

Pro/Anti-vaxxers in Brazil: a temporal analysis of COVID vaccination stance in Twitter

André Mediate de Sousa¹ and Karin Becker¹

¹Universidade Federal do Rio Grande do Sul, Brazil
{andremediate, karin.becker}@inf.ufrgs.br}

Abstract. Collective immunization is critical to combat COVID, but a large portion of the population in many countries refuses to be vaccinated despite the availability of vaccines. We developed a temporal analysis of pro/against stances towards COVID vaccination in Brazil using Twitter. We summarized the main topics expressed by pro/anti-vaxxers using BERTopic, a dynamic topic modeling technique, and related them to events in the national scenario. The anti-vaxxers were prevalent throughout 2020, expressing concerns about mandatory vaccination with a strong political bias. The pro-vaxxer movement significantly increased by late 2020 with the begging of immunization and became prevalent in 2021. This group expresses joy and anxiety to get vaccinated and criticisms towards the Federal Government.

CCS Concepts: • **Computing methodologies** → **Machine learning algorithms**.

Keywords: COVID, vaccination, dynamic topic modeling, BERTopic, Twitter, social media

1. INTRODUÇÃO

A pandemia do novo coronavírus apresenta-se como o maior desafio sanitário deste século, tendo custado milhões de vidas, implicando segundo [WHO 2021] a morte de 4,55 milhões de pessoas até o presente momento. A ciência avançou em largos passos sobre o conhecimento do vírus, seus tratamentos, seu contágio e medidas preventivas. Vacinas contra COVID foram desenvolvidas, testadas e aprovadas em tempo célere, sem abandonar requisitos legais de qualidade e segurança e eficácia farmacêutica dispensados a outros medicamentos. Contudo, observa-se mundialmente um alto nível de hesitação que tem prejudicado a cobertura da imunização coletiva [Debus and Tosun 2021].

Os movimentos anti-vacina têm crescido no mundo [Hornsey et al. 2018]. Em relação à COVID, estudos mostram hesitação em países europeus [Debus and Tosun 2021], e alta aceitação em alguns países asiáticos ou de economia média, entre eles o Brasil [Lazarus et al. 2020]. Estes estudos concluem a importância dos governos na definição de políticas públicas e da formação de alianças entre os partidos a fim de aumentar a confiança e o apoio público para atingir níveis mínimos de imunização coletiva.

No Brasil, a pandemia chegou em um cenário político altamente polarizado. Diferenças políticas têm influenciado medidas básicas de prevenção da COVID, como o distanciamento social [Ebeling et al. 2020], o uso obrigatório da máscara ou a utilização de medicamentos sem comprovação científica. A polarização política também tem influenciado a implantação de um Programa Nacional de Imunização (PNI) para COVID, apesar do país ser referência mundial em vacinação em larga-escala [Domingues et al. 2020]. O governo central é criticado por muitos por não ter empreendido esforços concretos ao longo de 2020 para garantir contratos com as indústrias farmacêuticas internacionais para a compra

Research partially supported by CAPES and FAPERGS (project 19/2551-0001862-2).

Copyright©2021. Permission to copy without fee all or part of the material printed in KDMiLe is granted provided that the copies are not made or distributed for commercial advantage, and that notice is given that copying is by permission of the Sociedade Brasileira de Computação.

de vacinas. O Presidente Bolsonaro manifesta-se de forma ambígua quanto à vacinação, o que pode comprometer a percepção da população quanto a sua importância segundo estudos [Lazarus et al. 2020]. Boa parte desta discussão sobre PNI COVID ocorreu em um clima de disputa política entre os futuros candidatos às eleições presidenciais de 2022, governador Dória e Bolsonaro.

A análise da percepção da COVID19 nas redes sociais é uma área de pesquisa muito ativa. A maioria dos trabalhos se concentra na identificação dos tópicos discutidos [Garcia and Berton 2021; Xue et al. 2020; Amara et al. 2021; Tao et al. 2020], tipicamente usando a técnica LDA. Um estudo temporal é proposto em [Sha et al. 2020], relacionando os tópicos com ações governamentais e administrativas. Um modelo para análise de identificação da influência política em posicionamentos pró/anti isolamento social foi proposto em [Ebeling et al. 2020]. Embora estudos mostrem que posicionamentos anti-vacina em geral são mais ligados a teorias conspiratórias e argumentos de liberdade individual [Hornsey et al. 2018; Cossard et al. 2020], no caso da vacinação COVID, estudos apontam para uma correlação com ideologias liberais extremistas [Curiel and Ramírez 2021; Fridman et al. 2021].

