

# Percepções sobre Desmatamento e Queimadas no Brasil: Uma Análise de Sentimentos, Tópicos e Engajamento no Bluesky

Bruno Henrique Freitas de Melo<sup>1</sup>, Raoni Simoes Ferreira<sup>2</sup>, Catarina de Souza Costa<sup>1</sup>

<sup>1</sup>Centro de Ciências Exatas e Tecnológicas - Universidade Federal do Acre (UFAC)  
Rio Branco – AC – Brasil

<sup>2</sup>Instituto de Computação – Universidade Federal do Amazonas (UFAM)  
Manaus – AM – Brasil

bruno.henrique@sou.ufac.br, raoni@icompu.ufam.edu.br, catarina.costa@ufac.br

**Abstract.** *This study analyzes perceptions and sentiments expressed by Brazilian users on Bluesky regarding deforestation and wildfires during August and September 2024. A dataset of approximately 126,000 public posts was analyzed using Natural Language Processing techniques. The methodology combined sentiment classification with the LeIA lexicon and topic modeling using Latent Dirichlet Allocation. Results indicate a predominance of indignation directed toward political and economic actors. Thematic analysis revealed discussions about political responsibility, everyday wildfire impacts, social consequences, and discourse framing fires as criminal acts.*

**Resumo.** *Este trabalho analisa percepções e sentimentos de usuários brasileiros no Bluesky sobre desmatamento e queimadas no Brasil durante agosto e setembro de 2024. Foi construído um conjunto de dados de aproximadamente 126 mil postagens, analisado com técnicas de Processamento de Linguagem Natural. A metodologia combinou classificação de sentimentos com o léxico LeIA e modelagem de tópicos por Latent Dirichlet Allocation. Os resultados indicam predominância de indignação direcionada a atores políticos e econômicos. A análise temática revelou discussões sobre responsabilização política, impactos cotidianos das queimadas, consequências sociais e criminalização dos incêndios.*

## 1. Introdução

As queimadas e o desmatamento no Brasil são problemas ambientais históricos, e o ano de 2024 foi marcado por uma temporada de queimadas especialmente devastadora no Brasil. Somente em agosto daquele ano, a área queimada no país aumentou 149% em relação ao ano anterior, totalizando 5,65 milhões de hectares devastados [Brasil de Fato 2024].

As consequências dessas queimadas impactaram diretamente diversas capitais brasileiras. Praticamente todas as capitais do país sofreram com a fumaça e a fuligem, afetando a qualidade do ar e a saúde da população [DW Brasil 2024].

Diante desse cenário de crise ambiental com impactos visíveis na vida urbana, a população usou às redes sociais como canal de expressão, informação e mobilização em tempo real. Precisa-se considerar o fato de que o Brasil possui uma das populações mais

conectadas do mundo, com cerca de 144 milhões de usuários de redes sociais em 2024 [Datareportal 2024].

Dentre essas plataformas, o Bluesky Social<sup>1</sup> destaca-se por sua arquitetura descentralizada e pelo crescimento acelerado de usuários brasileiros durante o terceiro trimestre de 2024. Esse aumento foi impulsionado pelo bloqueio temporário da rede social X<sup>2</sup> (anteriormente Twitter) no país em 2024 [BBC News Brasil 2024]. Desenvolvido como uma alternativa aberta ao X, o Bluesky é construído sobre o AT Protocol, uma arquitetura que permite interoperabilidade e garante maior transparência no acesso aos dados [Kleppmann et al. 2024].

Nesse contexto, este trabalho propõe analisar as percepções e sentimentos de usuários brasileiros no Bluesky sobre o desmatamento e as queimadas ocorridas entre agosto e setembro de 2024. A partir dessa análise, busca-se responder à seguinte pergunta de pesquisa: “*Quais são as percepções e sentimentos mais comuns expressos por brasileiros no Bluesky em relação ao desmatamento e às queimadas durante o período crítico de agosto a setembro de 2024?*”.

Para investigar esse fenômeno, utiliza-se uma abordagem quantitativa baseada em técnicas de mineração de texto para extrair padrões relevantes de grandes volumes de dados textuais não estruturados [Gaikwad et al. 2014] e abordagens de Processamento de Linguagem Natural (PLN) para identificar estruturas linguísticas e padrões semânticos em grandes coleções de textos [Jurafsky and Martin 2024, Nadkarni et al. 2011]. A partir dessas técnicas, identificam-se padrões emocionais, tópicos predominantes de discussão e as correlações entre a polaridade e o engajamento na rede.

O artigo está organizado da seguinte forma: A Seção 2 apresenta os trabalhos relacionados sobre análise de sentimentos, mineração de texto e estudos envolvendo o Bluesky. A Seção 3 descreve a metodologia adotada, incluindo os procedimentos de coleta, pré-processamento e análise dos dados. Na Seção 4 são apresentados e discutidos os resultados obtidos. Por fim, a Seção 5 apresenta as considerações finais, limitações do estudo e possíveis direções para trabalhos futuros.

