

Estado da publicação: O preprint não foi publicado em outro meio.

É nos comentários que a mágica acontece: estratégias visuais e argumentativas da extrema-direita nas redes sociais

Ana Carolina Trevisan Camilo Ferreira, Allan Herison Ferreira

<https://doi.org/10.1590/SciELOPreprints.13311>

Submetido em: 2025-09-10

Postado em: 2025-09-12 (versão 1)

(AAAA-MM-DD)

É NOS COMENTÁRIOS QUE A MÁGICA ACONTECE: ESTRATÉGIAS VISUAIS E ARGUMENTATIVAS DA EXTREMA-DIREITA NAS REDES SOCIAIS

ANA CAROLINA TREVISAN

ORCID: <https://orcid.org/0000-0002-2818-5858>

anacarolinatcf@gmail.com.br

IFILNOVA/Universidade NOVA de Lisboa. Lisboa, Portugal e
LAPS/Universidade de São Paulo, SP, Brasil

ALLAN HERISON FERREIRA

ORCID: <https://orcid.org/0000-0003-3606-2089>

allanherison@gmail.com.br

ICNOVA/Universidade NOVA de Lisboa. Lisboa, Portugal e
LAPS/Universidade de São Paulo, SP, Brasil

RESUMO: Este artigo apresenta um modelo metodológico para analisar o papel dos comentários em postagens multimodais nas redes sociais. Publicações com significado ambíguo e com forte apelo emocional — compostas por imagens ou vídeos sem legenda ou acompanhados de legendas com conteúdo vago — frequentemente dissimulam mensagens que comprometeriam o autor, desinformação, discurso de ódio, conteúdo xenofóbico e ataques a adversários. É, contudo, na seção de comentários que os sentidos implícitos na publicação original emergem, o engajamento se intensifica e os processos de radicalização se consolidam.

A proposta metodológica articula teoria da argumentação, ciência de dados e sociologia digital em um modelo híbrido que integra a análise argumentativa ou Social Argumentative Analysis Framework (SAAF) e o Model Context Protocol (MCP). O MCP oferece uma infraestrutura padronizada e auditável para integrar fluxos de dados heterogêneos, enquanto o método SAAF operacionaliza a detecção de esquemas argumentativos, falácias, linguagem emotiva e info-argumentos, comparando-os com bibliotecas contextuais que reúnem repertórios científicos, religiosos, partidários, ideológicos, entre outros que se demonstrem pertinentes para a análise.

Em sua versão mais ampla deste estudo, o recorte empírico abrange Brasil e Portugal, com foco em publicações realizadas por líderes partidários da extrema-direita nas plataformas X, Instagram e TikTok. O modelo evidencia como conteúdos audiovisuais ambíguos transferem para os comentários a carga interpretativa, espaço onde perfis coordenados ou automatizados semeiam leituras preferenciais reforçadas por comunidades interpretativas e de recepção.

As contribuições concentram-se em quatro pontos: (i) valorização dos comentários em redes sociais como unidade-chave de análise multimodal; (ii) integração entre esquemas argumentativos e repertórios de legitimação (bibliotecas) para comparação e análise; (iii) combinação de metodologias de análise por meio da integração MCP/SAAF; e (iv) introdução do conceito de meta-identidade, que descreve as dinâmicas envolvidas nas interações entre usuários, plataformas digitais e seus processos de classificação de identidades digitais por meio de políticas, parâmetros e algoritmos.

. **Palavras-chave:** Comentários em redes sociais, Esquemas argumentativos, Meta-identidade, Extrema-direita digital, Desinformação e pós-verdade

IN THE COMMENTS IS WHERE THE MAGIC HAPPENS: VISUAL AND ARGUMENTATIVE STRATEGIES OF THE FAR RIGHT ON SOCIAL MEDIA

ABSTRACT: This article presents a methodological model for analyzing the role of comments on multimodal social media posts. Posts with ambiguous meaning and strong emotional appeal — consisting of images or videos without captions or accompanied by vague captions — often conceal messages that could compromise the author, disinformation, hate speech, xenophobic content, and attacks on opponents. It is, however, in the comments section that the implicit meanings of the original post emerge, engagement intensifies, and processes of radicalization consolidate.

The methodological proposal brings together argumentation theory, data science, and digital sociology in a hybrid model that integrates the Social Argumentative Analysis Framework (SAAF) and the Model Context Protocol (MCP). The MCP provides a standardized and auditable infrastructure for integrating heterogeneous data flows, while SAAF operationalizes the detection of argumentation schemes, fallacies, emotive language, and info-arguments, comparing them with contextual libraries that include scientific, religious, partisan, ideological, and other repertoires deemed relevant for analysis.

In its broader version, this study's empirical scope covers Brazil and Portugal, focusing on posts by far-right party leaders on X, Instagram, and TikTok. The model demonstrates how ambiguous audiovisual content shifts the interpretative burden onto comments, a space where coordinated or automated profiles seed preferred readings reinforced by interpretative and reception communities.

The contributions are centered on four key points: (i) valuing social media comments as the core unit of multimodal analysis; (ii) integrating argumentation schemes and repertoires of legitimation (libraries) for comparative analysis; (iii) combining methodologies through MCP/SAAF integration; and (iv) introducing the concept of meta-identity, which describes the dynamics involved in interactions between users, digital platforms, and their processes of classifying digital identities through policies, parameters, and algorithms.

Keywords: Social media comments, Argumentation schemes, Meta-identity, Digital far right, Disinformation and post-truth

INTRODUÇÃO

Publicações com significado ambíguo e com forte apelo emocional — compostas por imagens ou vídeos sem legenda ou acompanhados de legendas com conteúdo vago — tornaram-se um formato recorrente na estratégia informacional da extrema-direita nas plataformas digitais. Nessas peças, conteúdos potencialmente desinformativos, xenófobos, racistas, com discursos de ódio contra minorias e adversários, para citar alguns, são frequentemente dissimulados por meio da ambiguidade visual e textual.

A expressão leve e bem-humorada “é nos comentários que a mágica acontece” que intitula este estudo aparece na fala de influenciadores¹ e em blogs de empresas de marketing digital ao se referir à importância do papel dos comentários, tanto pela riqueza de significados que trazem, quanto pelo seu papel de alavancar o engajamento de publicações nas redes sociais, importância essa que também é reconhecida no contexto corporativo e na comunicação estratégica (KURBANOV, 2025; PENNINGS, 2025, TECH FOCUS PRO, 2025).

É no espaço dos comentários que os sentidos mais explícitos emergem, onde se observam diferentes interpretações, acréscimos ou enquadramentos de sentido, bem como coordenação para reforçar conteúdos preferenciais e assim, ressignificar a publicação em si. Embora não deixe de ocorrer publicações mais explícitas, agressivas ou marcadamente posicionais, esse descolamento entre a superfície intencionalmente “neutra” ou mais amena do post e a camada interpretativa dos

¹ Um exemplo da área do entretenimento é o da influenciadora e atriz Vivian Linhares, que se especializou em reagir com humor aos comentários em vídeos em redes sociais, como Instagram TikTok e YouTube, cujo canal pode ser acessado em: <https://www.youtube.com/c/VivianLinhares>.

comentários é particularmente relevante em ambientes digitais marcados pelos 3Ps — polarização, populismo e pós-verdade —, nos quais repertórios antielites, anticiência, antissistema e retóricas conspiratórias reconfiguram a confiança pública e as formas de validação discursiva (CONRAD; NEWMAN, 2024; HAMELEERS, 2024; HARSIN, 2024; BERGMANN, 2024).

Apesar desse papel decisivo, a literatura tende a privilegiar a análise do conteúdo dos posts (especialmente textos), sub explorando os comentários. Duas razões se destacam: dificuldades técnicas de coleta e organização de camadas de interação em escala e a multimodalidade do material, que exige instrumentos para relacionar implicaturas visuais a cadeias argumentativas textualmente realizadas nos comentários (KJELDSEN, 2012; JEONG, 2008).

Se, por um lado, as plataformas digitais enfrentam pressões por maior transparência e proporcionalidades segundo diretrizes recentes de organismos internacionais (UNESCO, 2021; UNITED NATIONS, 2023), por outro, encontram um contexto de moderação mais flexível desde a posse do presidente norte-americano Donald Trump em 2025. No início deste ano, o governo norte-americano afrouxou as regras de moderação e transparência antes exigidas, blindando as big-techs sediadas nos EUA, mesmo que entrem em dissonância com as leis regionais dos demais países em que atuam, como no caso brasileiro (PRAZERES, 2025). Essas medidas abrem espaço para oportunidades táticas de contornar a fiscalização com mensagens ambíguas e de terceirizar o sentido à seção de comentários, onde ficam mais protegidas e menos sujeitas à fiscalização e às regras estabelecidas pela legislação de cada país.

Partimos, assim, da tese de que os comentários funcionam como camadas interpretativas: é neles que se fixam sentidos, se amplia o engajamento e se performam dinâmicas de radicalização por meio de esquemas argumentativos, falácias e linguagem emotiva (WALTON; REED; MACAGNO, 2008; WALTON, 2010; WALTON; MACAGNO, 2015; 2019). Essa dinâmica é reforçada por padrões de coordenação e automação (*bots* e perfis pouco orgânicos) que induzem leituras específicas e replicam, em rede, enquadramentos transnacionais (ROBLES et al., 2022; WASSERMAN; FAUST, 2012).