A proposta deste trabalho é realizar um estudo temporal, a fim de compreender os argumentos utilizados por brasileiros no Twitter para defender posicionamentos pró e contra a vacinação COVID. A análise foi desenvolvida utilizando aproximadamente de 6 milhões tweets coletados e pré-processados entre Janeiro de 2020 até Maio de 2021, período compreendendo desde o início da pandemia, até a vacinação de 35 milhões de brasileiro com uma dose pelo menos. Utilizamos a técnica de modelagem dinâmica BERTopic [Grootendorst 2020] e relacionamos os posicionamentos encontrados e respectivos argumentos de acordo com eventos ocorridos no período, os quais causam mudanças expressivas no volume de dados de texto que discutem determinado tópico associado a um momento específico [Dou et al. 2012]. Com esta análise, buscamos responder às seguintes perguntas de pesquisa:

Q1: Quais são os principais tópicos subjacentes aos posicionamentos pró/anti vacinação COVID no Brasil?

Q2: Os posicionamentos mudam com o decorrer do tempo?

Q3: Os tópicos subjacentes a cada posicionamento mudam ao longo do tempo?

Nossos resultados apontaram a prevalência do posicionamento pró-vacinação, em consonância com o legado de vacinação em larga escala do país e com o caráter de urgência da pandemia. Caracterizado principalmente por elogios à ciência e gratidão e expectativas sobre vacinação, encontramos igualmente críticas contra o presidente e as ações governamentais que retardaram a vacinação no Brasil. Já o posicionamento anti-vacinação é representado por manifestações de desconfiança quanto à segurança da vacina e defesa da liberdade individual (em oposição à obrigatoriedade), restrições à “vacina chinesa” e disputa política envolvendo Bolsonaro e Dória. O posicionamento anti-vacinação foi prevalente em 2020, principalmente como uma resposta política a eventos no cenário nacional, perdendo força após o início da vacinação no mundo e no Brasil, quando o posicionamento pró-vacinação ganhou expressão e tornou-se dominante.

O restante deste trabalho está estruturado como segue. A Seção 2 apresenta a fundamentação teórica e trabalhos relacionados. A Seção 3 descreve os dados e o método de análise proposto. A Seção 4 apresenta o resultado das análises visando responder às questões de pesquisa. A Seção 5 apresenta as conclusões e trabalhos futuros.

2. TRABALHOS RELACIONADOS

O Twitter vem sendo utilizado para estudo de diferentes fenômenos sociais. No caso do COVID, boa parte dos trabalhos se concentra na identificação dos tópicos discutidos [Garcia and Berton 2021; Xue et al. 2020; Amara et al. 2021; Tao et al. 2020], alguns com foco na influência política [Jiang et al. 2020; Sha et al. 2020]. Em relação ao Brasil, [Ebeling et al. 2020] analisou a influência política em posicionamentos pro/anti isolamento social no Brasil, enquanto que [Garcia and Berton 2021] comparou as percepções entre USA e Brasil. Nestes trabalhos, a técnica prevalente para a modelagem

de tópicos é o Latent Dirichlet Allocation (LDA) [Blei et al. 2003] e suas variações, que resulta em tópicos cujos termos têm maior probabilidade de co-ocorrência.

Mais recentemente foram propostas técnicas baseadas em similaridades de representações vetoriais de texto (embeddings). Derivado do Top2Vec, BERTopic [Grootendorst 2020] é um framework que engloba algoritmos para buscar automaticamente tópicos densos em uma coleção de documentos, assumindo que documentos semanticamente semelhantes formam tópicos. Requer como entrada um *corpus* e um modelo de representação de linguagem pré-treinado (BERT). Após a redução da dimensionalidade, encontra documentos semelhantes usando algum algoritmo de agrupamento (e.g. HDBScan, baseado em densidade, usado neste trabalho). Tem como desvantagem o grande número de clusters resultantes, prejudicando a interpretação de sua semântica, e para tal propõe estratégias de redução, tal como número máximo de clusters ou estratégias de junção. Para interpretação dos tópicos, BERTopic propõe associar aos termos a medida *c*-TF-IDF, que calcula o TF-IDF dos termos por classe. Assim, pode-se avaliar os termos mais importantes de cada tópico em relação aos demais, ou de um subconjunto de documentos de um tópico (por exemplo, em um período de tempo específico).