## 2. Trabalhos Relacionados

A análise de conteúdo em redes sociais por meio de técnicas de mineração de texto e análise de sentimentos tem sido amplamente utilizada para compreender percepções públicas sobre diferentes fenômenos sociais. Estudos nessa área demonstram que a aplicação de métodos computacionais permite identificar padrões discursivos, emoções predominantes e estruturas temáticas em grandes volumes de dados textuais provenientes de plataformas digitais.

Diversas pesquisas têm explorado especificamente o debate ambiental em redes sociais. A publicação “*Análise de sentimento em tweets relacionados ao desmatamento da floresta amazônica*” [Paes et al. 2022], por exemplo, investiga sentimentos expressos em tweets relacionados ao desmatamento da Amazônia brasileira, identificando predominância de emoções negativas e que os picos de atividade e de emoções negativas, como a raiva, estavam diretamente correlacionados com eventos de grande repercussão.

---

<sup>1</sup><https://bsky.app/>

<sup>2</sup><https://x.com/>

Além disso, trabalhos que aplicam mineração de texto para analisar discursos online em contextos sociais complexos demonstram o potencial dessas técnicas para revelar padrões temáticos e dinâmicas de debate em ambientes digitais. A publicação “*An analysis of violence against women based on victims report*” [Corrêa and Faria 2021], por exemplo, utiliza modelagem de tópicos para identificar cenários recorrentes em relatos de violência, evidenciando a capacidade dessas abordagens de extrair estruturas discursivas relevantes a partir de grandes corpora textuais.

Paralelamente, estudos recentes têm investigado novas plataformas descentralizadas, como o Bluesky, destacando suas características arquiteturais e seu potencial para novos modelos de interação social online [Kleppmann et al. 2024, Sahneh et al. 2024, Zheng et al. 2025]. Pesquisas recentes também passaram a explorar empiricamente o comportamento dos usuários na plataforma. Por exemplo, o estudo “*Emerging Digital Spaces: The Case of Bluesky’s Rise During 2024 Brazilian Elections*” [Buzelin et al. 2025] analisa as percepções dos usuários da rede Bluesky durante o período das eleições municipais brasileiras de 2024, também explorando o fenômeno da migração de usuários após o bloqueio do X.

Nesse contexto, este trabalho busca sua contribuição ao investigar percepções e repercussões associadas ao debate sobre desmatamento e queimadas no Bluesky. Até o momento da realização desta pesquisa, não foram encontrados estudos voltados à análise das discussões sobre queimadas e desmatamento no contexto brasileiro dentro do Bluesky, particularmente durante o período de agosto e setembro de 2024, marcado pela migração de usuários após o bloqueio temporário do X no Brasil.

### 3. Metodologia

Este estudo caracteriza-se como uma pesquisa quantitativa de natureza exploratória voltada à análise de percepções e padrões temáticos sobre desmatamento e queimadas no Brasil em publicações feitas por usuários brasileiros na rede social Bluesky. A janela de publicação usada neste estudo é de agosto a setembro de 2024. A abordagem metodológica baseia-se em técnicas de mineração de texto e PLN, combinando análise de sentimentos, modelagem de tópicos e análise de métricas de engajamento para examinar como o debate público digital se manifesta na plataforma. O fluxo geral das etapas é apresentado na Figura 1 e descritos nos parágrafos a seguir.

O código utilizado na coleta e processamento dos dados está disponível em repositório público<sup>3</sup>.

#### 3.1. Coleta de Dados

A coleta foi realizada por meio da API pública do Bluesky, baseada no AT Protocol. Utilizou-se o endpoint `app.bsky.feed.searchPosts` com os seguintes parâmetros: `limit=100` (máximo de postagens por página), `lang=pt` (filtro de idioma), `sort=latest` (ordenação cronológica), e `since/until` delimitando o período de 1º de agosto a 30 de setembro de 2024. A paginação foi gerenciada via parâmetro `cursor`.

Foram utilizados 11 termos de busca: *queimada*, *queimadas*, *desmatamento*, *fumaça*, *mudanças climáticas*, *aquecimento global*, *efeito estufa*, *amazônia*, *pantanal*,

<sup>3</sup><https://github.com/I-Am-BrunoHFMelo/Leia-Bluesky-Queimadas-2024/blob/master/coleta.py>