O objetivo deste artigo² é apresentar um modelo analítico robusto para endereçar esse fenômeno, integrando análise argumentativa (que denominamos como Social Argumentative Analysis

² Os conceitos e propostas metodológicas desenvolvidos neste estudo foram apresentados pelos autores em quatro diferentes ocasiões: 1) Reconciling perspectives: Approaching far-right argumentation and its communication strategies through visual arguments analysis. Coimbra, 11 out. 2025. 2) Meta-Identity and Digital Platforms: The Influence of

Framework ou SAAF) e ciência de dados por meio do Model Context Protocol, ou MCP (FERREIRA; TREVISAN, 2025a), além de repertórios contextuais e análise de redes. O MCP é mobilizado como infraestrutura de integração, rastreabilidade e governança (logs, permissões, metadados), combinando fontes externas com data marts internos em arranjo híbrido, com salvaguardas para riscos de integração (ANTHROPIC, 2024; HOU et al., 2025; RAY, 2025). O SAAF operacionaliza a identificação de esquemas, falácias e linguagem emotiva, bem como a comparação com repertórios (por exemplo, bibliotecas ou repertórios científicos, jornalísticos, partidários, religiosos, de nichos extremistas, dentre outros) para estimar coerência argumentativa e legitimidade/dissonância (FERREIRA; TREVISAN, 2025a; WALTON; REED; MACAGNO, 2008).

O recorte empírico do estudo mais amplo está orientado a Brasil e Portugal e a três plataformas — X, Instagram e TikTok —, com prioridade para peças audiovisuais de alto engajamento com legendas vagas e temas sensíveis. A perspectiva é transnacional, com foco em alianças e compartilhamento de repertórios políticos entre atores da extrema-direita europeia e latino-americana.

1. REFERENCIAL TEÓRICO

1.1. 3Ps e a ambiguidade estratégica em ambientes digitais

O debate em torno dos 3Ps — polarização, populismo e pós-verdade — consolidou-se como um dos eixos centrais para compreender a transformação do espaço público em sociedades contemporâneas. A polarização corresponde ao processo de radicalização das posições políticas e sociais, reduzindo as possibilidades de dissenso e erodindo o espaço comum de debate (CONRAD; NEWMAN, 2024; ROBLES et al., 2022). Já o populismo atua por meio da simplificação moral de conflitos complexos, estruturando-os em dicotomias como “povo” versus “elite” e reforçando repertórios discursivos centrados na oposição à ciência, às instituições democráticas e ao jornalismo profissional (CASAS-MAS et. al, 2025; HAMELEERS, 2024; HARSIN, 2024).

Algorithmic Decision-Making on Social and Personal Identities. Trabalho apresentado na sessão JS-220.3, V Forum of Sociology, International Sociological Association, Rabat, 11 jul. 2025. Disponível em: <https://isaconf.confex.com/isaconf/forum2025/meetingapp.cgi/Paper/159089>. Acesso em: 20 ago. 2025. 3) Critical Paths for Data Analysis in Social Sciences: Practical Application for Analyzing Data on Polarization, Populism, and Post-Truth (3 P's). Trabalho apresentado na sessão 544.2, V Forum of Sociology, ISA, Rabat, 11 jul. 2025 Disponível em: <https://isaconf.confex.com/isaconf/forum2025/meetingapp.cgi/Paper/158014> (Os textos resultantes estão em fase de análise por pares e, em caso de publicação, terão a devida referência atualizada neste artigo. 4) Argumentative Analysis of Visual and Audiovisual Content in Right-Wing Populist Digital Communication in Brazil and Portugal. Trabalho apresentado na sessão 262.3, V Forum of Sociology, ISA, Rabat, 11 jul. 2025 Disponível em: <https://isaconf.confex.com/isaconf/forum2025/meetingapp.cgi/Paper/160360>

Nesse contexto, a pós-verdade desloca o critério de validação discursiva da evidência para os elementos identitários e emocionais, reforçando a centralidade de apelos afetivos e de narrativas de pertencimento (BERGMANN, 2024). A literatura recente mostra que esses três fenômenos não operam isoladamente, mas de forma interdependente e cumulativa: a erosão do dissenso provocada pela polarização potencializa o populismo, que por sua vez mobiliza repertórios antielite e anti-ciência; ao mesmo tempo, a pós-verdade legitima essas construções ao reconfigurar os regimes de verdade para regimes de pós-verdade (HARSIN, 2015, 2024).

A noção de populismo epistêmico proposta por Harsin (2024) ilustra essa articulação, ao evidenciar como a contestação sistemática das autoridades cognitivas — ciência, jornalismo e instituições políticas — se integra à retórica populista. De modo semelhante, Bergmann (2024) demonstra como a retórica conspiratória é incorporada ao populismo, estruturando um modelo de comunicação política que combina desinformação, apelos emocionais e mobilização coletiva. Esses elementos compõem, portanto, um quadro no qual a ambiguidade estratégica em ambientes digitais deixa de ser um obstáculo incidental e se transforma em recurso central para atores da extrema-direita.

1.2. Teoria da argumentação: esquemas, falácias e info-argumentos

A teoria da argumentação contemporânea oferece instrumentos fundamentais para compreender como discursos ambíguos e comentários em redes sociais digitais estruturam sentidos políticos. O ponto de partida é a concepção de esquemas argumentativos como inferências defensíveis que podem ser testadas por meio de perguntas críticas (WALTON; REED; MACAGNO, 2008). Esses esquemas podem ser classificados em três grandes grupos: argumentos independentes da fonte (epistêmicos), argumentos dependentes da fonte e argumentos de raciocínio prático. No primeiro grupo estão os esquemas de descoberta, como o argumento por sinal, por ignorância ou pela melhor explicação, e aqueles que aplicam regras a casos, incluindo o argumento de causa-efeito, por analogia, por precedente e por exemplo. Já entre os argumentos dependentes da fonte, destacam-se os que se apoiam na autoridade ou posição de saber, como o argumento de opinião especializada ou de testemunho, além dos *ad hominem* (direto ou circunstancial) e dos argumentos baseados na aceitação popular. Por fim, os esquemas de raciocínio prático abrangem tanto os argumentos instrumentais, que conectam ação e motivo, quanto os argumentos de valores, incluindo os baseados em consequências positivas ou negativas, como aqueles que exploram desperdício, ameaça ou custos irreversíveis (WALTON; MACAGNO, 2015, p. 22).

As formas desviantes ou abusivas desses esquemas constituem as falácias, compreendidas não apenas como erros formais de lógica, mas como estratégias pragmáticas de manipulação (WALTON, 2010). Entre as mais recorrentes em contextos digitais estão: o ad hominem, o straw man (distorção da posição do oponente), a omissão ou ignorar qualificações, o raciocínio post hoc, a generalização apressada, a dicotomia falsa, a definição persuasiva e o uso de epítetos como forma de petição de princípio (WALTON; MACAGNO, 2015; MACAGNO; TREVISAN, 2024).

Para além de esquemas e falácias, destaca-se ainda a noção de info-argumentos, entendida como o uso de dados ou “factos” apresentados de forma sugestiva, que insinuam conclusões sem explicitar a cadeia inferencial que as sustenta (MACAGNO, 2025). Essa categoria é particularmente relevante na análise de postagens audiovisuais ambíguas e nos comentários que delas derivam, pois evidencia como a informação pode ser mobilizada de modo fragmentário e estratégico para reforçar determinadas interpretações.

1.3. Linguagem emotiva e repertórios de legitimação

Para Walton e Macagno (2014), as chamadas palavras emotivas não apenas descrevem fatos, mas carregam avaliações implícitas que orientam a interpretação de um enunciado. Termos como “traidor”, “invasor” ou “herói” exemplificam a forma como a carga emocional embutida em uma palavra pode funcionar como premissa não explícita, moldando a recepção e diminuindo a necessidade de justificar a conclusão por meio de argumentos racionais.

A análise da linguagem emotiva deve ser articulada com a noção de repertórios de legitimação, entendidos como “bibliotecas discursivas” utilizadas por diferentes comunidades para interpretar e avaliar a aceitabilidade de um argumento. Esses repertórios podem ser científicos, religiosos, partidários ou midiáticos, funcionando como matrizes de comparação que conferem legitimidade a certos enunciados e rejeitam outros. Na prática, comentários em redes sociais tornam-se um espaço privilegiado para observar como tais repertórios são mobilizados: enquanto alguns usuários legitimam um argumento a partir de referências científicas, outros recorrem a autoridades religiosas ou a discursos antielite para sustentar interpretações divergentes (MCP).

Assim, a comparação de palavras emotivas, falácias e esquemas argumentativos com repertórios de legitimação (bibliotecas formadas por conteúdo científico, religioso, extremista, e assim por diante) fornece as bases para compreender como se estruturam, bem como classificar e analisar

as interpretações coletivas expressas nos comentários que consolidam modos de pertencimento e engajamento e dão sentido às imagens publicadas.

1.4. Meta-identidade (dinâmicas - estados - resultados)

Os comentários em postagens audiovisuais podem revelar interpretações explícitas ou “leituras”, muitas vezes radicais, feitas por diferentes grupos com seus respectivos “jogos de linguagem” (WITTGENSTEIN, 2009) com conteúdo de ódio, racismo, xenofobia e ataque aos direitos humanos. Por estarem numa camada menos visível, e portanto, somente disponível àqueles que buscam mais informações ou que já tinham algum grau de engajamento com o assunto, desempenham um papel decisivo no engajamento e no impulsionamento da publicação principal.

Para além da moderação, interessa às plataformas digitais explorar esses mesmos padrões em benefício próprio, seja de forma direta, ampliando o tempo de permanência e engajamento dos usuários àquelas discussões; seja de forma indireta, ao reforçar cadeias de poder e grupos políticos que convergem com seus interesses estratégicos e econômicos.

As dinâmicas de funcionamento do que chamamos de meta-identidade, ou seja, um modelo de identidade formado pelo referenciamento cruzado de parâmetros, marcadores e classificações realizadas pelas plataformas digitais com base em dados pessoais, conteúdos e comportamentos dos usuários, ajudam a compreender os mecanismos por trás da preferência das redes sociais pelos agentes políticos de extrema-direita e pelo avanço de seus discursos populistas e polarizantes. (FERREIRA; TREVISAN, 2025b).

