

Characterizing Reactions and Comments Associated with News on Facebook

Samuel S. Guimarães¹, Fabrício Benevenuto (Orientador)¹

¹Departamento de Ciência da Computação
Univerisade Federal de Minas Gerais (UFMG)
Belo Horizonte, Brazil

samuelsg@ufmg.br, fabricio@dcc.ufmg.br

Abstract. *News consumption is increasingly done on social media websites. In this environment, all types of entities and people present themselves as news sources. These new outlets might focus on specific audiences, and some exhibit the news less objectively. Facebook is one of these platforms, which categorizes an extensive group of pages as a kind of news media. It is crucial to characterize all pages that disseminate information in this ecosystem to analyze this phenomenon. Our main objective is to create an in-depth diagnostic of news stories and opinions, focusing on Brazilian Facebook. Our contributions are: (i) a new method to measure the political bias of Facebook pages on a given country, and (ii) a detailed characterization of a comprehensive sample of these pages.*

1. Introduction

Today many people use social media websites to satisfy their daily news diet. According to a recent survey from the Pew Research Center [Shearer and Matsa 2018], nearly 68% of U.S. adults mainly get informed about current events on social media websites. For Brazilians, a similar study found this statistic to be 66%. It was also found that social media has passed print newspapers as a news source in Brazil since 2014 [Newman et al. 2019]. This new behavior changed how news is consumed and produced, lowering the barrier to entry, and thus promoting independent journalism, such as citizen journalism from the Arab spring [Russell 2011].

A recent work indirectly measured this shift, discovering 20,448 self-reported pages of U.S. news outlets located on Facebook [Ribeiro et al. 2018]. This independent journalism, also called **alternative news media**, still generates some debates about firm definitions, sometimes challenging the very definition of journalism [Newman 2011]. One article conceptualized some key dimensions where this journalism differs from the so-called **mainstream media**, which are *the producers, the content, the media organizations* formed, and *the media systems* where it lives [Holt et al. 2019].

In these dimensions proposed, **public figures** and political entities can be seen as a type of *producer*. Generally, they are a focus for the media as the subject of news stories. However, they can also use their official and personal accounts on social media to present information to the public in a more direct way, possibly replacing traditional news. The use of Twitter by Former U.S. President Donald Trump was analyzed in a similar light [Wells et al. 2020]. These alternative media outlets are increasing in number and power recently, sometimes with their **political bias** being the source of popularity [Kelkar 2019, Newman et al. 2019].

Therefore, understanding alternative media pages, their characteristics, and audiences can help us find trends in their effects in different communities, assessing their societal impacts. In Brazil, several alternative media outlets have emerged during both left and right-wing governments, mostly mixing activism and reporting, raising questions about their political accountability and compromise with the truth. Specifically, in 2018, Facebook dismantled what was called a coordinated disinformation web of 196 pages and 87 personal profiles [Haynes 2018]. The owners of these fake accounts hid the nature and origin of their publications, intending to propagate division and disinformation. Credible independent journalists can create these types of pages, but political groups, like in the cases presented, can start their outlets to shape public opinion or attack political enemies [Holt 2018, Kelkar 2019]. In this scenario, the mapping of these pages and their political bias can be crucial to increasing the transparency of this media environment.

We aim at finding possible solutions to that problem, with our main contribution in the dissertation being a new methodology to identify Facebook pages as mainstream media, alternative media, and public figures, also finding their political ideology bias, focusing on a given country. Along with finding the ideology of the pages, we make an overview of the media ecosystem through a characterization of the audience of these pages.

Furthermore, in this new media environment, people also use social media to influence their friends, sharing news stories they consider correct [Weeks and Holbert 2013]. Beyond that, comments sections are also used as a way to interact and discuss with other people about posts published on different pages. Some people are more often inclined to leave online comments on online news, including news websites. Among these people, 77.9% of them comment specifically on social media [Stroud et al. 2016]. A notable proportion of the comments on news posts can be considered toxic [Su et al. 2018, Hille and Bakker 2014]. According to a study [Reis et al. 2015], a similar trend happened with comments on news websites, which were negative independent of the headline.

Addressing that, our second contribution is a diagnostic about the comments associated with the news posts shared on Facebook, focusing on measuring their toxicity levels. To complement this analysis, we also examine Facebook reactions, which include *Love* and *Haha*, which are more emotional-specific versions of regular likes. As users can apply them to express more clear emotions, we assess their use to display negativity and anger, comparing these reactions with the previous toxicity score of the comments of each news post.

As a case study of the use of our methodology, and exploration of a media environment, we analyze the ecosystem of news inside Facebook, focusing on Brazil, which has a history of interaction between alternative media and politics [Haynes 2018]. Studying this ecosystem in Brazil is limited to few pieces of research, with most of them being case studies of existing groups of interest [Moretto and Ortellado 2018, Newman et al. 2019], even though Facebook is the most used platform for news reading in Brazil.