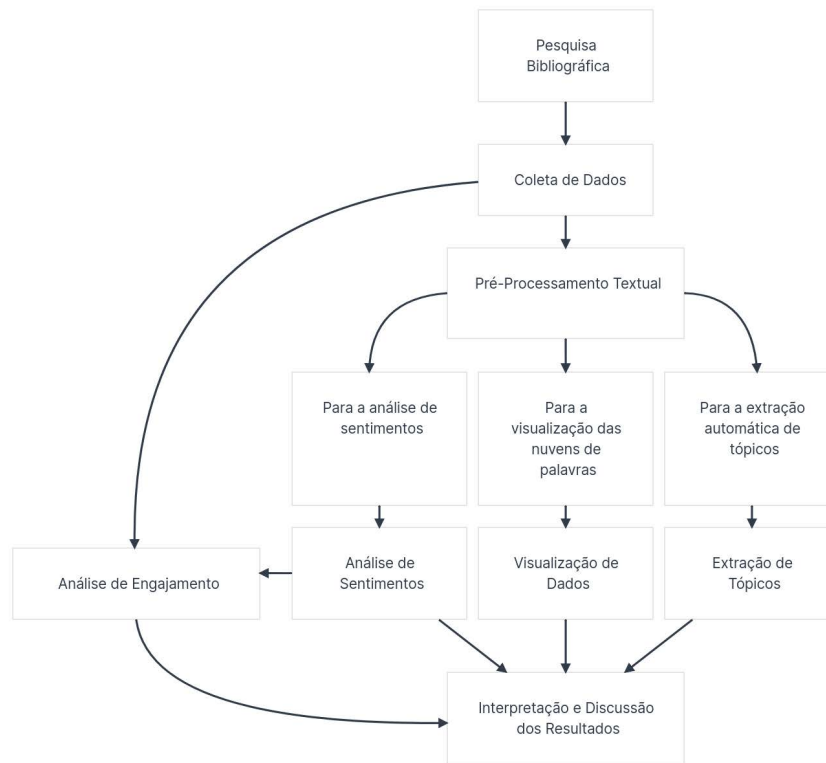


Figura 1. Fluxo metodológico adotado na pesquisa.

*agronegócio e incêndio*. Para cada postagem coletada, recuperaram-se respostas via `app.bsky.feed.getPostThread` com `depth=6` e `parentHeight=80`, limitando a profundidade da árvore de conversa. A coleta de respostas foi paralelizada com `ThreadPoolExecutor` (`max_workers=5`) e intervalos de 0,4s entre requisições para respeitar limites da API.

Além do conteúdo textual, foram coletadas métricas de engajamento: curtidas (`likeCount`), respostas (`replyCount`), republicações (`repostCount`) e citações (`quoteCount`).

### 3.2. Pré-processamento Textual

Após a coleta, os dados passaram por três fluxos de pré-processamento textual, adaptados às técnicas analíticas empregadas e ilustrados na Figura 1:

- **Análise de sentimentos (LeIA):** O texto foi mantido próximo da forma original, preservando pontuação expressiva, emojis e capitalização, uma vez que o modelo LeIA utiliza esses elementos como indicadores de intensidade emocional.
- **Modelagem de tópicos (LDA):** Aplicou-se normalização textual (minúsculas), remoção de URLs, menções, hashtags, emojis e números. Em seguida, realizou-se lematização com o `spaCy` (`pt_core_news_sm`), mantendo apenas substantivos, verbos e adjetivos. Também foi aplicada remoção de *stopwords* (lista padrão do NLTK combinada com uma lista customizada). Por fim, o dicionário foi filtrado por frequência, com os parâmetros `no_below=10` (palavras presentes em menos

de 10 documentos) e  $no\_above=0.3$  (palavras presentes em mais de 30% dos documentos).

- **Nuvens de palavras:** Para a visualização lexical, utilizou-se um pré-processamento simplificado voltado à representação da frequência de termos. Foram removidos URLs, menções, hashtags e *stopwords*, sem aplicação de lematização. Além disso, variantes gráficas de termos relevantes foram unificadas manualmente por meio de mapeamento lexical (por exemplo, *amazônia* → *amazonia*), a fim de evitar fragmentação visual das frequências.

### 3.3. Análise de Sentimentos

A análise de sentimentos foi conduzida utilizando o modelo LeIA (Léxico para Inferência Adaptada), uma adaptação para o português do modelo VADER, utilizado principalmente para análise de sentimentos em textos de redes sociais [Hutto and Gilbert 2014, Almeida 2018]. Esse tipo de abordagem baseia-se em léxicos e regras heurísticas para classificar textos em polaridades, retornando scores de positividade, neutralidade, negatividade e um *compound* normalizado entre -1 e +1. A classificação final utilizou os limiares padrão:

- Positivo:  $compound > 0,05$
- Neutro:  $-0,05 \leq compound \leq 0,05$
- Negativo:  $compound < -0,05$

A partir das classificações, calcularam-se métricas médias de engajamento por categoria de sentimento.

### 3.4. Modelagem de Tópicos

A coleção de dados textual foi representada no modelo *bag-of-words*, utilizando os filtros de frequência definidos na etapa de pré-processamento (o 3.2). Em seguida, aplicou-se o algoritmo Latent Dirichlet Allocation (LDA) [Blei et al. 2003] com uso da biblioteca Gensim.

