>

Marconi, M. A., & Lakatos, E. M. (2019). *Fundamentos de metodologia científica* (8ª ed.). Atlas.

Matamoros-Fernández, A., & Farkas, J. (2021). Racismo, discurso de ódio e mídias sociais: Uma revisão sistemática e crítica. *Television & New Media*, 22(2), 205–224. <https://doi.org/10.1177/1527476420982232>

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113.

Minayo, M. C. D. S. (1992). *O desafio do conhecimento: Pesquisa qualitativa em saúde*. Hucitec.

Mohammad, S. M. (2016). Sentiment analysis: Detecting valence, emotions, and other affectual states from text. In R. Calvo, S. D’Mello, J. Gratch, & A. Kappas (Eds.), *Emotion measurement* (pp. 201–237). Elsevier. <https://doi.org/10.1016/B978-0-08-100508-8.00009-6>

Mollas, I., Chrysopoulou, Z., Karlos, S., & Tsoumakas, G. (2022). ETHOS: A multi-label hate speech detection dataset. *Complex & Intelligent Systems*, 8(6), 4663–4678. <https://doi.org/10.1007/s40747-021-00608-2>

Mousa, M., Avolio, B., & Pahlevan Sharif, S. (2024). Emancipation and street-level bureaucracy in non-western contexts: The representation of neurodiverse individuals in public sector organizations. *Public Integrity*, 26(4), 439–454. <https://doi.org/10.1080/10999922.2023.2247867>

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.

Nuseir, M. T., El Refae, G. A., Aljumah, A., Alshurideh, M., Urabi, S., & Al Kurdi, B. (2023). Digital marketing strategies and the impact on customer experience: A systematic review. In *The effect of information technology on business and marketing intelligence systems* (pp. 21–44). Springer. https://doi.org/10.1007/978-3-031-12382-5_2

Obermaier, M., Schmid, U. K., & Rieger, D. (2023). Too civil to care? How online hate speech against different social groups affects bystander intervention. *European Journal of Criminology*, 21(3), 817–833. <https://doi.org/10.1177/14773708231156328>

Oliveira, E., Oliveira, L., & Baldi, V. (2023). Analysis of cyberaggression in social networks involving students and university environments. In *International Conference on Design and Digital Communication* (pp. 287–300). Springer. https://doi.org/10.1007/978-3-031-47281-7_23

Prytula, M. (2024). Fine-tuning BERT, DistilBERT, XLM-RoBERTa and Ukr-RoBERTa models for sentiment analysis of Ukrainian language reviews. *Machine Learning*, 3(4).

Ramos, G., Batista, F., Ribeiro, R., Fialho, P., Moro, S., Fonseca, A., Guerra, R., Carvalho, P., Marques, C., & Silva, C. (2024). A comprehensive review on automatic hate speech detection in the age of the transformer. *Social Network Analysis and Mining*, 14(1), Article 204. <https://doi.org/10.1007/s13278-024-01361-3>

Ray, G., McDermott, C. D., & Nicho, M. (2024). Cyberbullying on social media: Definitions, prevalence, and impact challenges. *Journal of Cybersecurity*, 10(1), tyae026. <https://doi.org/10.1093/cybsec/tyae026>

Saad, S., & Saberi, B. (2017). Sentiment analysis or opinion mining: A review. *International Journal on Advanced Science, Engineering and Information Technology*, 7(5), 1660–1666. <https://doi.org/10.18517/ijaseit.7.5.2137>

Schoenebeck, S., Lampe, C., & Triêu, P. (2023). Online harassment: Assessing harms and remedies. *Social Media + Society*, 9(1), 20563051231157297. <https://doi.org/10.1177/20563051231157297>

Shaikh, F. B., Meghji, A. F., Dahri, N. A., Kalhor, M., Ayyasamy, R. K., & Subramaniam, S. (2025). Predicting cyber-harassment in the digital age: An integrated model of psychological, technological, and environmental factors using SEM. *Iran Journal of Computer Science*. Advance online publication. <https://doi.org/10.1007/s42044-025-00347-4>

Singer, J. (1999). Why can't you be normal for once in your life? From a problem with no name to the emergence of a new category of difference. *Disability Discourse*, 59-67.

Tennakoon, H., Betts, L., Chandrakumara, A., Saridakis, G., & Hand, C. (2024). Exploring the effects of personal and situational factors on cyber aggression. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 18(3), 1–29. <https://doi.org/10.5817/cp2024-3-7>

Thompson, J. B. (2005). The new visibility. *Theory, Culture & Society*, 22(6), 31–51.

Tisdell, E. J., Merriam, S. B., & Stuckey-Peyrot, H. L. (2025). *Qualitative research: A guide to design and implementation* (5th ed.). John Wiley & Sons.

Van Steenberghe, R. (2024). The armed conflict in Gaza, and its complexity under international law: Jus ad bellum, jus in bello, and international justice. *Leiden Journal of International Law*, 37(4), 983–1017. <https://doi.org/10.1017/S0922156524000220>

Walker, N., & Raymaker, D. M. (2021). Toward a neuroqueer future: An interview with Nick Walker. *Autism in Adulthood*, 3(1), 5–10. <https://doi.org/10.1089/aut.2020.29014.njw>

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7), 5731–5780.

Wielki, J. (2020). Analysis of the role of digital influencers and their impact on the functioning of the contemporary online promotional system and its sustainable development. *Sustainability*, 12(17), 7138. <https://doi.org/10.3390/su12177138>

Wolfsfeld, G. (2018). The role of the media in violent conflicts in the digital age: Israeli and Palestinian leaders' perceptions. *Media, War & Conflict*, 11(1), 107–124. <https://doi.org/10.1177/1750635217727312>

Xian, Y., Lampert, C. H., Schiele, B., & Akata, Z. (2019). Zero-shot learning—A comprehensive evaluation of the good, the bad and the ugly. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(9), 2251–2265. <https://doi.org/10.1109/TPAMI.2018.2857768>

Yin, W., & Zubiaga, A. (2021). Towards generalisable hate speech detection: A review on obstacles and solutions. *PeerJ Computer Science*, 7, e598. <https://doi.org/10.7717/peerj-cs.598>