Na chave do conceito de meta-identidades e suas dinâmicas, podemos compreender o fenômeno identificando cinco das oito³ principais dinâmicas da meta-identidade: i) uma dinâmica performativa, que diz respeito ao modo como publicações que geram emoções reativas e alto nível de engajamento a conteúdos polêmicos são favorecidas pelas plataformas, ii) interacional, que corresponde ao registro das interações ponto-a-ponto entre usuários das plataformas tal como curtidas e os próprios comentários, iii) a executiva, que é definida por atribuições de peso, prioridade ou fluxos de preferência nas plataformas para determinados conteúdos serem mais ou menos expostos e que são definidas por tomadores de decisão das plataformas, iv) a dinâmica normativa que corresponde ao conjunto de leis e regras sociais que impactam as normas internas das plataformas digitais, e v) a

³ As oito dinâmicas da meta-identidade conforme Ferreira e Trevisan (2025b) são: interativa, performativa, executiva, somática, transacional, analítica, executiva, normativa e algorítmica.

dinâmica algorítmica, que integra parâmetros e regras das dinâmicas anteriores para definir o “caminho” de processamento que resultará em cálculos e ponderações que podem definir se uma publicação será mais ou menos potencializada dentro de uma plataforma (FERREIRA; TREVISAN, 2025b).

1.5. Imagens e vídeos como argumentos visuais

A centralidade do audiovisual nas plataformas digitais exige considerar que os argumentos não se limitam à linguagem verbal. Diversos estudos apontam que imagens e vídeos podem sustentar formas próprias de argumentação, ainda que não apresentem proposições textuais explícitas (KJELDSEN, 2012; JEONG, 2008). O poder das imagens reside em sua força evocativa: pela seleção de símbolos, metáforas visuais ou situações representadas, elas sugerem conclusões e orientam interpretações sem necessidade de cadeias inferenciais lineares. Cabe ao público completar as premissas que a imagem sugere, reconstruindo ativamente os seus argumentos (KJELDSEN, 2015).

Nessa perspectiva, imagens e vídeos podem ser compreendidos como argumentos visuais (KJELDSEN, 2015), cuja eficácia depende não apenas de seus elementos internos, mas da forma como são decodificados por comunidades que tecem suas próprias tramas de interpretação.

2. QUESTÃO CENTRAL E HIPÓTESES

Para orientar o desenho metodológico, foram formuladas hipóteses que estão em fase de implementação e teste de escala, mas que permitem a realização do exercício apresentado nesta publicação:

H1: Postagens ambíguas tendem a apresentar, nos comentários, maior densidade de falácias e de linguagem emotiva do que os próprios posts em redes sociais.

H2: Comentários coordenados ou automatizados tendem a induzir leituras específicas e ancoram repertórios característicos, como antielite, anticiência ou anti-imigração.

H3: Os padrões observados nos comentários podem ser utilizados para propor desenhos conceituais de classificação de perfis que eventualmente podem ser traduzidos em meta-identidades analíticas, distinguíveis pela predominância de determinados esquemas argumentativos, emoções recorrentes e repertórios de legitimação.

Com esse desenho, a pesquisa propõe um modelo analítico crítico e escalável, que permita avançar na compreensão do papel dos comentários na construção e circulação de narrativas radicais em ambientes digitais.

3. CORPUS E AMOSTRAGEM

O corpus da pesquisa mais ampla é constituído por postagens audiovisuais produzidas por atores da extrema-direita em plataformas digitais, considerando perfis públicos e, quando aplicável, contas verificadas. A escolha por este recorte justifica-se pela centralidade desses atores na produção e disseminação de conteúdos políticos que operam por meio de estratégias de ambiguidade discursiva e visual, frequentemente projetadas para evitar mecanismos de moderação.

A janela temporal definida contempla os três meses antes e após as principais eleições para o poder executivo, momentos em que os ânimos se acirram e em que as campanhas se intensificam nas redes sociais. Nestas situações, ocorrem com mais frequência tanto a produção de conteúdos com alto apelo emocional quanto os debates polarizados.

Os critérios de inclusão do corpus são: (i) conteúdos de natureza audiovisual (imagens estáticas, vídeos curtos ou montagens multimodais); (ii) legendas minimalistas ou ambíguas, isto é, textos curtos, vagos ou com formulações abertas que não explicitam proposições de forma direta; (iii) alto nível de engajamento, medido por métricas como número de visualizações, comentários, compartilhamentos e curtidas; e (iv) presença de temas sensíveis como imigração, segurança pública, saúde, corrupção, direitos civis ou políticas de identidade. A combinação desses critérios garante a seleção de conteúdos que, embora aparentemente simples ou neutros na superfície, provocam nos comentários processos intensos de interpretação, disputa e polarização.

As unidades de análise são definidas em dois níveis complementares: o post em si, entendido como a combinação entre o elemento audiovisual (imagem ou vídeo) e sua legenda; a camada de comentários, incluindo tanto os comentários diretos quanto as réplicas e threads associadas.

Esse duplo nível é fundamental, pois o objetivo não é analisar apenas a mensagem inicial, mas sobretudo os sentidos explicitados, negociados e radicalizados na esfera dos comentários. Essa abordagem possibilita observar como esquemas argumentativos, falácias e repertórios de legitimação emergem de forma mais explícita na interação entre usuários, em contraste com a ambiguidade estratégica do conteúdo original.

A construção do corpus respeitará princípios de ética e privacidade. Apenas serão coletados conteúdos públicos, disponíveis sem restrição de acesso. Os nomes de usuários comuns (excetuando-se pessoas públicas) são anonimizados de forma sistemática, garantindo que não possam ser identificados diretamente em relatórios ou publicações. Além disso, os dados serão armazenados em repositórios com protocolos de segurança e governança conforme as boas práticas de pesquisa em ciências sociais digitais.

Em síntese, o desenho do corpus e da amostragem busca equilibrar robustez analítica e responsabilidade ética. A escolha de perfis de extrema-direita e de conteúdos audiovisuais ambíguos com alto engajamento assegura a relevância empírica do material a ser analisado, enquanto a inclusão da camada de comentários como unidade autônoma de análise permite explorar o papel decisivo das interações na construção de sentidos, na mobilização emocional e na cristalização de meta-identidades analíticas.

4. PROCEDIMENTOS ANALÍTICOS

4.1. Coleta e pré-processamento

A etapa inicial do processo metodológico consiste na coleta sistemática dos dados provenientes das plataformas digitais selecionadas — X, Instagram e TikTok —, com foco em postagens audiovisuais e respectivas camadas de comentários. A coleta está sendo realizada por meio de técnicas de scraping e, quando disponíveis, por interfaces de programação de aplicações (APIs), assegurando a extração estruturada de conteúdos, metadados e relações de interação.

Cada unidade de dado é acompanhada de um conjunto mínimo de metadados descritivos organizados segundo o modelo 5W2H: quem (autor do post ou comentário, identificado se pessoa pública), quando (data e horário de publicação), onde (plataforma e, quando disponível, localização geográfica), como (formato do conteúdo — imagem, vídeo, texto), porquê (enquadramento declarado ou inferido), quanto (níveis de engajamento como curtidas, compartilhamentos e respostas) e o quê (conteúdo textual ou multimodal efetivamente publicado). Esse padrão garante a comparabilidade entre diferentes plataformas e contextos, além de possibilitar análises longitudinais.

O material coletado passa por um processo de de-duplicação, eliminando conteúdos idênticos ou repetidos de forma automática, o que reduz a redundância e aumenta a qualidade do corpus. Em seguida, é implementado um protocolo de anonimização que suprime ou codifica informações que possam identificar diretamente usuários comuns, assegurando a conformidade com princípios éticos

e de privacidade (ROBLES et al., 2022). Perfis de interesse público — tais como parlamentares, líderes partidários ou influenciadores políticos de grande alcance — serão mantidos identificados, em consonância com sua atuação como agentes públicos ou figuras de relevância política.

Por fim, serão construídas amostras estratificadas de acordo com três dimensões principais: (i) nível de engajamento (alto, médio e baixo, medido por interações agregadas), (ii) tema central (como saúde, imigração, corrupção, segurança pública, entre outros) e (iii) plataforma de circulação. Esse desenho amostral assegura que a análise não se limite a casos de grande repercussão, mas contemple também conteúdos medianos ou marginais que podem desempenhar papel relevante em processos de radicalização discursiva.

4.2. Etapa de codificação

Após a coleta e o pré-processamento, os dados são submetidos a uma etapa de codificação sistemática, cuja finalidade é transformar os conteúdos textuais e multimodais em unidades passíveis de análise comparativa. Para isso, será utilizado um codebook desenvolvido a partir da literatura da teoria da argumentação e dos estudos recentes sobre desinformação e comunicação estratégica (WALTON; REED; MACAGNO, 2008; WALTON; MACAGNO, 2015, FERREIRA; TREVISAN, 2025a).

Esse codebook contempla quatro dimensões principais:

- Esquemas argumentativos (razão prática, consequências, valores, autoridade, analogia, sinal etc.);
- Falácias (como straw man, ignorar qualificações, dicotomia falsa, post hoc, generalização apressada, definição persuasiva e epítetos);
- Info-argumentos, definidos como dados ou “fatos” apresentados de forma sugestiva, que insinuam conclusões sem cadeia inferencial explícita;
- Emotividade, incluindo a presença de palavras emotivas e quasi-definições que reforçam enquadramentos valorativos.

O processo de codificação é realizado inicialmente por meio de codificação humana. Após a etapa manual uma amostra será selecionada para testar um modelo de classificação automática. Na estratégia de aprendizado ativo (active learning), o modelo é progressivamente refinado a partir das anotações humanas. Esse procedimento não substitui a análise crítica dos pesquisadores, mas amplia

a capacidade de escalabilidade do modelo, e sua qualidade interpretativa será testada a partir da comparação com a codificação manual.