2. Goals

The goal of this project is to help to fill this research gap, presenting two contributions:

1. Creation and validation of a methodology to identify and measure the political bias of Facebook pages for a given country;

2. An in-depth characterization of ideology bias, public demographics, reactions in posts, and comment toxicity of a large sample of the three actors cited: mainstream media, alternative media, and public figures, in Facebook pages from Brazil.

3. Related works and Theme Overview

Understanding alternative media pages and their audience's interactions between themselves and with the content can help us find trends in how they affect different communities and assess their societal impacts. This examination is related to one of the challenges of the Open World and its implications in Information Systems, as described in [Araujo et al. 2017], bringing more transparency to the relationship of politics and media in social media platforms. More specifically, one of the main challenges we face is related to how subjective are the concepts of alternative news media, political bias, and the degrees of online hate.

3.1. Alternative Media

To address that, we first briefly describe some cases of alternative outlets changing the news and political landscape, giving a few examples of what we would consider alternative media. Recently, the impact of alternative media has been examined [Kelkar 2019, Holt 2018]. The ecosystem of alternative right-wing outlets in the USA [Kelkar 2019] creates an asymmetry in the perception of world affairs, being particularly relevant in the U.S. election of 2016. During this election, Russian troll farms spread propaganda on social media through fake news and fear-mongering, resulting in more hate on the platform. In Sweden and Germany, right-wing movements share the anti-system rhetoric with the alternative media to make their audiences participate in voting. In those cases, Immigration-critical alternative media gained a significant reach, and they coined a German word for the traditional news outlets, "Lügenpresse", or the Lying Press [Holt 2018]. These examples across the countries show that politicians and political movements purposely use alternative media to amplify their message, influencing their followers to distrust the established media landscape and trust the new form of journalism. Sometimes these alternative media outlets return significant media support to the politicians. We use the definition cited before [Holt et al. 2019], especially applying the dimensions of *producers* and *media organizations*.

3.2. Toxicity and related metrics

We also detail the contrast between toxicity and other types of hateful online behavior like online harassment, offensive content, hate speech, and incivility. Not surprisingly, many recent research efforts have attempted to operationalize the concept of hate speech, and other overlapping behaviors, by defining them in terms of measurable factors to be able to identify and counter them. The key challenge for that is that, even in our society, there is not a unanimously accepted definition of hate speech. So, the strategy to operationalize hate speech has been usually a data-driven approach in which humans were used for labeling a predefined set of messages as hate speech or non-hate speech.

A survey [Fortuna and Nunes 2018] noted that these studies used similar, but not identical, descriptions of what is hateful and how it is different from offensive content. In its survey, they found that the definition generally covers the presence of some aspects:

1. Hate speech has specific targets;

2. Hate speech is to incite violence or hate;
3. Hate speech is to attack or diminish;
4. Humour can be used to some degree.

Another effort [Ross et al. 2016] showed that even human annotators provided with hate speech definitions do not lead to highly reliable corpora, indicating that this task is difficult even for humans. More recently, the Jigsaw Group launched the Perspective API [Jigsaw 2017]. This API uses machine learning models to score the perceived impact a comment might have on a conversation, creating a toxicity index to measure how toxic a message can be perceived by a user. The concept of toxicity presented has an overlap with many of the reviewed subdivisions. However, it focuses on the fact that these types of messages disrupt the online environment, possibly ending the interactions.

For our purposes, the concept of toxic content is the most comprehensive type of hate that did not include actions that are not as harmful. We used the Perspective API to measure toxicity.

When we compare our proposed analysis with previous works, we see few articles attempting to create or use databases of Facebook news comments or post reactions like we intend to use. One of them [Khan and Chang 2019] built a dataset from Facebook data, collecting posts from the Amazon Facebook page for five years, using the Graph API. The authors used different neural networks to predict the number of distinct interactions the posts would receive. Meanwhile, another research [Kolhatkar and Taboada 2017] created a dataset focusing on the idea of the constructiveness in news comments, which was related to constructing a civil dialogue, using deep learning to evaluate the result. After building the dataset, they examined the relationship between toxicity and constructiveness using the Perspective API. Another article [de Pelle and Moreira 2017] created a dataset called *OffComBr*, which focused on one of the biggest news websites in Brazil, called G1¹. Their idea was to create a database that could be used in machine learning, labeling 1,250 comments from the website commenting section.

In all cases, the number of pages, the size of the content analyzed, or the objective was distinct from ours. Therefore, to the best of our knowledge, there is no previous work that compared the toxicity inside alternative media, mainstream media, and public figures Facebook pages in the current context.

3.3. Analysis of posts reactions

Meanwhile, dealing with the post reactions, a few articles were also relevant to our project. First, there is an analysis of 21,000 Facebook posts from 15 media pages from four different countries employed to test the relation of emojis in comments to reaction profiles of pages [Tian et al. 2017]. Examining approximately 57 million reactions and 8 million comments, the authors used emoji sentiment scores from previous work to calculate the message sentiment. They created reaction profiles of the pages using the proportion of each reaction in all the posts received. After applying clustering algorithms, they conclude that Facebook comments and the reactions to posts may work as an indication of user emotional state and possible attitudes.