Neste trabalho contribuímos com uma análise temporal dos tópicos expressos por posicionamentos pro/anti vacinação no contexto brasileiro da COVID usando BERTopic como fundamento à modelagem dinâmica.

3. MATERIAIS E MÉTODOS

Este trabalho realiza um estudo temporal para entender o comportamento de brasileiros ao defender no Twitter posicionamentos pró e contra a vacinação COVID. Utilizamos tweets relacionados à vacinação entre fevereiro/2020 e maio/2021, período que compreende desde o início da pandemia, até a vacinação de 35 milhões de brasileiros com pelo menos uma dose. Para responder às questões de pesquisa elencadas na Seção 1, utilizamos técnica de modelagem dinâmica BERTopic (Q1 e Q3), e a análise da distribuição de volume de tweets no tempo (Q2). O restante desta seção detalha os dados e a modelagem dinâmica dos tópicos.

3.1 Coleta e pré-processamento de tweets

Através da API do Twitter, coletamos aproximadamente 6 milhões de tweets usando com argumento de busca os termos “vacina”, “vacinação” ou ambos, no período de 29/02/2020 até 03/05/2021¹. Sobre os tweets retornados, extraímos as 5 hashtags mais frequentes que expressassem claramente apoio, e 5 hashtags mais frequentes contrárias a campanha da vacinação. Para cada hashtag selecionada, fizemos novas buscas individuais, com o objetivo de selecionar apenas os tweets condizentes com os posicionamentos que elas de fato representam. Com base nas buscas destas hashtags específicas obtivemos um total de 294.182 tweets, divididos em dois posicionamentos:

- *provaaxers*: com um total de 252.071 mil tweets, esse posicionamento representa as pessoas que se manifestam favoravelmente à vacinação por meio das hashtags #vacinaja, #vacinaparatodos, #vemvacina, #vacinasim e #todospelasvacinas;
- *antivaaxers*: com um total de 42.111 mil tweets, esse posicionamento representa as pessoas que se manifestam contrárias à vacinação por meio das hashtags #vacinaobrigatoriano, #eunaovoutomarmvacina, #naovoutomarmvacina, #vacinanao e #vacinaobrigatorianunca;

Os tweets de cada posicionamento passaram por pré-processamento, com normalização de texto e a remoção de pontuação, caracteres especiais, hashtags, urls, e menções a outros usuários. Também

¹Esta data limite visou excluir da análise tweets relacionados com a CPI da COVID, que mudou bastante o foco das manifestações no Twitter envolvendo vacinação.

foram descartados tweets com menos de três termos, e todos os retweets. Após este processo, restaram 163.134 mil tweets para os provaxxers, e 21.810 mil para os antivaxxers.

3.2 Modelagem de tópicos com BERTopic

Para responder à questão Q1, aplicamos o BERTopic aos tweets pré-processados de cada posicionamento. O número de tópicos extraídos está diretamente relacionado à quantidade de documentos, e aos parâmetros. Em nossa análise os parâmetros mais influentes foram: *min_topic_size* e *nr_topic*. Com os parâmetros default, identificamos 2.036 tópicos para os provaxxers e 58 para os antivaxxers. Como é difícil interpretar este grande número de tópicos, usamos alguns recursos oferecidos pelo framework: a) união dos documentos mais similares após treinar o modelo; b) plotagens dos agrupamentos em um plano bidimensional de acordo com a similaridade; c) análise dos tópicos usando termos com alto c-TF-IDF (unigramas e bigramas). Após algumas interações, definimos *nr_topic* = 9 para ambos posicionamentos, conforme a Figura 1. O *min_topic_size* foi 10 e 50 para provaxxers e antivaxxers, respectivamente.