4.3. Metodologia para análise de multimodalidade

Considerando a análise de conteúdos digitais, em especial no caso de imagens e vídeos publicados sem legendas ou com legendas vagas ou ambíguas associadas, a abordagem metodológica prevê três procedimentos principais:

- Extração de elementos visuais: coleta, armazenamento e identificação de objetos, personagens, cenários ou situações representadas na imagem ou no vídeo, bem como de quaisquer textos incorporados (legendas sobrepostas, símbolos, slogans).
- Levantamento de possíveis significados correntes de símbolos, expressões populares ou referências presentes na imagem que possam ou não servir como gatilho de ressignificação para as diferentes comunidades e seus repertórios de interpretação.
- Relação com os comentários: os elementos visuais e suas possíveis implicaturas são interpretados à luz das leituras manifestadas nos comentários. Essa triangulação permite avaliar como o público decodifica ou ressignifica as imagens, revelando sentidos que não estavam explícitos no material audiovisual inicial.

O princípio orientador é, portanto, invertido em relação às análises tradicionais: em vez de partir da imagem para compreender os comentários, parte-se dos comentários para obter as diversas interpretações para a imagem, considerando-os como repertórios interpretativos acessíveis que estabilizam a polissemia dos conteúdos visuais. Essa escolha metodológica permite capturar a dimensão interativa da comunicação digital e reforça o papel dos comentários como *locus* privilegiado da radicalização discursiva.

4.4. Sinais de coordenação/automação

Um dos benefícios deste modelo metodológico é a possibilidade de identificar padrões de coordenação e automação em comentários de postagens audiovisuais. Esses padrões podem indicar a presença de *bots* ou de redes de perfis coordenados, cuja função é induzir interpretações específicas, manipular métricas de engajamento e reforçar determinados repertórios discursivos (ROBLES et al., 2022).

A detecção se apoia em um conjunto de heurísticas operacionais:

- Repetição lexical: análise da recorrência de palavras ou frases idênticas, em comentários publicados por diferentes perfis, em curtos intervalos de tempo, sugerindo scripts automatizados. Não raro, frases reproduzidas por diferentes canais trazem uma grafia errada, como um erro de digitação ou erro gramatical que são propagados por diferentes contas num mesmo período.
- Bursts temporais: identificação de picos de atividade concentrados em poucos minutos, com grande volume de comentários semelhantes, frequentemente associados a campanhas organizadas. Um exemplo comum é o de comentários de apoio ou de reforço de significado publicados poucos segundos após a publicação principal.
- Assinaturas temporais: padrões de publicação contínua, em horários pouco plausíveis para atividade humana, indicativos de automatização parcial ou total.

A incorporação desses sinais ao modelo contribui para evidenciar não apenas a presença de perfis inorgânicos, mas também as dinâmicas de orquestração coletiva que potencializam a ambiguidade inicial das postagens, fixando sentidos complementares ou reinterpretações nos comentários.

5. MODELO ANALÍTICO: MCP/SAAF E REPERTÓRIOS CONTEXTUAIS

5.1. Visão geral

Para uma análise completa, incluindo post e comentários, propomos um modelo analítico que combina a metodologia de análise argumentativa com infraestruturas computacionais de dados para lidar com a complexidade dos discursos digitais em ambientes multimodais.

A base da análise dos esquemas argumentativos é o que denominamos Social Argumentative Analysis Framework (SAAF), operado sobre o Model Context Protocol (MCP), que fornece uma arquitetura capaz de integrar dados heterogêneos e, ao mesmo tempo, assegurar rastreabilidade, auditabilidade e governança (ANTHROPIC, 2024; HOU et al., 2025; RAY, 2025).

A lógica do modelo organiza-se em um pipeline modular, composto por camadas sucessivas que equilibram escalabilidade técnica com sensibilidade interpretativa. Em sua primeira etapa, realiza-se a coleta e pré-processamento do material discursivo — postagens audiovisuais e comentários — com registro sistemático de metadados. Em seguida, o material é encaminhado ao MCP, que funciona como uma interface padronizada para ativar simultaneamente diferentes módulos de análise.

O MCP, opera em duas instâncias principais: (i) a argumentativa, dedicada à identificação de esquemas de raciocínio, falácias e usos de linguagem emotiva, utilizando um codebook que servirá como base para a codificação (WALTON; REED; MACAGNO, 2008), e (ii) a de legitimidade, que compara os discursos coletados a repertórios contextuais (bibliotecas) previamente organizados. Dessa forma, cada proposição não é apenas avaliada em termos de coerência lógica, mas também em função do contexto de legitimação em que se insere, seja ele científico, religioso, partidário ou midiático.

A modularidade garante que cada camada do pipeline possa ser adaptada a diferentes objetos empíricos, sem perder a comparabilidade entre contextos. Além disso, a presença de logs auditáveis e a separação entre bases internas (data marts ou data warehouse) e externas (plataformas digitais) preservam os princípios de governança, transparência e segurança recomendados por organismos internacionais (UNESCO, 2021; UNITED NATIONS, 2023).

Com isso, o modelo permite que a análise crítica de interações discursivas — especialmente os comentários em posts audiovisuais ambíguos — seja realizada de maneira escalável, mas sem abdicar do rigor metodológico e da centralidade da interpretação sociológica.

5.2. SAAF (Social Argumentative Analysis Framework)

O Social Argumentative Analysis Framework (SAAF) constitui a dimensão analítica do modelo, permitindo operacionalizar a teoria da argumentação em diálogo com repertórios contextuais organizados em bases de dados. O SAAF estrutura-se em duas instâncias complementares: a argumentativa e a de legitimidade.

Na instância argumentativa, o foco recai sobre a detecção e classificação de esquemas argumentativos presentes em comentários e postagens. São contemplados os esquemas argumentativos conforme sistematizados por Walton, Reed e Macagno (2008). Paralelamente, identifica-se a ocorrência de falácias entendidas como desvios pragmáticos de esquemas válidos (WALTON, 2010; WALTON; MACAGNO, 2015). A esta camada soma-se a análise de linguagem emotiva e de quasi-definições, recursos que ampliam a carga avaliativa de termos e conceitos (WALTON; MACAGNO, 2019). O resultado é a atribuição de um score de coerência argumentativa para cada comentário ou postagem, permitindo comparar padrões de consistência entre diferentes atores e contextos (MACAGNO; TREVISAN, 2024).

Na instância de legitimidade, o SAAF integra a análise dos conteúdos a repertórios contextuais previamente organizados em datamarts temáticos, que podem incluir áreas como saúde, imigração,

segurança, religião ou corrupção (FERREIRA; TREVISAN, 2025a). Esses repertórios ou bibliotecas funcionam como matrizes de comparação, possibilitando a medição da proximidade semântica e ideológica entre o discurso analisado e os referenciais de comunidades distintas — científicas, jornalísticas, partidárias, religiosas ou mesmo nichos extremistas. Além da proximidade, são identificadas assinaturas narrativas recorrentes, tais como enquadramentos antivacina, anti-mídia ou antiglobalistas, que estruturam modos de legitimação.

O produto desta etapa é a atribuição de scores de legitimidade ou dissonância para cada biblioteca, o que permite estimar se um comentário ou postagem encontra respaldo em determinada comunidade interpretativa ou se opera em oposição a ela. A combinação das duas instâncias — coerência argumentativa e legitimidade contextual — possibilita identificar zonas de risco discursivo, em que falácias e linguagem emotiva coincidem com altas taxas de legitimação em bibliotecas de conteúdos polarizados.

Dessa forma, o SAAF constitui uma ferramenta metodológica que não apenas sistematiza a análise de argumentos e falácias, mas também evidencia como esses elementos se articulam com repertórios de legitimação social, permitindo compreender a formação de meta-identidades analíticas em ambientes digitais.

5.3. MCP (Model Context Protocol)

O Model Context Protocol (MCP), até que venha a ser superado por novas opções sistêmicas e tecnológicas, constitui o integrador do modelo proposto, na medida em que oferece uma interface padronizada para conectar diferentes fluxos de dados e agentes de análise. Sua principal função é assegurar a integração controlada de informações, permitindo rastrear de forma transparente o contexto de origem de cada proposição — incluindo autor, plataforma, timestamp e metadados associados —, ao mesmo tempo em que pode ser desenhado para manter um registro auditável de todas as operações (ANTHROPIC, 2024; HOU et al., 2025; RAY, 2025).

A arquitetura do MCP é concebida para lidar com ambientes de dados heterogêneos e distribuídos, permitindo a combinação de fontes externas com repositórios internos. Nessa lógica, adota-se um modelo híbrido: de um lado, data marts ou data warehouses internos, destinados ao armazenamento de dados sensíveis e de alta criticidade; de outro, integrações externas controladas que permitem a conexão com plataformas digitais abertas, desde que mediadas por protocolos de

governança e autorização. Essa configuração garante maior segurança e adaptabilidade, preservando ao mesmo tempo a interoperabilidade necessária para análises comparativas em escala.

Entretanto, a adoção do MCP envolve também riscos específicos que demandam salvaguardas. Entre eles destacam-se: (i) a persistência inadvertida de privilégios após atualizações de sistema; (ii) a possibilidade de escalada de acessos entre ferramentas conectadas; (iii) colisões de nomes em ambientes extensíveis, que podem gerar ambiguidades ou falhas de mapeamento; e (iv) a poluição de contextos persistentes, que compromete a precisão analítica (HOU et al., 2025).

Para mitigar esses riscos, são previstos mecanismos como: (a) isolamento de ambientes de execução por meio de sandboxes; (b) uso de listas brancas (whitelists) para restringir ferramentas autorizadas; (c) rotação periódica de credenciais; (d) manutenção de logs imutáveis de operações; e (e) a presença de human-in-the-loop, garantindo que decisões críticas de categorização e atribuição de legitimidade passem por validação humana. Tais salvaguardas alinham-se às recomendações de organismos internacionais que alertam para a necessidade de equilibrar inovação técnica com governança democrática e transparência (UNESCO, 2021; UNITED NATIONS, 2023).

5.4. Agregação e meta-identidade

A etapa de agregação no modelo proposto busca articular os diferentes níveis de análise em um processo cumulativo, no qual unidades discursivas individuais se transformam em padrões coletivos mais amplos. Esse movimento é essencial para compreender como interações localizadas em posts e comentários contribuem para a formação de meta-identidades analíticas.