O LDA assume que as palavras que ocorreram simultaneamente e de forma comum são agrupadas em conjuntos de tópicos (com seus respectivos valores de probabilidade). Cada tópico consiste em uma distribuição de probabilidade do vocabulário de palavras do conjunto de dados de postagens. Por fim, cada postagem do conjunto de dados é representada nos termos desses tópicos.

Assim, para determinar o número ótimo de tópicos (K) para esse conjunto de dados, realizou-se *grid search* no intervalo  $K \in [4,12]$  utilizando a métrica de coerência  $c_v$ . Nos experimentos executados, o melhor modelo LDA obtido foi aquele definido para  $K=4$  (coerência  $c_v = 0,495$ ), pois apresentou melhor interpretabilidade dos temas identificados. O modelo obtido foi treinado com  $passes=12$  e  $iterations=500$ , com hiperparâmetros  $\alpha$  e  $\eta$  em *auto*. Após o treinamento, cada publicação foi representada pela distribuição de probabilidade associada aos tópicos identificados, sendo considerado o tópico dominante aquele com maior probabilidade.

## 4. Resultados

Os resultados das análises foram apresentados por meio de diferentes visualizações. Foram geradas nuvens de palavras (*wordcloud*) para a coleção de dados textual global,

para cada termo de busca e para os tópicos identificados pelo modelo LDA. Além disso, a relação entre polaridade de sentimento e métricas de engajamento foi analisada por meio de *heatmaps* (mapas de calor) produzidos com as bibliotecas *matplotlib* e *seaborn* [Gu 2022, Heimerl et al. 2014].

#### 4.1. Caracterização dos Dados Coletados e Evolução Temporal

O conjunto de dados analisado compreende 126.164 publicações em português coletadas entre 1º de agosto e 30 de setembro de 2024, totalizando 124.807 registros após remoção de nulos. Destes, 65.212 (51,7%) correspondem às postagens originais e 60.952 (48,3%) respostas a essas postagens.

A evolução temporal (Figura 2) revela atividade em agosto e explosão abrupta no início de setembro, com pico em 12 de setembro (aproximadamente 7.200 postagens). Esse padrão correlaciona-se diretamente com o bloqueio do X no Brasil a partir de 30 de agosto, sugerindo que parte significativa do discurso analisado representa migração de usuários e tensões pré-existentes, não necessariamente formação da comunidade no Bluesky. A redução progressiva no final de setembro indica diminuição da atenção imediata ao tema, embora o volume permaneça superior ao patamar pré-crise.

O conjunto de dados utilizado foi anonimizado. Os materiais suplementares contendo todas as análises, estão disponível em um repositório de dados online<sup>4</sup>.