O framework BERTopic inclui recursos para a visualização dos tópicos encontrados e dos termos mais significativos. Também é possível dispor o volume dos documentos de cada tópico ao longo do tempo, e extrair termos significativos em intervalos temporais específicos. Estes recursos foram utilizados para interpretar os tópicos de uma forma em geral, e para analisar a evolução dos tópicos na questão de pesquisa Q3.

O código foi desenvolvido em python no ambiente do Jupyter Notebook com auxílio de bibliotecas como NLTK e pandas. Os notebooks estão disponíveis em um repositório público².

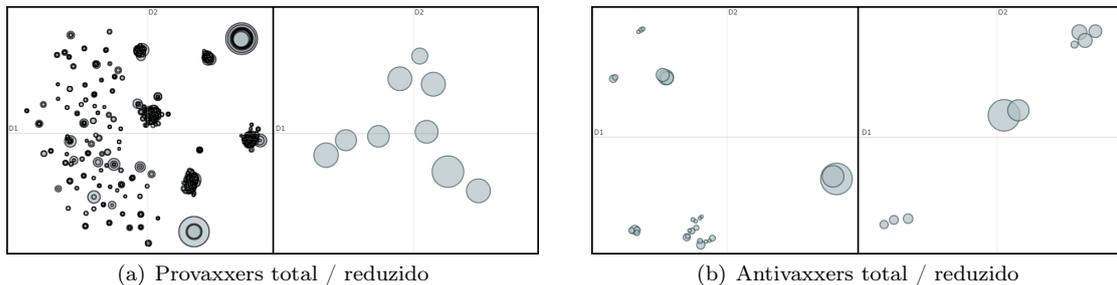


Fig. 1. Redução de tópicos

4. RESULTADOS

4.1 Q1: Quais são os principais tópicos subjacentes aos posicionamentos pró/anti vacinação COVID no Brasil?

A Tabela 1 mostra a representação global dos principais tópicos extraídos conforme apresentado na Seção 3.2, a respectiva quantidade de tweets e os 5 termos (unigrama ou bi-grama) mais representativos. Os tópicos estão ordenados pela quantidade de tweets. Note-se que nem todos os tweets foram encaixados em tópicos, já que em métodos de agrupamento baseado em densidade, áreas não densas são consideradas ruído.

Os provaxxers estão engajados em promover a campanha da vacinação e fazem severas críticas à postura e políticas adotadas pelo governo federal no enfrentamento à pandemia. As expressões “tomar vacinas”, “defenda vacinas”, “dose vacina” e “viva ciência”, presentes nos tópicos 0, 1 e 2, demonstram

²<https://github.com/mediote/twAnalytics>

Table I. Representação global de tópicos por posicionamento

Tópico	Tweets	Palavras - Provaxxers
0	8366	vacina, tomar vacinas, vacina vacina, gado, vacinas
1	5602	defenda vacinas, defenda, vacinas, vacinas mantenha, mantenha informado
2	5874	ciencia, viva ciencia, ciencia viva, cientistas, madison reyes
3	5624	genocida, invadir hospitais, hospitais, impeachment, violentamente contra
4	5126	saude, ministerio saude, ministro saude, saude anos, vacinacao
5	4339	dose, primeira dose, dose vacina, vacinada, segunda dose
6	4595	vacinacao, vacinar, vacinada, vacinados, vacinado
7	5556	brasil, vacinometro, vacinometro brasil, brasileiro, brasileiros
8	2505	tratamento, tratamento precoce, news, fake news, informe morte
Tópico	Tweets	Palavras - Antivaxxers
0	5680	vacina, tomar, tomar vacina, vacina nao, vacinas
1	2567	yuri, coming, is coming, brasil, melody yuri
2	1253	brasil, brasileiros, povo brasileiro, vai
3	1111	vereadores, vereadores senadores, prefeitos, governadores prefeitos, políticos
4	936	china, chinesa, chines, vacina chinesa, chineses
5	522	tag, twitter, hashtag, mascara, levantem tag
6	444	gado, gado nao, vacina, tomar, tomar vacina
7	380	laboratorio, dose, efeitos colaterais, efeitos, ratos
8	282	constrangido submeter, intervenção cirurgica, cirurgica, vida tratamento, tratamento medico

que as argumentações prevalentes são em defesa da campanha da vacinação, elogios à ciência, gratidão e expectativa das pessoas pela vacina. Um número significativo de tweets são voltados a críticas. No tópico 3, os termos “genocida” e “impeachment” mostram forte oposição ao presidente, enquanto que os termos do tópico 4 indicam preocupações com (falta de) ações do ministro e do Ministério da Saúde. Os tópicos 5, 6 e 7 referem-se à contabilização do andamento da vacinação. O tópico 8 demonstra uma preocupação com desinformação e fake news.