No nível micro, a unidade básica de análise é a proposição, expressa em um comentário ou em uma parte de um post. Cada proposição é classificada segundo esquemas argumentativos, falácias, uso de linguagem emotiva e repertórios de legitimação ativados. O objetivo nesse estágio é identificar as estruturas argumentativas elementares que circulam nas interações digitais.

No nível meso, a análise se amplia para o post ou thread como um todo, incluindo o autor e o tópico em questão. Aqui, observa-se como múltiplas proposições se articulam para construir sentidos mais complexos, possibilitando o delineamento de perfis argumentativos. Esses perfis representam combinações recorrentes de esquemas, falácias e emoções em torno de determinados temas ou atores políticos, funcionando como indicadores de consistência discursiva em escala intermediária.

No nível macro, a agregação alcança entidades mais abrangentes — grupos sociais, plataformas digitais ou mesmo contextos nacionais. Esse nível permite observar como repertórios argumentativos são replicados e adaptados em diferentes comunidades e países, favorecendo a identificação de padrões transnacionais de circulação. Nesse estágio, as combinações estáveis de esquemas, falácias e repertórios convertem-se em elementos classificatórios que podem ser utilizados como meta-identidades analíticas, que são formas de classificação baseadas na categorização científica e na interpretação epistêmica (FERREIRA; TREVISAN, 2025b).

As saídas analíticas previstas para esta etapa de publicação não apresentam dados empíricos, mas sim, painéis conceituais que ilustram o potencial do modelo. Entre eles, destacam-se: (i) a identificação de zonas de risco argumentativo, nas quais falácias e linguagem emotiva coincidem com forte legitimação em repertórios polarizados; (ii) séries temporais que cruzam coerência argumentativa com legitimidade contextual, permitindo observar a evolução de narrativas; e (iii) mapas de repertórios por comunidade, que evidenciam como diferentes grupos mobilizam bibliotecas discursivas específicas (científicas, religiosas, partidárias, midiáticas, entre outras).

Ao conjugar esses três níveis, o modelo propõe um caminho para compreender a dinâmica pela qual comentários aparentemente dispersos se consolidam em estruturas coletivas, com disputas por significado e por narrativas. Entende-se, portanto, o post e seus comentários como uma estrutura interligada que se retroalimenta de significado.

6. EXEMPLO DE ANÁLISE DE POST MULTIMODAL

A título de exemplo, analisamos a seguir uma publicação realizada pelo então Presidente Jair Bolsonaro em 30 de setembro de 2022 - três dias antes do primeiro turno das últimas eleições presidenciais brasileiras. A publicação oferece um exemplo da estratégia de ambiguidade calculada por meio do uso de imagens desacompanhadas de legenda para a finalidade de gerar debate nas redes sociais.

O post traz uma foto de Bolsonaro jovem, ao lado de uma onça pintada acorrentada, sem nenhuma legenda, a não ser um link para uma publicação de uma conta dedicada à biodiversidade brasileira. A contagem de comentários chega a cerca de 1 mil, conforme ilustra a Figura 1. No entanto, para este exercício, foram analisados os 240 comentários disponíveis na plataforma, sem contar as respostas a estes comentários. A publicação também contou com 4 mil repostagens, 35 mil likes e 93 salvamentos.

Sua publicação ocorreu logo após um debate realizado pela rede Band de televisão, em que estiveram presentes, além dos principais concorrentes no primeiro turno, os então candidatos Lula da Silva e Jair Bolsonaro, bem como outros candidatos com menor expressividade como Soraya Tronicke do partido União Brasil e o autodenominado Padre Kelmon (apoiador das políticas de Bolsonaro), que protagonizaram um embate direto repleto de acusações e falas incomuns que acabaram por viralizar no noticiário brasileiro e nas redes sociais por meio de memes. Nesse debate, a candidata Soraya Tronicke comparou-se a uma onça, animal típico da fauna brasileira e muito encontrado na região pantaneira do Estado de Mato Grosso, de onde ela é natural. Segundo ela mesma mencionou durante o debate, o Mato Grosso é conhecido por ter “mulheres que viram onça” (PORTINARI, 2022), fazendo alusão ao seu perfil combativo, especialmente dirigido contra as políticas de Jair Bolsonaro, e portanto, contra o candidato Kelmon, a quem chamou de “padre falso”, “padre fake” e “padre de festa junina”, o que depois foi realmente confirmado pela instituição religiosa a que ele se dizia pertencer.



Figura 1. Publicação feita pela conta de @jairbolsonaro na rede social X.

É importante contextualizar que, nesta altura, como se tornou público nos anos seguintes à eleição do Presidente Lula da Silva, Bolsonaro liderava um plano de Golpe de Estado com apoio de

setores das Forças Armadas, ministros, outros agentes políticos e de segurança e de apoiadores concentrados em frente a diversos quartéis do exército, incluindo o de Brasília. Neste período, uma grande quantidade de publicações em diversas redes sociais espalhava notícias falsas, desinformação e ataques às instituições, colocando em dúvida a lisura das urnas e o resultado das eleições que estavam prestes a ser realizadas no Brasil (AGÊNCIA BRASIL, 2025).

Do total de 240 comentários, 139 (58%) tinham como temática o apoio eleitoral a um ou outro candidato, e não expressavam relação com a imagem do post, divididos entre 62% de apoio e 38% de oposição. Tal comportamento justifica-se pela proximidade do primeiro turno das eleições com índices acirrados de competitividade. Do total de comentários, 37 (15%) interpretaram a imagem como uma alusão ao debate, e destes, a maioria (89%) era composta por ataques contra a candidata Soraya. Para além da temática do debate ou das eleições, 24 (10%) comentários revelaram outros níveis de interpretação, mais simbólicos, relacionados à foto do post. Dentre os 4 comentários de oposição nesse sentido, dois interpretaram a imagem como uma tentativa de o Presidente parecer “valentão” ou “machão”, outro como a forma de Bolsonaro de ver as mulheres de forma geral (“animais, seres inferiores”) e como uma tentativa de oprimir o povo (onça como símbolo nacional acorrentada e dominada). Já entre os comentários de apoiadores, 7 referiam-se à capacidade de domar a onça (expressão figurada usada para expressar força, valentia), 5 referiam-se à “hora da onça beber água” (expressão popular que significa momento decisivo, perigoso, originado de um saber indígena sobre os costumes da onça de beber água à noite, em estado de extremo alerta e preparada para o ataque). Outros 2 comentários se referiam ao fato de haver ali na postagem algum sinal, seja pela afirmativa: “é pedir demais para os esquerdistas interpretarem, né, Presidente”, seja pela pergunta: “Presidente, o senhor tá querendo dizer algo?”.

Um comentário digno de nota, é o do ex-assessor de Bolsonaro Filipe Martins, publicado um minuto depois do post de Bolsonaro e que figura em primeiro lugar na lista de comentários, com 3 mil likes e 258 repostagens. O comentário também não traz texto, apenas uma imagem da obra do pintor Benedito Calixto, de 1893, intitulada Evangelho nas Selvas.



Figura 2. Publicação feita pela conta de @filgmartin na rede social X.

Na pintura, um padre jesuíta domina uma onça ao empunhar o crucifixo e a Bíblia em meio à selva brasileira. Uma interpretação muito aceita é a de que o autor trata da colonização dos povos nativos, representados como feras, por meio do poder religioso evangelizador dos jesuítas, em especial, o Padre Anchieta, que marca suas obras (PHILIPPOV, 2023). Apesar da interpretação corrente no mundo erudito e religioso, trata-se de outra imagem que abre uma janela de possibilidades de interpretação para o público daquele comentário, acrescentando símbolos e significados e reforçando a ambiguidade do post principal. Nas respostas a este comentário, uma pessoa o associa ao Padre Anchieta, outra ao Padre Kelson, e outro se indigna pelo fato de o personagem estar “convertendo uma onça”.

6.1. Análise dos comentários

A imagem aparentemente neutra de um líder em sua juventude interagindo com um símbolo nacional (a onça) poderia passar despercebida a quem não estivesse imerso na realidade do país naquele momento. Para os apoiadores dos dois principais candidatos, a imagem acionou significados diferentes, e para quem mergulhou nos comentários, sentidos mais explícitos e outros mais complexos e simbólicos emergiram.

Este caso ilustra o descolamento entre a superfície intencionalmente "neutra" do post e a camada interpretativa desenvolvida pelos usuários, particularmente relevante em ambientes digitais marcados pelos 3Ps - polarização, populismo e pós-verdade. A análise dos 240 comentários coletados demonstra como repertórios antielites, anticiência e retóricas conspiratórias reconfiguram a confiança pública e as formas de validação discursiva, conforme descritos por Harsin (2024) e Bergmann (2024).

Apesar do papel decisivo dos comentários na fixação de sentidos, a literatura tende a privilegiar a análise do conteúdo do post, sub explorando sistematicamente essas camadas de interação. Este estudo tem como objetivo examinar de que modo os comentários funcionam como camadas interpretativas onde se fixam sentidos, se amplia o engajamento e se performam dinâmicas de radicalização por meio de esquemas argumentativos, falácias e linguagem emotiva (WALTON; REED; MACAGNO, 2008; WALTON, 2010).

Com base na análise dos comentários, é possível identificar com clareza as expectativas, anseios e disposições tanto dos apoiadores quanto dos críticos. Esses elementos revelam visões de mundo, valores e projetos políticos profundamente antagônicos, que se acirraram no contexto eleitoral polarizado de 2022.

Entre os apoiadores de Bolsonaro, observa-se uma expectativa pela figura de um líder forte e corajoso, associado a uma imagem paternalista e autoritária, capaz de proteger o país de supostas ameaças internas e externas. Esse imaginário é exemplificado em enunciados como “Enquanto eles pregam o ódio, a Cristofobia, Bolsonaro apenas amansa a onça” e “Nunca foi tão fácil escolher um lado”, que revelam a disposição de admirar gestos performativos de força e controle.