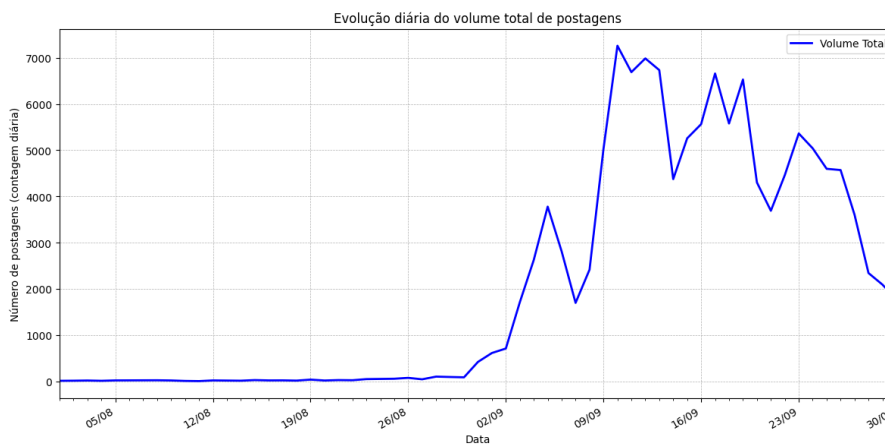


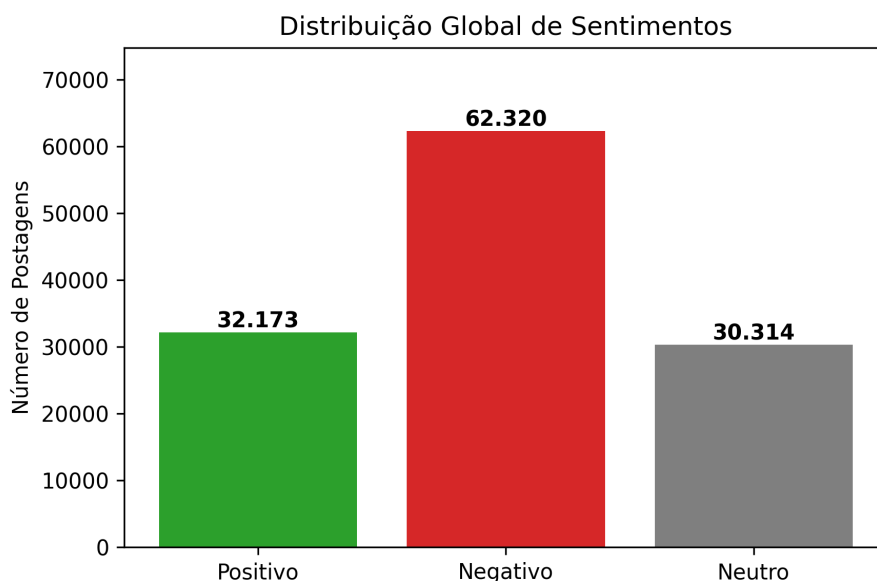
Figura 2. Evolução diária de postagens

#### 4.2. Análise de Sentimentos e Engajamento

A distribuição de sentimentos (Figura 3) revela predominância de polaridade negativa: 49,93% (62.320 postagens) contra 25,78% positivas (32.173) e 24,29% neutras (30.314).

A intensidade média do sentimento foi calculada a partir da média do escore *compound* gerado pelo modelo LeIA para cada grupo de polaridade (Seção 3.3). As postagens classificadas como negativas apresentaram média de -0,519, enquanto as positivas apresentaram média de 0,412. Isso indica que, além de mais frequentes, as manifestações negativas também tendem a apresentar maior intensidade emocional.

<sup>4</sup><https://github.com/I-Am-BrunoHFMelo/Leia-Bluesky-Queimadas-2024/tree/master/resultados>



**Figura 3. Frequência de postagem por categorias de sentimentos na base de dados.**

O engajamento médio por categoria (Figura 4) confirma tendência de maior repercussão para conteúdo negativo: 11,08 curtidas médias contra 10,09 das positivas, com vantagem semelhante em respostas (0,65 vs. 0,59), republicações (1,87 vs. 1,67) e citações (0,30 vs. 0,22). Postagens neutras apresentam engajamento substancialmente inferior em todas as métricas.

Entretanto, análise granular por termo de busca revela exceções instrutivas. As visualizações lexicais detalhadas associadas a esses termos encontram-se disponíveis no material suplementar do artigo. Para *fumaça*, postagens positivas obtiveram média de 3,79 curtidas contra 3,04 das negativas devido a mensagens de alívio pela dissipação da fumaça e chegada de chuvas (ex.: “Céu amanheceu azul... A chuva que limpou tudo”, 455 curtidas, *compound* 0,40).

De forma mais pronunciada, “efeito estufa” exibiu 57,01 curtidas médias em postagens positivas contra 10,74 em negativas. Análise qualitativa revela fenômeno duplo. Parte do engajamento direciona-se a conteúdos genuinamente construtivos — como por exemplo: discussões sobre o papel do Brasil no G20 climático ou o Protocolo de Montreal — sugerindo que, para temas científicos globais, o público valoriza informação propositiva. Simultaneamente, a anomalia foi amplificada por limitações do modelo LeIA: termos quantitativos absolutos (‘maior’, ‘líder’) foram classificados como positivos independentemente de contexto crítico (e.g., ‘Amazônia foi MAIOR emissora de gases’, 7.239 curtidas, *compound* 0,76). A dificuldade do modelo em discernir ironia ou crítica velada distorceu parcialmente a categoria.

### 4.3. Análise Lexical e Modelagem de Tópicos

Com o objetivo de identificar os termos mais recorrentes, foi gerada uma nuvem de palavras a partir de todas as publicações coletadas (Figura 5). Ela evidencia quatro eixos discursivos distintos que estruturam o debate. Análises léxicas individuais por termo de