From a different perspective, one article deals with the problem of prediction involving reactions, trying to evaluate alternatives for predicting these reactions to posts

¹<https://g1.globo.com/>

on supermarket chains pages [Moers et al. 2018]. Indeed, their final model was able to predict the reaction distribution with a Mean Square Error of 0.135, using their dataset of 8,103 posts on the customer service page of 12 US/UK big supermarket chains. In their analysis, they found a problem in using Facebook like that was initially the only reaction, with the early results showing that results become increasingly meaningless when including likes due to its large predominance in the data. A similar analysis leads us to remove the likes in our study too.

Finally, [Basile et al. 2017] also tried to predict results using the data from Facebook reactions, using them as proxies for predicting news controversy. They say these controversies are “situations where, even after lengthy interactions, opinions of the involved participants tend to remain unchanged and become more and more polarized towards extreme values”. Based on this definition, they hypothesize that a news story has a higher controversy if it has a large number of reactions divided into two or more of the different emotions represented, measuring this by the use of the entropy of these reactions. In their results, they use a simple type of regression to prove the ability to predict the controversy of a text by using the cited definition. Although they used a small and local dataset, predominantly with content from Italy, we also used this concept of controversy to compare the reactions to the toxicity received by posts.

3.4. Political Bias

Finally, previous work in journalism has discussed different kinds of biases present in the media, with diverse definitions for media bias and its more specific forms. Looking at these definitions, we see that some subdivide bias by how it affects the news articles [D’Alessio and Allen 2000, Mullainathan and Shleifer 2002, Entman 2007], while others focus on the objectives of the insertion of the biased perspectives [Druckman and Parkin 2005]. As the intention is difficult to measure, it is not useful, for our purposes, to use a definition that focuses on the objectives of bias. However, to evaluate the diverse ways that news outlets transform the content, we need to analyze the content produced. We expect that the amount of data examined might interfere with the accuracy of the calculated bias. Processing more information means that it is more computationally costly.

Because of that, our approach is closer to the work of [Mullainathan and Shleifer 2002], which tries to quantify the way the content was altered by defining a type of bias they called editorial slant. This slant is defined as “the quantity and tone of a newspaper candidate coverage as influenced by its editorial position”, choosing to investigate the effects on the people who use the different news outlets in their news diet.

As our main contribution is a new method to calculate political bias, we introduce four studies to be compared with our methodology. First, the Pew Research Center [Mitchell 2014] analyzed the audiences from 36 news outlets by interviewing 2,901 people and asking them which media outlets they knew, the ones they read, their ideological self-identification, and how much they trust in media. This interview allowed them to determine how different political biases affected the news perception.

Another research [Budak et al. 2016] employed content analysis to measure the overall ideological bias of 15 major U.S. news outlets by compiling 803,146 articles pu-

blished over an entire year. They later classified 10,502 specific political stories using 749 human judges. The overall value was measured using these articles, and the results showed little divergence in coverage by outlets of different bias, except in scandals.

Like our strategy, [Bakshy et al. 2015] also used Facebook data to calculate the ideology of 500 websites shared on that social media. They classified 226,000 URLs using the political affiliations from 10.1 million users that declared their bias from an initial seven million shared over six months. Their results showed that social media content could cross ideological lines and reach people from opposite perspectives.

Lastly, [Ribeiro et al. 2018] used the Facebook Marketing API² to gather data on the proportion of users inside different sections of the political spectrum, then calculated a bias score for 20,448 U.S. media outlets. Their result provides a demographic analysis of the American audience, especially using the division of the users in Very Conservative, Conservative, Moderate, Liberal, and Very Liberal, to generate the bias metric. However, their approach cannot be exploited outside of the U.S. because the political leaning of the page audience is not readily available.

4. Methodology

To execute our intended contributions, we start by finding all the relevant Facebook pages to analyze in the media ecosystem. We separate these pages into a preliminary dataset and two versions of an extended dataset, each for a separate analysis.

Our first strategy was exploiting the Marketing API from Facebook to find pages based on a predetermined list of interests related to politics, based on the same API by related work [Ribeiro et al. 2018], thus creating the preliminary set. Subsequently, we created a new approach where we used a snowball process that finds pages based on an initial list of media pages, searching for new pages related to Facebook *interests* associated with the starting ones. These interests are an extensive set of topics that Facebook infers from the users' engagement, generally used in the creation and management of ads on Facebook.

Using this methodology, we created the extended dataset. One previous work that studied the media division based on political sides was the source for our starting list of pages [Moretto and Ortellado 2018]. Finally, we used more comprehensive page types to create our second extended dataset, executing the same steps again. Then, we focused on analyzing the pages found on each iteration of this snowball process. We contrast the differences and show how we filtered these pages to reach the final datasets. In total, our preliminary set found 63 pages, the first version of the extended one was composed of 156, and, finally, the second version encountered 767 pages.