Outro eixo central é a ansiedade pela preservação de valores tradicionais e religiosos, frequentemente vinculados ao catolicismo, à família e à moralidade conservadora. Comentários como “O Brasil nasceu católico. Terra da Santa Cruz” e “O pior (dos inimigos) é o comunismo” mostram o desejo de resistir a transformações progressistas e a reafirmação de uma identidade nacional atrelada a referências religiosas.

O medo do comunismo e da esquerda aparece recorrentemente como ameaça existencial. Afirmações como “Votar nele (Lula) é ensinar aos nossos filhos: que o crime compensa!” e “Eles só pregam ódio” retratam a oposição como corrupta, anti-Brasil e perigosa, reforçando a expectativa de uma vitória eleitoral que afaste definitivamente o retorno do PT ao poder.

Há também o desejo de reconhecimento e valorização, sobretudo em regiões historicamente associadas a outro campo político, como no caso do Nordeste. Comentários como “O Nordeste vai surpreender e vai ajudar na reeleição do capitão” e “governo da igualdade nacional” expressam o anseio de inclusão simbólica em um projeto nacionalista que valorize grupos que se sentiam marginalizados antes de 2018. Além disso, nota-se a necessidade de defesa contra críticas, com respostas defensivas que interpretam qualquer oposição como ataque direto ao grupo, como exemplificado em “Aí é pedir demais para os esquerdistas interpretarem né presidente?”.

No polo oposto, entre os críticos de Bolsonaro, emergem expectativas e anseios bastante distintos. A exigência de coerência e integridade é recorrente, com críticas que denunciam discrepâncias entre discurso e prática, sobretudo em temas ambientais. Comentários como “Em compensação está destruindo a Amazônia, quanta hipocrisia” e “Passar a mão em onça em corrente é fácil, o símbolo nacional da biodiversidade você tentou destruir” apontam para a busca de um governo que pratique o que prega.

A preocupação com o meio ambiente é um dos pontos mais fortes entre os críticos, que veem a gestão Bolsonaro como ameaça direta ao meio ambiente e ao futuro do planeta. Frases como “Onças do Pantanal mandam lembranças” e “Você há de pagar toda a perversidade causada à Amazônia” revelam um horizonte de expectativas pautado pela urgência climática e pela demanda de políticas sérias de preservação.

A rejeição à performatividade de coragem, vista como bravata, também é destacada. “Típico de covarde. Só pega na onça se ela tiver amarrada” e “Foto antiga, de uma onça acorrentada e provavelmente drogada... pra bancar o machão” ilustram o desprezo pela encenação de força em contextos controlados, contraposta à expectativa de lideranças genuinamente corajosas.

O desejo de justiça e responsabilização emerge em críticas à impunidade de lideranças políticas. Comentários como “Ex-presidiário Bolsonaro preso em 1987...” e “Onça pediu rachadinha?” evocam a necessidade de responsabilização por erros e irregularidades associadas ao governo e à sua trajetória. Por fim, manifesta-se o medo do autoritarismo e da degradação democrática. “Vamos votar em prol da democracia e não vamos deixar o ódio e a mentira vencer de novo” e “Não somos coveiro mas, domingo a gente enterra o Bolsonaro” reforçam a percepção de Bolsonaro como ameaça às instituições democráticas, alimentando o anseio de uma vitória eleitoral que reafirme direitos civis e barre projetos de natureza autoritária.

O recorte analítico aqui apresentado oferece uma demonstração empírica robusta da tese central que orienta nossa investigação mais ampla: os comentários não são meras reações, mas sim, espaços onde sentidos ambíguos se tornam explícitos e se cristalizam em leituras compartilhadas através de padrões de coordenação e enquadramentos transnacionais (ROBLES et al., 2022; WASSERMAN; FAUST, 2012).

A linguagem nos comentários reflete um grau profundo de polarização onde o outro lado não é apenas um adversário político, mas um inimigo a ser derrotado ("escorpiões", "capeta", "escória"). Para os apoiadores, a postagem era um sinal de força e esperança diante do cenário das sondagens que davam vantagem ao seu rival. Para a oposição, era mais uma expressão de hipocrisia a ser combatida nas redes sociais antes da eleição. O contexto explica por que os comentários são tão passionais, falaciosos e tribais: a eleição era vista não como uma disputa política normal, mas como uma batalha existencial pelo futuro do país.

Podemos pensar o contexto digital nos termos de Bourdieu (1991) como um campo de batalha simbólica em que plataformas como o X tornam-se arenas onde se disputa o poder de definir a realidade. Neste pequeno microcosmo social como espaço de lutas para imposição de significados legítimos, a disputa não se dava em torno do significado da imagem, mas pelo poder de definir a narrativa dominante em um momento crítico.

Na chave da semiótica da cultura e da teoria da recepção (ECO, 1986; JAUSS, 1984), a imagem não possui um significado fixo, mas se abre à cooperação interpretativa do leitor/espectador. O sentido emerge quando os usuários projetam sobre o texto ou a imagem os seus próprios repertórios culturais, horizontes de expectativa e experiências. Assim, a interpretação não é mera decodificação, mas um processo ativo de atribuição de significado, no qual cada recepção atualiza e reconfigura a obra. Os usuários não estavam simplesmente reagindo, estavam fechando o significado aberto e ambíguo da postagem original com base em seus horizontes de expectativa (JAUSS, 1984) e repertórios culturais.

Do ponto de vista de Fish (1985), os significados de textos não existem como entidades fixas e independentes, mas são produzidos e estabilizados pelas práticas interpretativas de grupos de leitores. Esses grupos, ou “comunidades interpretativas”, compartilham convenções, códigos, estratégias de leitura e repertórios culturais que moldam como os textos são compreendidos. Com base nessa perspectiva, pode-se considerar que, no espaço daquela publicação, os apoiadores e os opositores funcionaram como comunidades interpretativas distintas. Cada comunidade com suas

próprias estratégias para ler a imagem (como prova de força ou de hipocrisia), e os comentários servindo para reforçar internamente essas estratégias e consolidar a identidade do grupo. Por esse prisma, também podemos concluir que os comentários, na medida em que são lidos, têm a capacidade de moldar interpretações que não surgiriam de outra forma. A própria imagem publicada sem texto e, portanto, sem uma expressão clara do que o autor quer comunicar, atrai o leitor para os comentários, que pode ali confirmar ou confrontar uma interpretação, bem como expressar uma diferente das anteriores, criando uma trama de significados que vai continuar a ser tecida pelos comentários seguintes. Este espaço dos comentários torna-se, portanto, um ponto estratégico para formar opiniões, e não à toa, muitos agentes políticos empregam, por assim dizer, colaboradores, grupos profissionais e até robôs para intervir nesses espaços e acirrar polêmicas a fim de gerar engajamento e visibilidade nas redes sociais.

Uma análise argumentativa preliminar dos comentários aponta para o predomínio do uso de falácias, em especial, a Ad Hominem, que visa atacar o interlocutor em seu caráter, tais como, no caso de Bolsonaro ("capeta", "mentiroso", "covarde"), de Lula ("ladrão") ou da oposição ("esquerdistas incapazes"). As comunidades ideológicas, as falácias do Espantalho (Straw Man) são identificadas nas caricaturas ao oponente ("comunistas querem destruir a identidade católica"; "esquerda só prega ódio"); e também Apelo à Emoção, ao apelar principalmente para sentimentos (medo, orgulho, nojo) em vez de razão.

Quanto à qualidade dos argumentos, pode-se afirmar que em sua grande maioria são falaciosos e pouco ou nada argumentativos, com foco em atacar o adversário. Comparando-se os discursos de apoiadores e críticos, observa-se que os críticos foram os que mais exploraram a lógica formal e a factualidade, ancorando sua argumentação em contradições reais e em dados concretos, ainda que por vezes recorressem a falácias. Já os apoiadores se basearam sobretudo em apelos emocionais e identitários, com pouca substância lógica ou factual, destacando-se mais pelo volume de interações do que pela consistência argumentativa. Os memes, por sua vez, cumpriram a função de escape criativo, operando menos como argumento e como recurso humorístico e de ataque, voltados ao compartilhamento, sem compromisso com a veracidade ou com a lógica. Do ponto de vista argumentativo, a imagem analisada foi bem aproveitada pelos críticos, que souberam transformá-la em contra-argumentos persuasivos e fundamentados em fatos, enquanto os apoiadores a utilizaram como símbolo identitário e ataque aos adversários, com os memes aumentando o engajamento e reforçando interpretações sobre o resultado do debate.

7. RESULTADOS ESPERADOS (PROJEÇÕES CONCEITUAIS, SEM DADOS)

Embora a pesquisa ainda não apresente o resultado da análise de todos os dados empíricos coletados, é possível projetar resultados conceituais esperados a partir da fundamentação teórica e metodológica delineada. Esses resultados funcionam como hipóteses plausíveis sobre o comportamento discursivo em ambientes digitais dominados por atores da extrema-direita, especialmente quando se analisam posts audiovisuais ambíguos e suas respectivas camadas de comentários.

Com base nas análises amostrais e no que diz respeito aos padrões argumentativos, espera-se que os comentários apresentem uma maior densidade de falácias em comparação com as postagens originais. Entre essas falácias, destacam-se as de caráter *ad hominem* — utilizadas para desqualificar opositores sem enfrentar o mérito de seus argumentos — e as dicotomias falsas, que reduzem questões complexas a escolhas simplificadas entre “povo” e “elite” ou “liberdade” e “tirania”. A essas somam-se o uso recorrente de epítetos depreciativos e de definições persuasivas, estratégias que enquadram conceitos aparentemente neutros em molduras valorativas. Um elemento adicional esperado é a mobilização de argumentos de autoridade baseados em referências familiares ou religiosas, empregados como substitutos de fontes científicas ou jornalísticas, reforçando repertórios anti-intelectuais.