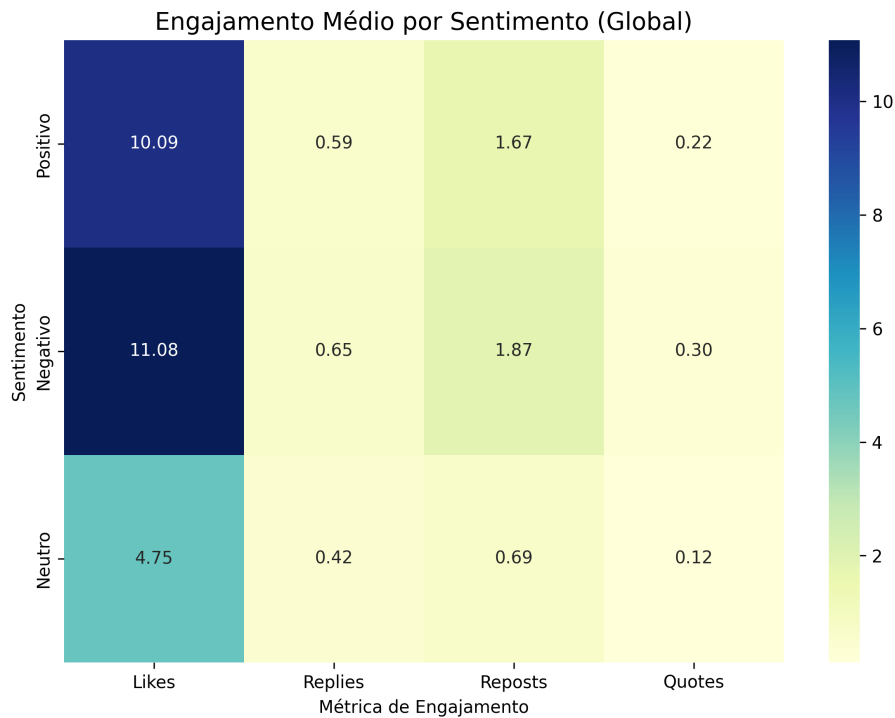


Figura 4. Engajamento médio por categoria de sentimento.

busca — disponíveis no material complementar<sup>5</sup> — revelam nuances específicas de cada eixo.

O **eixo central** (*queimadas, incêndio, amazônia*) mescla fenômeno físico (fogo, fumaça), atores políticos (governo, Lula, agro) e localização (Brasil, Amazônia), funcionando como confluência dos demais eixos. O **eixo político-econômico** (*desmatamento, agronegócio*) orbita atribuição de responsabilidade, com vocabulário de legalidade (ilegal, crime, IBAMA) e acusação (culpa, direita). O **eixo sensorial** (*fumaça*) traz léxicos de impacto corporal imediato: cheiro, respirar, garganta, nariz, casa, cidade — e, notavelmente, *chuva* como anseio coletivo. O **eixo científico** (*efeito estufa, mudanças climáticas, aquecimento global*) adota registro formal (emissões, carbono, gases, planeta), embora *aquecimento global* funcione como conceito-ponte, conectando vocabulário científico a impactos cotidianos (calor, morrer).

Notavelmente, *pantanal* foi singularmente moldado por interferência cultural: a proeminência de *novelas, renascer* e vocabulário afetivo (tristeza, esperança, paixão) reflete a influência da telenovela.

A modelagem de tópicos via LDA (Tabela 1) confirma e qualifica essa estrutura. O número ótimo de tópicos ( $K=4$ , coerência  $cv=0,495$ ) foi determinado por *grid search* no intervalo [4,12]. Os tópicos identificados mapeiam-se diretamente aos eixos lexicais: Tópico 1 (Responsabilidade Político-Institucional) ao eixo político-econômico; Tópico 2 (Relatos do Cotidiano) ao eixo sensorial; Tópico 3 (Consequências Sociais) articulando agronegócio, poder e impacto humano; Tópico 4 (Ação Criminosa) enquadrando incêndios como questão de legalidade e fiscalização.

<sup>5</sup><https://github.com/I-Am-BrunoHFMelo/Leia-Bluesky-Queimadas-2024/tree/master/resultados>

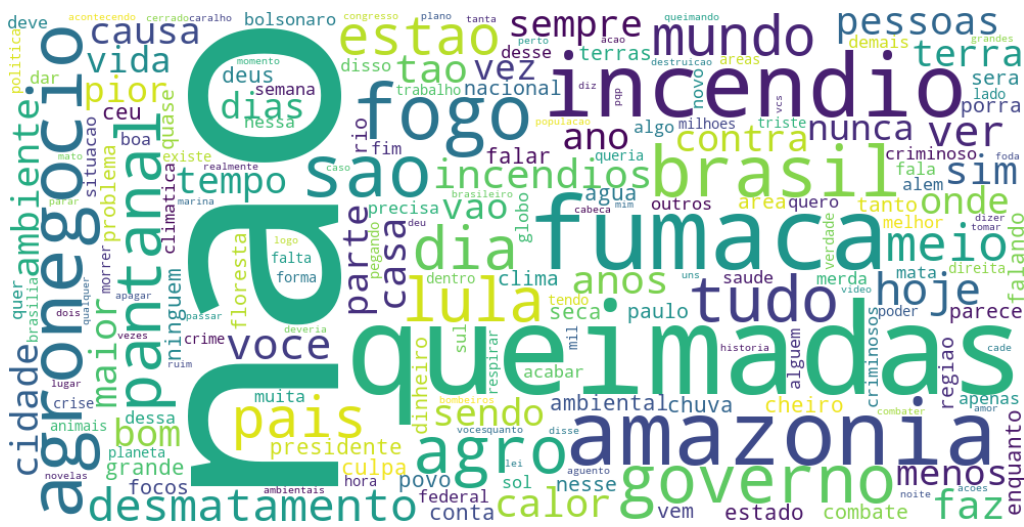


Figura 5. Nuvem de palavras representativa a todos os dados coletados.