From this result, we divide each case into the three mentioned types. Initially using categories given by Facebook, we were able to separate pages from **public figures**. For media pages, we classify them as *mainstream media* if they are registered in any Brazilian official press organizations, like the National Association of Journals (ANJ). If the media outlet is not registered, we consider it as *alternative media*.

Our next step is the method to measure political bias and our analysis of the pages' characterization. In the proposed methodology, we infer the ideological leaning from an

²developers.facebook.com/docs/marketing-apis

initial group of pages that self-reported their ideology, using one graph created by our snowball process for finding pages, propagating the political stance to other related pages using their relationship on Facebook. We then create an affinity score based on one tool related to the Marketing API called *Audience Insights* and later use it to compute the political leaning.

For example, consider two pages, which we will call a and b , associated with Facebook interests in those pages. When we search the tool using the interest in a , it presents a list of related pages, which includes b in this example. For all the pages found, including b , the *Audience Insights* provides three metrics: Monthly Active People (MAP), Audience, and Affinity score, shown in Figure 1. In this example, MAP is the size of the group of monthly active users on page b . The audience counts the active users on page b also interested on page a . Then, the affinity score given by Facebook measures how likely a user interested in a is to like page b compared to a random user.

Page	Relevance	Audience	Facebook	Affinity
O Globo			4.4m	
Época			1.9m	135x
Revista ISTOÉ	3	835.2K	1.9m	134x
GloboNews	4	797.8K	1.9m	127x

Number of MAP (Monthly Active People) for the page

The Pages that are most likely to be relevant to your audience based on affinity, Page size and the number of people in your audience who already like that Page.

How likely your audience is to like a given Page compared to everyone on Facebook.

Figura 1. Facebook Audience Insights tool.

However, this affinity score from Facebook has one problem. While the affinity score allows us to compare related pages found using the same interest in the snowball process, it is not straightforward to compare pages found by searching different interests because it is not normalized. We, therefore, propose a new normalized affinity score. Using this new metric, we construct a graph and, by applying a semi-supervised learning method, we propagate the ideology to all pages, using the following steps:

1. We calculate our new affinity score For each page found as interest using:

$$\mathcal{A}_{a,b} = \left(\frac{\text{Audience}_{a,b} + \text{Audience}_{b,a}}{\text{MAP}_a + \text{MAP}_b} \right)$$

2. We build an undirected weighted graph with pages as nodes and edges established when one page was found related to the other on the Audience Insights, with edge weight equal to the complement of affinity: $w(u,v) = 1 - \mathcal{A}_{u,v}$;
3. We employ the Floyd-Warshall algorithm [Floyd 1962] to calculate the distance between all pairs of nodes of the graph;
4. We identify which pages from the 13 selected categories can be confirmed as left-wing or right-wing and then label accordingly;

5. We apply one graph-based SSL method to classify the remaining pages as left or right, using the graph as a parameter;
6. We define the political bias as the likelihood of a page being classified as right-wing minus the probability of the page being left-wing, giving a value between -1 (left) to 1 (right). As we use cross-validation, we average this bias considering all folds.

We experiment with some existing graph-based SSL algorithms for step (5) against a supervised learning baseline. They are classic label propagation (LP) [Zhou et al. 2003], label propagation with smooth function classes (Smooth LP) [Zhu et al. 2003], and spectral graph transducer (SGT) [Joachims 2003]. As a baseline method, we apply the K-nearest neighbors algorithm (KNN) [Cover and Hart 1967] using only the known part of the graph in supervised learning. We perform 10-fold cross-validation and report the area under the ROC curve (AUC) for all instances, using this metric and cross-validation scheme to perform a grid search on the best hyper-parameters.

To better describe the ecosystem of media pages, beyond the bias measure by the method above, we also assess the toxicity of the posts and comments on the Facebook pages in our dataset. For that, we use Google Perspective API³, which provides different models to identify abusive textual content. The toxicity score measures how likely it is for a comment to be a “rude, disrespectful, or unreasonable comment that is likely to make people leave a discussion”.

Thus given a text, the models return a probability of that content being toxic or an attack, which we call their scores. When this text is confusing or misspelled, the model returns no score, as it does not recognize the text as comments. The API could not estimate the toxicity score for 8.17% of the posts and 9.13% of the comments in our dataset. Those texts contained misspelling and very informal writing of those texts. We evaluate the results from this API using a sample of 2.000 comments manually labeled by three volunteers. They used this definition, making possible the contrast between the method and human analysis, which can, for example, indicate the efficacy of separating toxic content from satire. We find that the Perspective API can be as good as a human.

Lastly, as another characterization of our found pages, we examine the reactions given to the posts by the users using the Facebook CrowdTangle platform⁴ tool. In this platform, we can access a very large dataset of historical data from Facebook and Instagram with detailed data. As our focus is on Facebook, we only deal with data from that platform and do not gather data from other available. For each page post, we collected each of the five possible Facebook reactions (Love, Haha, Wow, Sad and Angry) normalized by their proportion in the total, creating a profile by page. With these profiles, we did a clustering of page profiles to analyze possible patterns between the political bias and page types. We used the K-means algorithm for clustering. We evaluate these reactions ignoring Likes because it is the immense majority of all interactions and also ignoring Care because it was not available until 2019 [Moers et al. 2018].