No nível dos repertórios contextuais, antecipa-se um padrão de alta legitimação intragrupo, em que comunidades alinhadas à extrema-direita reforçam mutuamente seus discursos, ao mesmo tempo em que produzem forte dissonância com repertórios científicos e jornalísticos. Esse efeito de espelhamento contribui para a formação de bolhas interpretativas, nas quais argumentos circulam de forma redundante e adquirem a força de consenso interno. Tais bolhas atuam como matriz de reforço das meta-identidades analíticas, estruturando coletivos que se reconhecem e se mobilizam a partir de repertórios compartilhados.

Quanto à dimensão multimodal, espera-se que os posts audiovisuais mais pobres em proposições explícitas induzam comentários ricos em interpretações extremadas. Imagens ou vídeos curtos, acompanhados de legendas vagas, funcionam como estímulos para que os comentários preencham os sentidos ausentes, muitas vezes de maneira radicalizada. Nesse processo, bots e contas

coordenadas desempenham papel fundamental ao semear leituras preferenciais que depois são amplificadas por outros usuários.

No plano transnacional, projeta-se a circulação de frames discursivos comuns entre diferentes contextos, conectando especialmente atores da extrema-direita na Europa e na América Latina. Entre os frames esperados destacam-se: (i) anti-imigração, associado à retórica da defesa nacional e da xenofobia; (ii) anti-OMS, mobilizado como ataque às instituições multilaterais de saúde; e (iii) anti-“globalistas”, narrativa conspiratória que integra tanto discursos europeus quanto latino-americanos. Esse compartilhamento de repertórios sugere a existência de um ecossistema comunicativo transnacional, em que estratégias e narrativas são adaptadas e re-contextualizadas para diferentes públicos, mas conservam uma base comum.

Em síntese, os resultados esperados apontam para a confirmação da tese central: os comentários são o espaço privilegiado da construção de sentidos explícitos, onde falácias, linguagem emotiva e repertórios de legitimação se articulam de forma a reforçar engajamento, radicalização e meta-identidades coletivas.

8. DISCUSSÃO

Do ponto de vista teórico, a proposta contribui ao integrar a análise de esquemas argumentativos e falácias (WALTON; REED; MACAGNO, 2008; WALTON; MACAGNO, 2015) com o estudo de repertórios de legitimação (MACAGNO; TREVISAN, 2024). Essa combinação permite compreender como os comentários em redes sociais funcionam como um verdadeiro dispositivo de legitimação discursiva, no qual sentidos ambíguos inicialmente sugeridos em postagens audiovisuais são transformados em interpretações explícitas, polarizadas e muitas vezes radicais.

No plano metodológico, o uso articulado do Model Context Protocol (MCP) e do Social Argumentative Analysis Framework (SAAF) responde a um desafio central das ciências sociais digitais: equilibrar escala de análise com auditabilidade crítica. Ao padronizar fluxos de dados e compará-los com repertórios contextuais, o modelo assegura transparência e governança, enquanto a escolha do comentário como unidade-chave permite integrar de forma inovadora a dimensão multimodal do discurso.

A categoria de meta-identidade reforça esse quadro ao clarificar como as interações discursivas se consolidam em três níveis analíticos: dinâmicas, marcadas pela coordenação e repetição em comentários; estados, configurados como perfis argumentativos relativamente estáveis; e resultados,

que correspondem à cristalização de identidades coletivas em torno de narrativas políticas. Esse encadeamento possibilita rastrear trajetórias de radicalização, iluminando os processos pelos quais interpretações fragmentárias se tornam comunidades discursivas coesas (FERREIRA; TREVISAN, 2025).

Ainda assim, a proposta apresenta limitações que precisam ser explicitadas. A primeira refere-se à dependência da anotação humana, indispensável para treinar e validar os modelos, mas sujeita a vieses interpretativos. Em segundo lugar, há o risco de viés de plataforma, uma vez que cada rede social impõe condições próprias de coleta e moderação. Além disso, permanece a ambiguidade persistente em imagens e vídeos, que mesmo interpretados por meio de comentários não são reduzidos a significados unívocos. Finalmente, existe o risco de “jogar com o indicador”, isto é, de os próprios atores analisados adaptarem suas estratégias ao saberem quais métricas ou sinais estão sendo monitorados.

Para mitigar esses desafios, propõem-se alguns caminhos: (i) uso de triangulação metodológica, combinando abordagens qualitativas e quantitativas; (ii) rotatividade de amostras, de modo a reduzir vieses relacionados a períodos ou temas específicos; (iii) auditorias éticas independentes, assegurando que decisões analíticas sigam princípios de transparência e respeito à privacidade; e (iv) adoção de protocolos de governança de dados, alinhados a recomendações internacionais (UNESCO, 2021; UNITED NATIONS, 2023).

Em conjunto, essas reflexões ressaltam tanto a originalidade quanto a robustez metodológica do modelo, ao mesmo tempo em que reconhecem suas limitações e propõem mecanismos para enfrentá-las de forma crítica e responsável.

9. IMPLICAÇÕES (POLÍTICAS PÚBLICAS, PLATAFORMA E PESQUISA)

Os resultados conceituais projetados neste trabalho permitem delinear algumas implicações relevantes para diferentes âmbitos: políticas públicas, plataformas digitais e pesquisa acadêmica.

No plano das políticas públicas, o modelo oferece instrumentos para apoiar a regulação proporcional e transparente do ambiente digital. A possibilidade de construir indicadores de risco argumentativo (densidade de falácias, presença de linguagem emotiva ou manipulação de esquemas) e de legitimidade contextual (grau de ressonância em repertórios científicos, jornalísticos, partidários ou extremistas) abre caminho para mecanismos de monitoramento que respeitem a liberdade de expressão, mas permitam sinalizar zonas críticas de manipulação discursiva. Esses indicadores não

substituem a decisão humana, mas podem servir como apoio à moderação, reforçando princípios de proporcionalidade, transparência e governança democrática (UNESCO, 2021; UNITED NATIONS, 2023).

No campo da literacia digital, a pesquisa contribui para o desenvolvimento de uma espécie de “biblioteca de truques retóricos”, voltada à educação midiática. A sistematização de esquemas argumentativos, falácias e repertórios recorrentes possibilita treinar cidadãos e comunidades a reconhecer padrões manipulativos, fortalecer sua resiliência a discursos desinformativos e compreender a lógica das interações online. Trata-se de um recurso pedagógico que vai além do *fact-checking*, ao evidenciar não apenas a veracidade das informações, mas os modos de construção da legitimidade discursiva.

Quanto às pesquisas futuras, o modelo propõe um caminho para aumentar a replicabilidade nas ciências sociais digitais. O uso de um codebook validado, em conjunto com os logs do MCP, assegura rastreabilidade e auditabilidade dos processos analíticos, reduzindo a opacidade frequentemente associada a métodos baseados em inteligência artificial. Além disso, a arquitetura metodológica aqui delineada é transferível para outros campos temáticos além do populismo e da extrema-direita, podendo ser aplicada, por exemplo, à análise de narrativas sobre mudanças climáticas, processos eleitorais ou imigração.

Assim, as implicações do modelo vão além da análise específica dos comentários em postagens ambíguas: oferecem contribuições práticas para a regulação, ferramentas pedagógicas para a educação midiática e avanços metodológicos para a pesquisa em ciências sociais digitais.

CONSIDERAÇÕES FINAIS

O ponto de partida desta pesquisa foi a tese de que “é nos comentários que a mágica acontece”: é nesse espaço interativo que se consolidam os sentidos mais explícitos das postagens audiovisuais ambíguas da extrema-direita. Se os posts originais operam com legendas vagas e imagens de forte apelo emocional, são os comentários que cristalizam interpretações, mobilizam falácias, ativam repertórios e promovem processos de engajamento e radicalização.

O modelo aqui proposto buscou oferecer meios analíticos para ver, medir e comparar essa dinâmica sem perder a profundidade interpretativa das ciências sociais. Ao articular a teoria da argumentação (esquemas, falácias e linguagem emotiva) com a noção de repertórios contextuais e a

categoria de meta-identidade, e ao operar sobre uma arquitetura técnica baseada em MCP e SAAF, a pesquisa combina escala computacional com sensibilidade crítica.

As contribuições concentram-se em três eixos principais: (i) teórico, ao integrar esquemas argumentativos, repertórios de legitimação e meta-identidade; (ii) metodológico, ao desenvolver um pipeline modular que assegura auditabilidade e governança; e (iii) aplicado, ao propor indicadores e painéis conceituais que podem apoiar tanto a regulação proporcional quanto a literacia digital.

Como próximos passos, delineiam-se três frentes de avanço: (i) a implementação piloto do modelo em recortes empíricos delimitados, a fim de calibrar procedimentos; (ii) a validação interplataforma, testando a robustez do método em contextos diversos como X, Instagram e TikTok; e (iii) a publicação aberta do codebook e dos repertórios contextuais em versões anonimizadas, garantindo replicabilidade e promovendo a circulação crítica desses instrumentos no campo das ciências sociais digitais.

Dessa forma, a proposta não encerra um processo, mas abre possibilidades para futuras contribuições em inovação metodológica, análise de dados, bem como para embasar políticas públicas de regulação e transparência das plataformas digitais, reafirmando a importância de estudar os comentários como *locus* fundamental da construção de sentidos em ambientes digitais.

REFERÊNCIAS

AGÊNCIA BRASIL. Conheça a cronologia da tentativa de golpe de Estado, segundo a PGR.

Agência Brasil, Rio de Janeiro, 19 fev. 2025. Disponível em:

<https://agenciabrasil.ebc.com.br/politica/noticia/2025-02/conheca-cronologia-da-tentativa-de-golpe-de-estado-segundo-pgr>. Acesso em: 10 set. 2025..

ANTHROPIC. Introducing the Model Context Protocol. 2024. Disponível em:

<https://www.anthropic.com/news/model-context-protocol>. Acesso em: 1 set. 2025.

BERGMANN, Eirikur. Three-Step Rhetorical Model of Conspiratorial Populism. In: CONRAD, Maximilian; NEWMAN, Saul (Orgs.). *Post-Truth Populism: A New Political Paradigm*. Cham: Palgrave Macmillan, 2024. p. 217-245.