Tabela 1. Principais tópicos identificados pelo modelo LDA.

| Tópico | Interpretação                           | Principais termos  |
|--------|---|--|
| 1      | Responsabilidade político-institucional | brasil, governo, agro, lula, país, amazônia, ambiente, culpa |
| 2      | Relatos do cotidiano e impacto direto   | fogo, fumaça, cidade, casa, bom-beiro, apagar, dia           |
| 3      | Consequências sociais e econômicas      | agronegócio, poder, dinheiro, vida, povo, trabalho           |
| 4      | Criminalização e combate aos incêndios  | incêndio, criminoso, crime, foco, área, estado, pantanal     |

Uma visualização interativa dos tópicos gerados pelo modelo LDA está disponível online<sup>6</sup>

#### 4.4. Discussão dos Resultados

O achado mais proeminente é a dominância de um sentimento de profunda indignação. A predominância quantitativa (49,93%) e de intensidade (-0,519) do sentimento negativo, somada à onipresença do termo “não”, aponta para um discurso centrado na frustração e na crítica. A modelagem de tópicos permite qualificar a natureza dessa indignação: ela é eminentemente política e institucional. O Tópico 1 (Responsabilidade Político-Institucional), com termos como governo, agro, culpa e Lula, e o Tópico 4 (Ação Criminalosa e Combate ao Fogo), com incêndio, criminoso e crime, mostram que o debate focou na atribuição de responsabilidade, percebendo a crise como consequência de falha de governança e atos intencionais.

A análise das publicações coletadas no Bluesky revela que a reação dos usuários brasileiros à crise de queimadas e desmatamento em 2024 apresenta um padrão discursivo predominantemente crítico. A análise de sentimentos indica a predominância de polaridade negativa, sugerindo que o debate foi amplamente marcado por frustração e

<sup>6</sup><https://i-am-brunohfmelo.github.io/Leia-Bluesky-Queimadas-2024/>

indignação. A modelagem de tópicos complementa esse resultado ao demonstrar que grande parte das discussões se organiza em torno da atribuição de responsabilidades políticas e institucionais, além de narrativas relacionadas ao impacto cotidiano das queimadas e às consequências sociais do problema ambiental. A análise lexical reforça essa interpretação ao evidenciar a presença recorrente de termos associados tanto ao evento ambiental quanto a atores políticos e econômicos. Em conjunto, os resultados indicam que o debate online sobre queimadas no período analisado combinou críticas institucionais, relatos de impacto direto e discussões mais amplas sobre as implicações sociais e ambientais do fenômeno.

A regra geral observada foi que conteúdo negativo gerou maior repercussão, apontando emoções negativas como catalisadoras de mobilização online em eventos de crise [Beigi et al. 2016]. No entanto, a maior repercussão de postagens positivas para *fumaça* — ligado ao Tópico 2, de relatos do cotidiano — sugere que, para impactos diretos e sensoriais, mensagens de alívio e resolução geram forte ressonância.

A anomalia mais significativa, contudo, ocorreu com *efeito estufa*, onde conteúdo classificado como positivo obteve engajamento massivo. Análise qualitativa revelou fenômeno duplo: por um lado, parte do engajamento direcionou-se a conteúdos genuinamente construtivos; por outro, o escore `compound` médio foi inflado por limitação do LeIA, que classificou como positivas publicações noticiando o Brasil como “maior” ou “líder” em emissões.

Conclui-se, portanto, que a reação pública digital à crise variou principalmente entre a cobrança política, o lamento pessoal e, em certo grau, a valorização do debate científico.

## 5. Considerações Finais

Diante dos resultados apresentados, é possível responder à pergunta central deste estudo: “Quais são as percepções e sentimentos mais comuns expressos por brasileiros no Bluesky em relação ao desmatamento e às queimadas no Brasil durante o período crítico de agosto a setembro de 2024?”. A análise revelou que a percepção predominante foi de intensa indignação e frustração, associada a atores políticos e econômicos percebidos como responsáveis pela crise ambiental. No entanto, esse sentimento central coexistiu com outras formas de discurso igualmente relevantes, incluindo relatos pessoais sobre os impactos diretos da fumaça e do fogo no cotidiano e discussões de caráter científico e informativo.

De forma geral, os resultados indicam que a reação pública digital à crise ambiental não se configurou como uma manifestação monolítica de negatividade, mas sim como um ecossistema discursivo complexo e multifacetado. Embora a indignação política tenha ocupado posição central no debate, ela conviveu com outros espaços de discussão, como a experiência cotidiana da população afetada e a circulação de conteúdos educativos sobre fenômenos ambientais. Esse cenário sugere que diferentes interpretações da crise estão associadas a vocabulários, preocupações e formas de engajamento distintas dentro da esfera pública digital.

Um resultado particularmente relevante foi a observação de um aparente paradoxo no engajamento das publicações. Embora conteúdos de polaridade negativa tenham

apresentado maior propensão à viralização, alguns conteúdos de natureza informativa e científica, assim como mensagens de alívio, também alcançaram níveis mais altos quando voltados à positividade.