³<https://www.perspectiveapi.com>

⁴<https://www.crowdtangle.com>

5. Results

To find the best learning algorithm to be used in our method for finding political leaning, we tested four different alternatives and compared them with the four datasets with various ways of measuring bias. After tuning the hyper-parameters and testing the algorithms, we chose the Smooth LP with σ equal to 0.55 and the RBF kernel as the SSL algorithm from step (5) of our political bias method. The best RBF kernel found was not the standard implementation available on the scikit-learn library but our version that used our affinity score directly instead of calculating the euclidean distance on it.

Using this result, we find a high correlation of our methodology with most of them, especially with the ones that also used Facebook data [Ribeiro et al. 2018, Bakshy et al. 2015], as seen in Figure 2. Figure 3 shows the graph of the found pages in the snowball process with the political bias divided into three political classes.

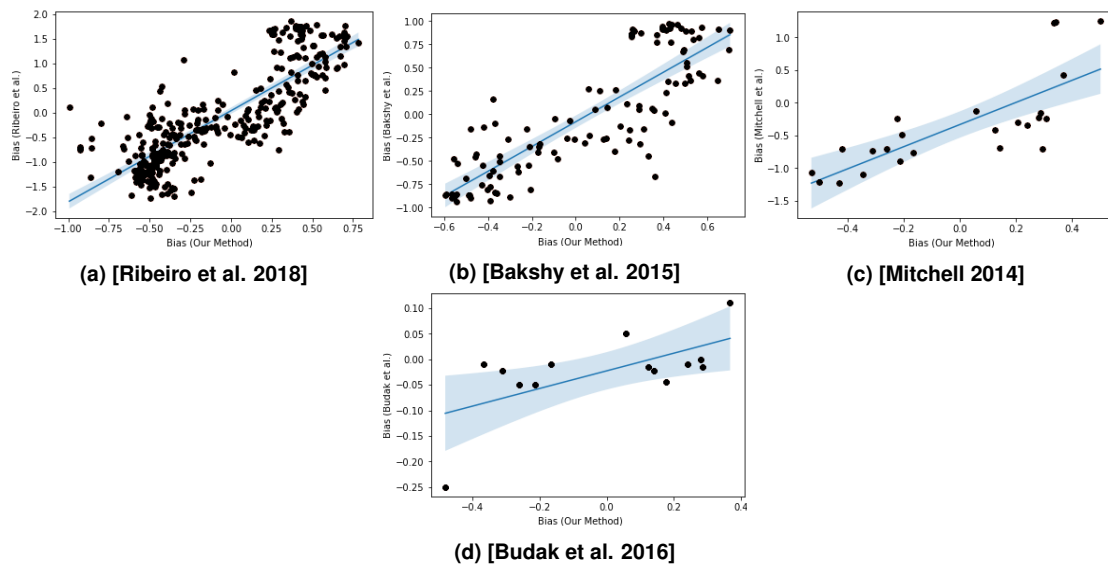


Figure 2. Comparison of our method to the baseline political leanings.

After calculating the political bias, we further characterize the pages by examining the age, gender, relationship status, and education of their audiences. In this analysis, we found that public figures had an older and more male audience, similar to trends present on the finding of other works [Moretto and Ortellado 2018], with the addition of an evaluation of relationships. There was also a mild correlation between ideology and the self-reported education level, in which pages further to the left had an audience that described themselves as more educated. This form of difference in culture and behavior was also found in other researches [Shi et al. 2017], increasing the evidence of the accuracy of our method.

Later, we show the results from using the Perspective API to measured the toxicity of Facebook posts and comments obtained from the pages. We show how both toxic posts, and comments, are related, comparing how these texts are affected by the pages' characteristics from where they were collected. We found a series of relevant trends. In general, comments that reply to other messages are more toxic than the original comment, and there are more of these toxic texts on pages with more toxic publications, even if the toxicity of the comments and the post itself were not highly correlated.



Figura 3. Graph of one dataset of pages with political bias.

Besides that, news media pages received more toxicity than public figures, but when a post mentions these public figures, we find an increase in the proportion of toxic messages.

Figure 4a illustrates this difference, showing that it extends to the entire distribution, with toxicity being statistically lower for public figures. Figure 4b similarly shows that in terms of the proportion of toxic comments (that is, comments whose toxicity is higher than 0.8), the averages of the percentage of these comments in publications of public figures are lower than that of publications in the media with averages of 9.58% and 16.53%, respectively. Figure 4c presents an indication that the traditional media has a distribution of the toxicity of the posts with less toxicity, showing that the higher proportion of toxic comments does not come from many posts with toxic text. Among the media pages, all but one have more than 10% of toxic comments, while half of the pages of public figures do not show such a level of toxicity in their messages.