O movimento dos antivaxxers é contra a obrigatoriedade da vacina, questiona sua segurança, focando também em disputas políticas. Os argumentos prevalentes são a obrigatoriedade da vacinação (tópico 0) e sua segurança (tópicos 6 e 7). A inconstitucionalidade da vacina obrigatória é o argumento do tópico 8, onde o artigo 15 do Código Civil ("Ninguém pode ser constrangido a submeter-se, com risco de vida, a tratamento médico ou intervenção cirúrgica") é usado para promover a ideia de que a obrigatoriedade da vacina não está amparada na lei. Ao analisar amostras de tweets, observamos que os tópicos 2 e 3 giram em torno disputa política envolvendo Bolsonaro e Dória. Este tema é recorrente junto a argumentações também em outros tópicos. O tópico 4 caracteriza-se por forte resistência à vacina Coronovac, por suas supostas origens chinesas. Há tópicos (1 e 5) focados em gerar engajamento e fortalecer este posicionamento, através de menções usadas para levantar hashtags artificialmente, comportamento este não identificado entre os provaxxers. É interessante notar a conotação do termo “gado”, usado nos dois grupos para se referir as pessoas com argumentos contrários ao posicionamento defendido.

4.2 Q2: Os posicionamentos mudam com o decorrer do tempo?

A Figura 2 distribui os tweets coletados de cada posicionamento ao longo do tempo. Os primeiros a se posicionarem em relação à vacinação de forma mais expressiva foram os antivaxxers em 17/10/2020, data em que é comemorado o Dia Nacional da Vacinação. É possível que este levante seja uma resposta a declarações de João Dória sobre a obrigatoriedade da vacina em São Paulo. Observa-se que até o início de dezembro há pouca atividade dos provaxxers, havendo um leve aumento no fluxo de tweets em torno do dia 04/12/2020, quando alguns países deram os primeiros passos rumo à vacinação de suas populações.

Em 16/12/2020, percebemos um grande aumento nas atividades dos antivaxxers, com um pico de 10.722 mil tweets, indicando uma possível reação ao começo da vacinação em nível global, e um contra-ponto ao início do posicionamento mais intenso dos provaxxers. A partir desse ponto, as manifestações dos provaxxers cresceram e se tornaram prevalentes. Os antivaxxers ainda se manifestaram de forma tímida em abril 2021, em reação a medidas do governo de São Paulo que condicionavam certas atividades da sociedade à vacinação, com pico do dia 21/04/2021.

