9^{th} International Conference on Computational Social Science IC^2S^2
July 17-20, 2023, Copenhagen, Denmark

A Crowdsourced Map of YouTube Recommendations during the Brazilian 2022 Elections

Keywords: YouTube, network science, crowdsourcing, elections, Brazil

Extended Abstract

Internet access plays a key role in shaping public opinion in Brazil. Social networks have been consistently used as information sources, and traditional press sources face distrust. According to a survey published by *We Are Social*, Brazilians are online about 10 hours a day, being the second country in the daily internet ranking, after the Philippines. Also according to the organization, 75% of the Brazilian population are online, and 70.3% of the population actively social media [4], and according to the Reuters Digital News Report, since 2020 social networks (currently, 64%) have been ahead of television (currrently, 55%) as a source of information in Brazil. [2]

Much like for many recent democratic elections and other transitions of power, the run-up to the Brazilian general elections was rife with claims of illegitimacy of the electoral process, as well as misleading content regarding multiple aspects of Brazilian politics, such as the potential role of the congress, of Brazil's Supreme Electoral Court, and of the Brazilian army in a potential presidential coup. Concerned with the spread of online misinformation during the electoral period, Brazil's Supreme Electoral Court signed agreements with WhatsApp, Twitter, Kwai and other major platforms to fight misleading content about the electoral processes of the 2022 general elections. Each company has signed a memorandum of understanding outlining goals to be accomplished. It is, however, notoriously difficult to track the effectiveness of such agreements, and to keep platforms accountable – which calls for more studies assessing the content shared in (and often, recommended by) these platforms.

In this project, we focus our analysis on the videos shared and recommended on YouTube during the Brazilian 2022 elections. This research contributes to a growing literature on the sociotechnical nature of recommender systems, and particularly on the examination of harmful content on YouTube [3]. Approaches trying to capture YouTube recommendations tend to rely on 'seed' channels and accounts, or on browsing with anonymous accounts [3], to different degrees of success. Here, we adopt an innovative methodology, relying on the voluntary tagging of YouTube videos by volunteer internet users, through a browser plug-in developed for this project which anonymous internet volunteers used to tag any videos they considered to be political. Simultaneously, we asked volunteers to answer an anonymous survey, asking demographic information such as age and ethnicity, as well as questions about what topics were considered political by each volunteer. Altogether, this project was designed around three central questions: first, understanding what Brazilians consider political on YouTube, second, assessing how the videos flagged by volunteers portray the 2022 elections, and third, making a "map" of the YouTube recommendation landscape during the Brazilian Elections, and analysing the network emerging from YouTube recommendations coming from the flagged videos.

With the help of 65 volunteers between October 10th 2022 and November 30th 2022, this project generated a total of 1,281 tagged videos and 9,519 recommended videos, from a total of 2,195 YouTube channels, covering a considerable period between the first and second round of the Brazilian elections, as well as the period immediately after the election. For every video, we then collected video metadata, including title, description, channel ID, view counts and like

counts, as well as any videos recommended by YouTube on the moment the video was tagged, along with their respective metadata. In addition, our team developed a script to collect the full transcription of these videos in text format.

To analyse the video transcripts, we used a Latent Dirichlet Allocation topic model, set to 70 topics. This analysis shows that political content on the platform was highly diverse, and included topics such as corruption and social justice, as well as a majority of videos on the actual electoral process (e.g. ballots, vote counting, parties). We also find a large fraction of videos from far-right YouTube channels, promoting anti-democratic discourse, questioning electoral procedures and calling for a coup d'état to keep Jair Bolsonaro (who was the president at the time) in power. Naturally, the dataset produced in this paper is biased by design, since it depended on the individual biases of the volunteers who participated in the crowdsourced data collection – and indeed, we find that most volunteers reported identified as learning towards left-wing politics, which could explain part of the overrepresentation of far-right videos.

In our analysis of this data, we present what types of videos Brazilians consider political on YouTube, which events were highlighted during the 2022 election, and what types of videos are recommended by YouTube based on the content tagged by volunteers. We also identify the political leaning of channels with tagged videos, and assess whether the recommendations of tagged videos reflect the same theme and opinion of the videos recommending them. Finally, from a sample of videos, we analyse the presence of disinformation, political violence, and anti-democratic speech towards political figures and institutions, as well as hate speech targeting gender, sexual orientation, race, color, and ethnicity.

Finally, we examine the network of YouTube recommendations built from the crowdsourced dataset. This network is shown in Figure 1. In the figure, every node represents a YouTube video, edges represent video recommendations, and colours represent communities in the network, identified using the Louvain method for community detection [1]. By grouping videos that tend to recommend each other into communities, we are able to identify "rabbit holes" where users are likely to end up after following a number of recommendations. We then describe the content and political leaning present in each network, ultimately painting a sketch of a "map of YouTube during the Brazilian 2022 elections".

Our findings have important implications for policymakers and researchers interested in understanding the role of social media in modern democracies. They suggest that platforms such as YouTube have the potential to shape political discourse in significant ways, and that more research is needed to understand the mechanisms behind this influence. Finally, our project provides a methodological contribution, by aggregating YouTube crowdsourced recommendation data from a wide range of volunteers. By developing and applying innovative research methods to pressing problems such as the role of online platforms in modern democracy, we hope to push the envelope in how we understand the digital world.

References

[1] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.

[2] Nic Newman, Richard Fletcher, Craig T. Robertson, Kirsten Eddy, and Rasmus Kleis Nielsen. Digital News Report, Reuters Institute & University of Oxford, 2022.

[3] Manoel Horta Ribeiro, Raphael Ottoni, Robert West, Virgílio AF Almeida, and Wagner Meira Jr. Auditing radicalization pathways on youtube. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pages 131–141, 2020.

[4] We Are Social. Digital 2022 Brazil Report, (February 2022).

Figure 1: Largest component of the co-recommendation network. In this figure, every node represents a YouTube video, edges represent video recommendations, and colours represent communities in the network, identified using the Louvain method for community detection [1].