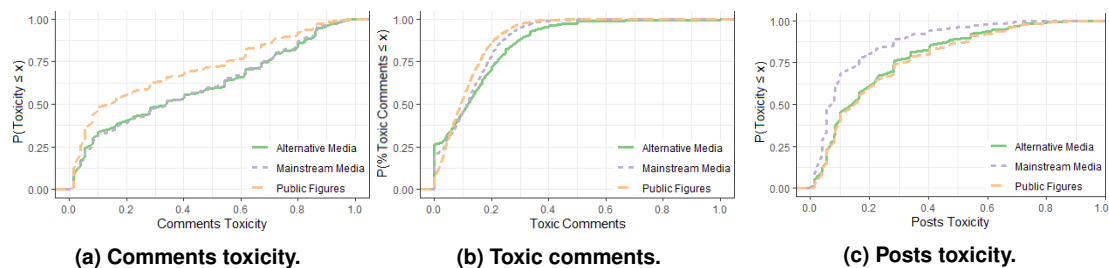


Figura 4. Cumulative Distribution Function (CDF) of toxicity in comments and posts by page type.

The posts’ reactions, also analyzed, confirm these trends from our toxicity results. Using the second extended dataset, we examine correlations between the toxicity and these reactions. We show page reaction profiles that indicated relevant differences between the pages, with some of them also related to our page type division.

When we look into the relationship between reactions and page types, the statistical significance is not found for the more negative emotions, *Angry* and *Sad*. However, for *Haha* and *Wow*, we discovered a relation of the proportion of the reactions that are: Mainstream Media > Alternative Media > Public Figures, as shown in Figures 5a to 5c. The exact reversed relationship appears for the *Love* reaction. This discrepancy between *Haha* and *Love* might indicate that the first of the two might be used as a negative reaction sometimes.

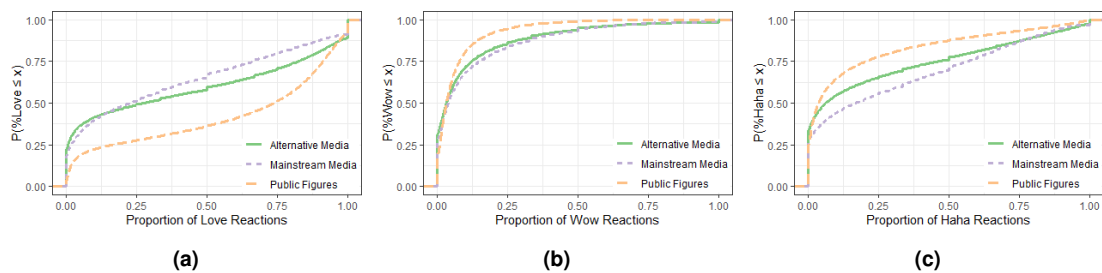


Figure 5. Cumulative density Function (CDF) of the reactions proportions grouped by page type

Meanwhile, Figures 6a to 6d shows that for the two more negative responses, posts from left-wing pages receive significantly more of those reactions than the other political sides. When we consider only the *Angry* reaction, we see that the users give statistically less of it to right-wing pages compared to centrist ones.

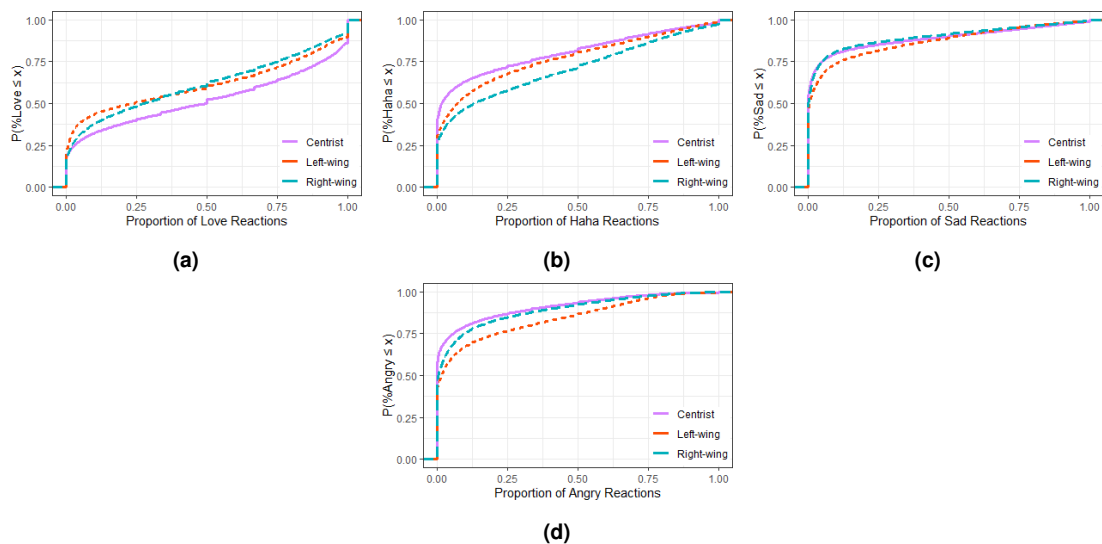


Figure 6. Cumulative Distribution function (CDF) of the reactions proportions grouped by political bias.