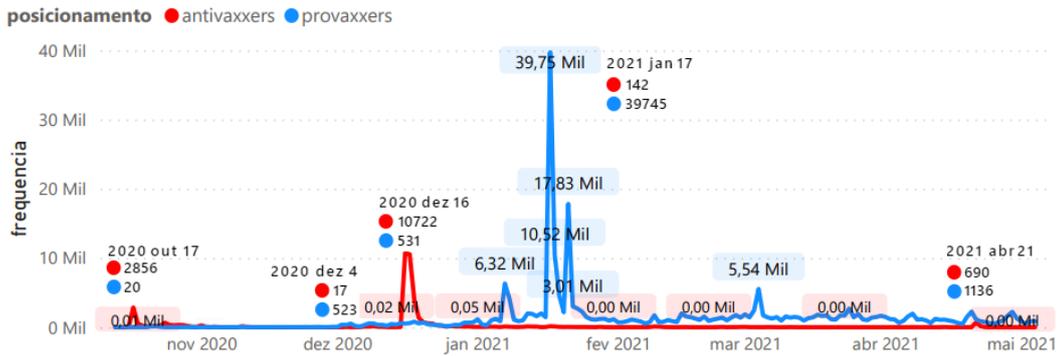


Fig. 2. Evolução dos posicionamentos no tempo

O intervalo de maior atividade dos provaxxers ocorreu entre janeiro e fevereiro de 2021, com um pico de 39.745 mil tweets no dia 17/02/2021. Este período ficou marcado no Brasil por vários acontecimentos, entre eles: a) Discussões políticas entre Bolsonaro e Dória; b) Situação de calamidade em Manaus, devido a uma variante do coronavírus e consequente falta de oxigênio nos hospitais; c) Aprovação do uso emergencial da Coronavac e outras vacinas; e d) Vacinação da primeira pessoa no Brasil, a enfermeira Mônica.

4.3 Q3: Os tópicos subjacentes a cada posicionamento mudam ao longo do tempo?

Ao adicionar uma dimensão temporal à representação global dos principais tópicos extraídos, o BERTopic nos permite criar uma distribuição e visualizar como os tópicos são representados em diferentes momentos no tempo. Com a dimensão temporal, a representação do tópico no momento t pode diferir da representação global do tópico. A Figura 3 mostra a distribuição dos tópicos descritos na Tabela 1 ao longo do tempo para os provaxxers (a) e antivaxxers (b). Para avaliar a evolução dos tópicos ao longo do tempo, descrevemos abaixo o tópico com maior volume de tweets em cada posicionamento. Uma análise completa considerando todos os tópicos está disponível no repositório público.