CASAS-MAS, Belén; RODRÍGUEZ-SÁEZ, Arturo; GUTIÉRREZ, Inmaculada; BENAVIDES, Gerardo Javier; ROBLES, José Manuel. Voices of discontent: unpacking populist rhetoric in Spain and the rise of anti-European sentiment. *Media and Communication*, v. 13, 2025.

CONRAD, Maximilian; NEWMAN, Saul. Conclusions: ‘Mainstream’ Alarmism or ‘Critical’ Complacency? How to Approach Post-truth Populism. In: NEWMAN, Saul; CONRAD, Maximilian (ed.). *Post-Truth Populism: A New Political Paradigm*. Cham: Palgrave Macmillan, 2024. p. 333-340. DOI: <https://doi.org/10.1007/978-3-031-64178-7>

ECO, Umberto. *Lector in fabula: a cooperação interpretativa nos textos narrativos*. São Paulo: Perspectiva, 1986.

FERREIRA, Allan; TREVISAN, Ana. Critical Paths for Data Analysis in Social Sciences: Practical Application for Analyzing Data on Polarization, Populism, and Post-Truth (3Ps). 2025a. DOI: <https://doi.org/10.5281/zenodo.15765674>

FERREIRA, Allan Herison; TREVISAN, Ana Carolina. Meta-identidade e plataformas digitais: disputas por transparência e controle na criação algorítmica das identidades pelos sistemas de IA. *SciELO Preprints*, 2025b. DOI: <https://doi.org/10.1590/SciELOPreprints.13161>

HAMELEERS, Michael. The Epistemic Dimension of Populist Communication: Can Exposure to Populist Communication Spark Factual Relativism? In: CONRAD, Maximilian; NEWMAN, Saul (ed.). *Post-Truth Populism: A New Political Paradigm*. Cham: Springer Nature Switzerland, 2024. p. 121-144. DOI: https://doi.org/10.1007/978-3-031-64178-7_6.

HARSIN, Jayson. Post-truth Politics and Epistemic Populism: About (Dis-) Trusted Presentation and Communication of Facts, Not False Information. In: NEWMAN, Saul; CONRAD, Maximilian (ed.). *Post-Truth Populism: A New Political Paradigm*. Cham: Springer Nature Switzerland, 2024. p. 25-64. DOI: https://doi.org/10.1007/978-3-031-64178-7_2

HARSIN, Jayson. Regimes of posttruth, postpolitics, and attention economies. *Communication, Culture & Critique*, v. 8, n. 2, p. 327-333, 2015. DOI: <https://doi.org/10.1111/cccr.12097>

HOU, Xinyi; ZHAO, Yanjie; WANG, Shenao; WANG, Haoyu. Model context protocol (MCP): Landscape, security threats, and future research directions. arXiv preprint, arXiv:2503.23278, 2025. Disponível em: <https://arxiv.org/abs/2503.23278>. Acesso em: 2 ago. 2025.

JAUSS, Hans Robert. Toward an aesthetic of reception. *International Fiction Review*, v. 11, n. 2, 1984.

JEONG, Se-Hoon. Visual Metaphor in Advertising. *Journal of Marketing Communications*, v. 14, n. 1, p. 59-73, 2008.

KJELDSEN, Jens E. Pictorial Argumentation in Advertising. In: VAN EEMEREN, Frans H.; GARSSEN, Bart (Orgs.). *Topical Themes in Argumentation Theory: Twenty exploratory studies*. Dordrecht: Springer Netherlands, 2012. p. 239-255.

KJELDSEN, Jens E. The study of visual and multimodal argumentation. *Argumentation*, v. 29, n. 2, p. 115-132, 2015.

KURBANOV, Iskandar. 9 AI-Powered Social Media Engagement Strategies for 2025. SendFame.com, 02 ago. 2025. Disponível em: <https://sendfame.com/blog/social-media-engagement-strategies>. Acesso em: 01 set. 2025

MACAGNO, Fabrizio. Info-arguments: Dialogical Ambiguity, Argument Interpretation, and the Problem of Meaning in Argumentation. *Argumentation*, 2025. Disponível em: <https://doi.org/10.1007/s10503-025-09667-y>

MACAGNO, Fabrizio; TREVISAN, Ana Carolina. Strategic communication in the Covid-19 pandemic: Uses of arguments and manipulative tactics in institutional social media communication. In: ILIE, Cornelia (ed.). *Manufacturing Dissent: Manipulation and Counter-Manipulation in Times of Crisis*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2024. (Pragmatics and Beyond New Series, vol. 339). p. 240-283. DOI: 10.1075/pbns.339.08mac.

PHILIPPOV, Karin. Considerações acerca da representação dos indígenas nas pinturas anchietanas de Benedito Calixto. *Imagem: Revista de História da Arte*, v. 2, n. 2, 2023.

PORTINARI, Natália. Quem é Soraya Thronicke, ex-aliada de Bolsonaro que perguntou se Padre Kelmon não tem ‘medo de ir para o inferno’. *O Globo*, Brasília, 30 set. 2022. Disponível em: <https://oglobo.globo.com/politica/eleicoes-2022/noticia/2022/09/quem-e-soraya-thronicke-ex-aliada-de-bolsonaro-que-perguntou-se-padre-kelmon-nao-tem-medo-de-ir-para-o-inferno.ghtml>. Acesso em: 1 set. 2025.

PRAZERES, Leandro. Como aliança entre Trump e big techs aumenta pressão sobre governo Lula e STF. *BBC News Brasil*, 21 jan. 2025. Disponível em: <https://www.bbc.com/portuguese/articles/cx2k80ekl3ko>. Acesso em: 10 ago. 2025.

RAY, Partha Pratim. A survey on model context protocol: Architecture, state-of-the-art, challenges and future directions. *Authorea Preprints*, 2025. Disponível em: <https://doi.org/10.22541/au.172250325.99455264/v1>

- ROBLES, José Manuel et al. When negativity is the fuel: Bots and political polarization in the COVID-19 debate. *Comunicar: Media Education Research Journal*, v. 30, n. 71, p. 59-71, 2022.
- TECH FOCUS PRO. 8 Steps to Blow Up on Social Media. Vocal Media, [s. l.], 2025. Disponível em: <https://vocal.media/education/8-steps-to-blow-up-on-social-media>. Acesso em: 1 set. 2025.
- UNESCO. When countering digital disinformation while respecting freedom of expression becomes a fine balancing act. 2021. Disponível em: <https://www.unesco.org/en/articles/when-countering-digital-disinformation-while-respecting-freedom-expression-becomes-fine-balancing>. Acesso em: 10 ago. 2025.
- UNITED NATIONS. Information integrity on digital platforms. Our Common Agenda Policy Brief 8. New York: United Nations, June 2023. Disponível em: <https://www.un.org/en/common-agenda>. Acesso em: 10 ago. 2025.
- WALTON, Douglas. Why fallacies appear to be better arguments than they are. *Informal Logic*, v. 30, n. 2, p. 159-184, 2010.
- WALTON, Douglas; MACAGNO, Fabrizio. A classification system for argumentation schemes. *Argument & Computation*, v. 6, n. 3, p. 219-245, 2015.
- WALTON, Douglas; MACAGNO, Fabrizio. Emotive meaning in political argumentation. *Informal Logic*, v. 39, n. 3, p. 229-261, 2019.
- WALTON, Douglas; REED, Chris; MACAGNO, Fabrizio. *Argumentation Schemes*. Cambridge: Cambridge University Press, 2008.
- WASSERMAN, Stanley; FAUST, Katherine. *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press, 2012.
- WITTGENSTEIN, Ludwig. *Philosophical investigations*. Oxford: Wiley-Blackwell, 2009.

DECLARAÇÃO DE DISPONIBILIDADE DE DADOS DA PESQUISA:

O conjunto de dados de apoio aos resultados deste artigo foi disponibilizado em um repositório do GitHub e pode ser acessado em <https://github.com/anacarolinatcf-hash/anpocs49>.

FINANCIAMENTO:

O presente trabalho foi realizado com apoio da Fundação para a Ciência e Tecnologia (FCT - Portugal) e com apoio do Instituto de Comunicação da NOVA (ICNOVA), Universidade NOVA de Lisboa, Portugal. Códigos de financiamento das bolsas doutorais: 2022.14807.BD (Autor 1) e UI/BD/153755/2022 (Autor 2).

CONTRIBUIÇÃO DAS/DOS AUTORES/AS:

Ferreira, Allan e Trevisan, Ana Carolina: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration, Supervision, Funding acquisition.

DECLARAÇÃO DE CONFLITO DE INTERESSE:

Os autores declaram que não há conflito de interesses a mencionar.

DECLARAÇÃO USO DE INTELIGÊNCIA ARTIFICIAL:

Os autores declaram que fizeram uso de ferramentas de Inteligência Artificial de modo assistivo para revisar o conteúdo, ajustar estilo, testar estruturas e traduzir partes do texto.

MINIBIOGRAFIAS DOS/DAS AUTORAS DO PAPER

Ana Carolina Trevisan é doutoranda em Ciências da Comunicação na NOVA de Lisboa, mestra em Sociologia pela USP e bacharel em Sociologia e Psicologia. Bolsista da FCT pelo IFILNOVA, pesquisa "Estratégias argumentativas nas redes sociais da extrema direita no Brasil e Portugal".

Allan Herison Ferreira é doutorando em Ciências da Comunicação na Universidade NOVA de Lisboa, com mestrado em Sociologia pela USP e bacharelado em Ciências Sociais pela USP. Bolsista da FCT afiliado ao ICNOVA, pesquisa "Estratégias identitárias em filmes digitais amadores".

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores declaram que uma vez que um manuscrito é postado no servidor SciELO Preprints, o mesmo só poderá ser retirado mediante pedido à Secretaria Editorial do SciELO Preprints, que afixará um aviso de retratação no seu lugar.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença [Creative Commons CC-BY](#).
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.