The same right-leaning pages have, with statistical significance, the posts with the most of *Haha* reactions of all three ideologies. Finally, for the *Wow* reaction, we did not

find any statistically significant differences.

Lastly, we also used the entropy used in information theory to measure the posts' controversy, using the reaction percentages. This entropy also showed to have a dependent relation to the proportion of toxic messages the post receives, similar to the results of toxicity.

6. Academic and Social Relevance

When we consider the academic relevance of our work, the impact of this dissertation starts with the two published papers [Guimarães et al. 2020a, Guimarães et al. 2020b], the first focused on the toxicity analysis, and another one presenting our political bias method. A third extended paper with more details of our methodology, and more characterizations of the pages found, including the post reactions, is under review to be published in a journal. Another relevant factor of this work is apparent when evaluating our method compared to other researches. One of its advantages is its applicability, as it is appropriate to any country where the Marketing API is available, and it is also more automated than other strategies. Finally, the characterization of the media ecosystem we did also was able to bring more transparency to the environment. Moreover, this transparency also becomes one of the principal social impacts of our work. Using a partnership with Facebook that gave us access to a platform called Crowdtangle, which has more detailed information on pages, we created a live visualization of the content generated by the pages we found, on the left and right, alternative or mainstream media.

Using this methodology and the new platform access, aim at the 2022 Brazilian Election, we will add a tool to the **Eleições sem Fake project**. Since 2018, this research project has the main objective of increasing transparency to the dissemination of content, whether organic or through ads on social media, trying to help to mitigate the problem of disinformation. This new system⁵ will offer an insight into the content that is discussed by different types of media. Thus we hope to provide an easier way to identify political bubbles and detect early the narratives from opposing perspectives that are created and spread on social media.

7. Acknowledgments

This work was partially supported by the Ministério Público de Minas Gerais (MPMG), project Analytical Capabilities, as well as grants from CNPq, CAPES, and Fapemig.

Referências

- Araujo, R., Maciel, R., and Boscaroli, C. (2017). I grandsi-br: Grandes desafios de pesquisa em sistemas de informação no brasil (2016-2026). *Relatório Técnico. Comissão Especial de Sistemas de Informação (CE-SI) da Sociedade Brasileira de Computação (SBC)*. 67p.
- Bakshy, E., Messing, S., and Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132.

⁵Preliminary available on <https://www.dcc.ufmg.br/~samuel.guimaraes/PoliticalBias>

- Basile, A., Caselli, T., and Nissim, M. (2017). Predicting controversial news using facebook reactions. In *Proceedings of the 4th Italian Conference on Computational Linguistics*, pages 12–17.
- Budak, C., Goel, S., and Rao, J. M. (2016). Fair and balanced? quantifying media bias through crowdsourced content analysis. *Public Opinion Quarterly*, 80(S1):250–271.
- Cover, T. and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27.
- D’Alessio, D. and Allen, M. (2000). Media bias in presidential elections: A meta-analysis. *Journal of communication*, 50(4):133–156.
- de Pelle, R. P. and Moreira, V. P. (2017). Offensive comments in the brazilian web: a dataset and baseline results. In *Proceedings of the 6th Brazilian Workshop on Social Network Analysis and Mining*, pages 510–519.
- Druckman, J. N. and Parkin, M. (2005). The impact of media bias: How editorial slant affects voters. *The Journal of Politics*, 67(4):1030–1049.
- Entman, R. M. (2007). Framing bias: Media in the distribution of power. *Journal of communication*, 57(1):163–173.
- Floyd, R. W. (1962). Algorithm 97: shortest path. *Communications of the ACM*, 5(6):345.
- Fortuna, P. and Nunes, S. (2018). A survey on automatic detection of hate speech in text. *ACM Computing Surveys (CSUR)*, 51(4):85.
- Guimarães, S. S., Reis, J. C. S., Lima, L., Ribeiro, F. N., Vasconcelos, M., An, J., Kwak, H., and Benevenuto, F. (2020a). Identifying and characterizing alternative news media on facebook. In *Proceedings of the 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 448–452.
- Guimarães, S. S., Reis, J. C. S., Ribeiro, F. N., and Benevenuto, F. (2020b). Characterizing toxicity on facebook comments in brazil. In *Proceedings of the Brazilian Symposium on Multimedia and the Web*, pages 253–260.
- Haynes, B. (2018). Facebook retira do ar rede ligada ao mbl antes das eleições. <https://br.reuters.com/article/idBRKBN1KF1MI-OBRDN>. Accessed September 30th, 2020.
- Hille, S. and Bakker, P. (2014). Engaging the social news user: Comments on news sites and facebook. *Journalism Practice*, 8(5):563–572.
- Holt, K. (2018). Alternative media and the notion of anti-systemness: Towards an analytical framework. *Media and Communication*, 6(4):49–57.
- Holt, K., Ustad Figenschou, T., and Frischlich, L. (2019). Key dimensions of alternative news media. *Digital Journalism*, 7(7):860–869.
- Jigsaw (2017). Perspective api. <https://www.perspectiveapi.com/>. Accessed May 24th, 2019.
- Joachims, T. (2003). Transductive learning via spectral graph partitioning. In *Proceedings of the 20th International Conference on Machine Learning*, pages 290–297.