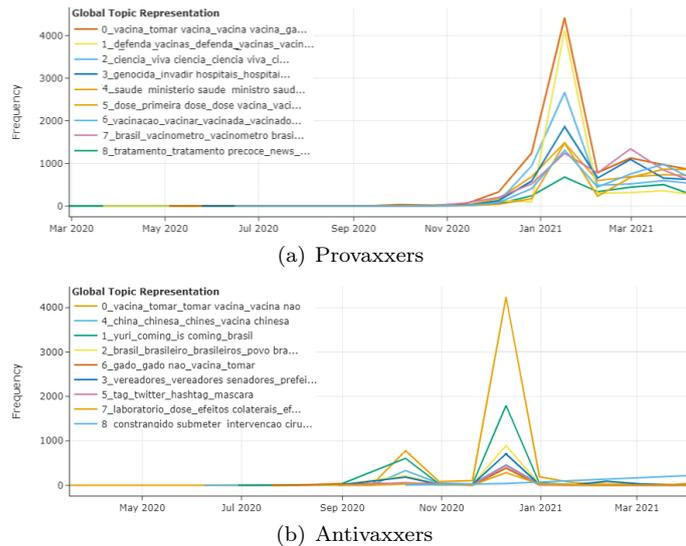


Fig. 3. Distribuição de tópicos

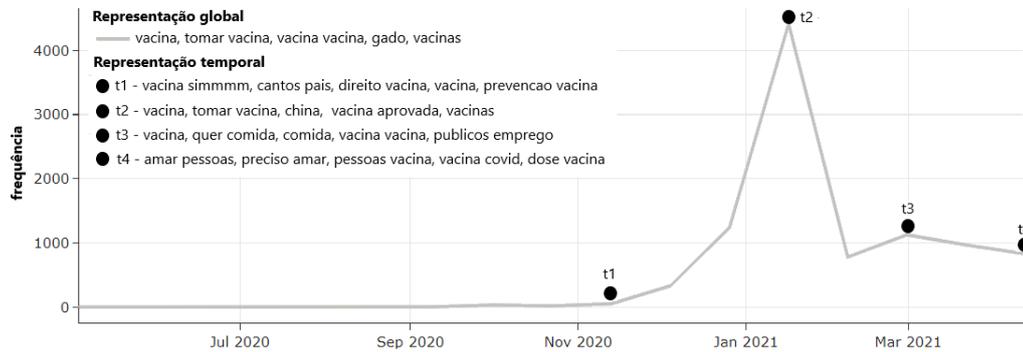


Fig. 4. Evolução do tópico de apoio à vacinação no tempo

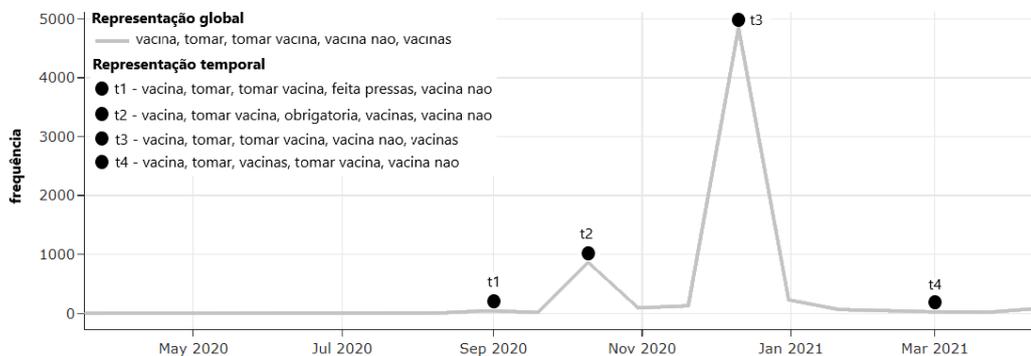


Fig. 5. Evolução do tópico contra a vacinação no tempo

Ao acompanhar a evolução do tópico 0 dos provaxxers na Figura 4, percebemos a seguinte evolução: em $t1$ as pessoas começaram a se manifestar a favor da vacinação; em $t2$, o pico no volume de tweets, indica reações à aprovação do uso emergencial da Coronavac; em $t3$, os termos “comida” e “emprego” indicam uma preocupação com as pessoas que estão em situação de fome e desemprego, causada pelas restrições impostas pelos governantes para conter o avanço do vírus dada a ausência de vacinas; e finalmente em $t4$, já com a vacinação em andamento, notam-se manifestações de gratidão e amor das pessoas por terem sido vacinadas.

A evolução do tópico 0 dos antivaxxers, conforme a Figura 5, mostra a seguinte evolução: em $t1$, preocupações com a segurança e eficácia da vacina, por ter sido desenvolvida em tempo recorde; em $t2$, as reações são contra sua obrigatoriedade; em $t3$ e $t4$ os argumentos são parecidos e se mostram contrários à vacinação em geral.

5. CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho identificamos dois posicionamentos principais sobre vacinação no Brasil (provaxxers e antivaxxers). Ao contrário dos trabalhos relacionados, que empregam LDA, utilizamos a técnica de modelagem BERTopic para compreender seus principais argumentos. Nossa análise temporal incluiu como os posicionamentos foram defendidos ao longo do período observado, e a evolução nos argumentos representados pelos tópicos ao longo do tempo.

Os antivaxxers fizeram manifestações expressivas somente ao longo de 2020, centradas na obrigatoriedade da vacina, insegurança, desconfiança da vacina “chinesa” e disputas políticas. Observou-se um esforço para levantar hashtags representativas deste pensamento artificialmente. Já a manifestação dos provaxxers é cinco vezes maior, mas intensificou-se somente quando do início da vacinação no mundo.

Em 2021 tornou-se prevalente, tendo seu pico associado à tragédia de Manaus e o início da vacinação no Brasil. Os provaxxers se manifestam a favor da ciência, gratidão e expectativas sobre vacinação, e ao mesmo tempo que mostram forte resistência ao presidente e ações governamentais que retardam a vacinação no Brasil. Boa parte dos tweets em ambos os posicionamentos envolvem discussões políticas, principalmente entre Bolsonaro e Dória. Concluímos que os antivaxxers reagem a eventos isolados e não constituem uma campanha efetiva contra a vacinação. O posicionamento provaxxer, é consideravelmente maior, e denota um esforço contínuo de defesa pró-vacinação, confirmando que o Brasil faz parte dos países receptivo à vacinação [Lazarus et al. 2020].

Como trabalhos futuros, contamos evoluir com a expansão desta análise a outros temas relacionados à vacinação (e.g. *someliers* de vacinas), desenvolver um método semi-automático que facilite a avaliação dos tópicos BERTopic em larga escala e ao longo do tempo, usar métodos de aprendizado de máquina para detecção de posicionamentos, identificação de teorias conspiratórias e comparação com movimentos semelhantes em outros países.