Do ponto de vista científico, este estudo contribui para a literatura ao explorar o Bluesky como fonte de dados para análise social no contexto brasileiro. Como plataforma emergente, ainda pouco investigada na literatura acadêmica, o Bluesky representa um novo espaço de observação da opinião pública digital. Além disso, ao combinar análise de sentimentos, análise lexical e modelagem de tópicos, o trabalho buscou oferecer uma compreensão mais detalhada da diversidade de discursos presentes nas discussões sobre a crise ambiental.

Entretanto, alguns limites devem ser considerados na interpretação dos resultados. O perfil demográfico dos usuários do Bluesky pode não ser representativo da população brasileira em geral, o que restringe a generalização dos achados. Além disso, o recorte temporal de dois meses oferece apenas um retrato momentâneo do debate público, não capturando sua evolução ao longo de períodos mais extensos. Também é importante considerar que ferramentas automatizadas de análise textual, como modelos de sentimento e algoritmos de modelagem de tópicos, podem apresentar limitações na interpretação de nuances da linguagem, como ironia ou sarcasmo.

Como direções para pesquisas futuras, sugere-se a realização de estudos comparativos que apliquem metodologias semelhantes a outras redes sociais, permitindo avaliar diferenças entre ecossistemas digitais distintos. Análises longitudinais também poderiam contribuir para compreender a evolução do discurso público ao longo do tempo. Por fim, abordagens qualitativas complementares, como análise de discurso ou entrevistas com usuários, podem ajudar a aprofundar a interpretação dos padrões identificados por métodos computacionais.

## Referências

- Almeida, R. J. A. (2018). Leia - léxico para inferência adaptada. <https://github.com/rafjaa/LeIA>. Acesso em: 6 nov. 2024.
- BBC News Brasil (2024). A rede social que conseguiu um milhão de usuários e bateu recordes de interações após bloqueio do x no brasil. Acesso em: 28 out. 2024.
- Beigi, G. et al. (2016). An overview of sentiment analysis in social media and its applications in disaster relief. In Pedrycz, W. and Chen, S.-M., editors, *Sentiment Analysis and Ontology Engineering*, volume 639 of *Studies in Computational Intelligence*. Springer, Cham.
- Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022.
- Brasil de Fato (2024). Agosto concentra quase metade das queimadas do ano; cerrado é o bioma mais prejudicado. Acesso em: 28 out. 2024.
- Buzelin, A., Bento, P., Aquino, Y., Chagas, A., Dutenhofner, P., Malaquias, S., Estanislau, V., Santana, C., Dayrell, L., Mesquita, L., Locatelli, M., Pappa, G., Jr., W. M., and Almeida, V. (2025). Emerging digital spaces: The case of bluesky's rise during 2024 brazilian elections. In *Anais do XIV Brazilian Workshop on Social Network Analysis and Mining*, pages 269–282, Porto Alegre, RS, Brasil. SBC.

- Corrêa, I. T. and Faria, E. R. d. (2021). An analysis of violence against women based on victims' reports. In *Proceedings of the XVII Brazilian Symposium on Information Systems*, pages 1–8, New York, NY, USA. ACM.
- Datareportal (2024). Digital 2024: Brazil. Acesso em: 28 out. 2024.
- DW Brasil (2024). Queimadas poluíram ar em quase todas as capitais do brasil; veja dados. Acesso em: 28 out. 2024.
- Gaikwad, S. V., Chaugule, A., and Patil, P. (2014). Text mining methods and techniques. *International Journal of Computer Applications*, 85(17).
- Gu, Z. (2022). Complex heatmap visualization. *iMeta*, 1(3):e43.
- Heimerl, F. et al. (2014). Word cloud explorer: Text analytics based on word clouds. In *2014 47th Hawaii International Conference on System Sciences*, pages 1833–1842, Waikoloa, HI, USA.
- Hutto, C. and Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth International Conference on Weblogs and Social Media (ICWSM-14)*, Ann Arbor, MI.
- Jurafsky, D. and Martin, J. H. (2024). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*. Stanford University, University of Colorado at Boulder, 3 edition.
- Kleppmann, M. et al. (2024). Bluesky and the at protocol: Usable decentralized social media. In *Proceedings of the ACM Conext-2024 Workshop on the Decentralization of the Internet (DIN '24)*, pages 1–9. ACM.
- Nadkarni, P. M., Ohno-Machado, L., and Chapman, W. W. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association*, 18(5):544–551.
- Paes, V. J. et al. (2022). Análise de sentimento em tweets relacionados ao desmatamento da floresta amazônica. In *Anais do XI Brazilian Workshop on Social Network Analysis and Mining (BraSNAM)*, pages 1–12, Porto Alegre, RS, Brasil. SBC.
- Sahneh, E. S. et al. (2024). The dawn of decentralized social media: An exploration of bluesky's public opening.
- Zheng, E.-T. et al. (2025). How is science discussed on bluesky?