REFERENCES

- AMARA, A., HADJ TAIEB, M. A., AND BEN AOUICHA, M. Multilingual topic modeling for tracking COVID-19 trends based on Facebook data analysis. *Applied Intelligence* 51 (5): 3052–3073, 2021.
- BLEI, D. M., NG, A. Y., AND JORDAN, M. I. Latent dirichlet allocation. *J. Mach. Learn. Res.* vol. 3, pp. 993–1022, 2003.
- COSSARD, A., DE FRANCISCI MORALES, G., KALIMERI, K., MEJOVA, Y., PAOLOTTI, D., AND STARNINI, M. Falling into the echo chamber: The italian vaccination debate on twitter. In *Proc. of the Int. AAAI Conference on Web and Social Media*. Vol. 14. pp. 130–140, 2020.
- CURIEL, R. P. AND RAMÍREZ, H. G. Vaccination strategies against COVID-19 and the diffusion of anti-vaccination views. *Scientific Reports* 11 (1): 6626, 2021.
- DEBUS, M. AND TOSUN, J. Political ideology and vaccination willingness: implications for policy design. *Policy sciences*, 2021.
- DOMINGUES, C. M. A. S., MARANHÃO, A. G. K., TEIXEIRA, A. M., FANTINATO, F. F. S., AND DOMINGUES, R. A. S. The Brazilian National Immunization Program: 46 years of achievements and challenges. *Caderno de Saúde Pública* vol. 36, 2020.
- DOU, W., WANG, X., RIBARSKY, W., AND ZHOU, M. Event detection in social media data. In *IEEE VisWeek Workshop on Interactive Visual Text Analytics-Task Driven Analytics of Social Media Content*. pp. 971–980, 2012.
- EBELING, R., SÁENZ, C. C., NOBRE, J. C., AND BECKER, K. Quarenteners vs. cloroquiners: a framework to analyze the effect of political polarization on social distance stances. In *Anais do VIII Symposium on Knowledge Discovery, Mining and Learning*. SBC, pp. 89–96, 2020.
- FRIDMAN, A., GERSHON, R., AND GNEEZY, A. COVID-19 and vaccine hesitancy: A longitudinal study. *PLoS ONE* 14 (4): e0250123, 2021.
- GARCIA, K. AND BERTON, L. Topic detection and sentiment analysis in twitter content related to covid-19 from brazil and the usa. *Applied Soft Computing* vol. 101, pp. 107057, 2021.
- GROOTENDORST, M. Bertopic: Leveraging bert and c-tf-idf to create easily interpretable topics. vol. <https://doi.org/10.5281/zenodo.4381785>, 2020.
- HORNSEY, M. J., HARRIS, E. A., AND FIELDING, K. S. The psychological roots of anti-vaccination attitudes: A 24-nation investigation. *Health Psychology* 37 (4): 307–315, 2018.
- JIANG, J., CHEN, E., YAN, S., LERMAN, K., AND FERRARA, E. Political polarization drives online conversations about covid-19 in the united states. *Human Behavior and Emerging Technologies* 2 (3): 200–211, 2020.
- LAZARUS, J. V., RATZAN, S. C., PALAYEW, A., GOSTIN, L. O., LARSON, H. J., RABIN, K., KIMBALL, S., AND ELMOHANDES, A. A global survey of potential acceptance of a COVID-19 vaccine. *Nature Medicine*. 10.1038/s41591-020-1124-9, 2020.
- SHA, H., HASAN, M. A., MOHLER, G. O., AND BRANTINGHAM, P. J. Dynamic topic modeling of the COVID-19 twitter narrative among U.S. governors and cabinet executives. *CoRR* vol. abs/2004.11692, 2020.
- TAO, G., MIAO, Y., AND NG, S. COVID-19 topic modeling and visualization. In *24th Intl. Conf. on Information Visualisation (IV)*. IEEE, pp. 734–739, 2020.
- WHO. Who coronavirus (covid-19) dashboard., 2021.
- XUE, J., CHEN, J., CHEN, C., ZHENG, C., LI, S., AND ZHU, T. Public discourse and sentiment during the covid 19 pandemic: Using latent dirichlet allocation for topic modeling on twitter. *PLOS ONE* 15 (9): 1–12, 09, 2020.