- Kelkar, S. (2019). Post-truth and the search for objectivity: political polarization and the remaking of knowledge production. *Engaging Science, Technology, and Society*, 5:86–106.
- Khan, S. A. and Chang, H.-T. (2019). Comparative analysis on facebook post interaction using dnn, elm and lstm. *PloS ONE*, 14(11):1–26.
- Kolhatkar, V. and Taboada, M. (2017). Constructive language in news comments. In *Proceedings of the First Workshop on Abusive Language Online*, pages 11–17.
- Mitchell, A. (2014). *Political polarization & media habits*. Pew Research Center.
- Moers, T., Krebs, F., and Spanakis, G. (2018). Semtec: social emotion mining techniques for analysis and prediction of facebook post reactions. In *Proceedings of the 10th International Conference on Agents and Artificial Intelligence*, pages 361–382.
- Moretto, M. and Ortellado, P. (2018). Quanto mais velhos, mais polarizados: Perfil dos usuários que interagem com páginas de notícias no facebook. Technical Report 1, Monitor do Debate Político no Meio Digital.
- Mullainathan, S. and Shleifer, A. (2002). Media bias. Technical report, National Bureau of Economic Research.
- Newman, N. (2011). Mainstream media and the distribution of news in the age of social media. RISJ Report 2011, Reuters Institute for the Study of Journalism.
- Newman, N., Fletcher, R., Kalogeropoulos, A., and Nielsen, R. (2019). *Reuters Institute digital news report 2019*. Reuters Institute for the Study of Journalism.
- Reis, J., Benevenuto, F., Vaz de Melo, P., Prates, R., Kwak, H., and An, J. (2015). Breaking the news: First impressions matter on online news. In *Proceedings of the 9th International AAAI Conference on Web-Blogs and Social Media*, pages 357–366.
- Ribeiro, F. N., Henrique, L., Benevenuto, F., Chakraborty, A., Kulshrestha, J., Babaei, M., and Gummadi, K. P. (2018). Media bias monitor: Quantifying biases of social media news outlets at large-scale. In *Proceedings of the 12th International AAAI Conference on Web and Social Media*, pages 290—299.
- Ross, B., Rist, M., Carbonell, G., Cabrera, B., Kurowsky, N., and Wojatzki, M. (2016). Measuring the reliability of hate speech annotations: The case of the european refugee crisis. In *Bochumer linguistische Arbeitsberichte 17, 3rd Workshop on Natural Language Processing for Computer-Mediated Communication*, pages 6–9.
- Russell, A. (2011). The arab spring— extra-national information flows, social media and the 2011 egyptian uprising. *International Journal of Communication*, 5:10.
- Shearer, E. and Matsu, K. E. (2018). News use across social media platforms 2018. <https://www.journalism.org/2018/09/10/news-use-across-social-media-platforms-2018/>. Accessed December 15th, 2019.
- Shi, Y., Mast, K., Weber, I., Kellum, A., and Macy, M. (2017). Cultural fault lines and political polarization. In *Proceedings of the 9th ACM on Web Science Conference*, pages 213–217.

- Stroud, N. J., Van Duyn, E., and Peacock, C. (2016). Survey of commenters and comment readers. <https://mediaengagement.org/research/survey-of-commenters-and-comment-readers/>. Accessed February 21th, 2020.
- Su, L. Y.-F., Xenos, M. A., Rose, K. M., Wirz, C., Scheufele, D. A., and Brossard, D. (2018). Uncivil and personal? comparing patterns of incivility in comments on the facebook pages of news outlets. *New Media & Society*, 20(10):3678–3699.
- Tian, Y., Galery, T., Dulcinati, G., Molimpakis, E., and Sun, C. (2017). Facebook sentiment: Reactions and emojis. In *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media*, pages 11–16.
- Weeks, B. E. and Holbert, R. L. (2013). Predicting dissemination of news content in social media: A focus on reception, friending, and partisanship. *Journalism & Mass Communication Quarterly*, 90(2):212–232.
- Wells, C., Zhang, Y., Lukito, J., and Pevehouse, J. C. W. (2020). Modeling the formation of attentive publics in social media: The case of donald trump. *Mass Communication and Society*, 23(2):181–205.
- Zhou, D., Bousquet, O., Lal, T. N., Weston, J., and Schölkopf, B. (2003). Learning with local and global consistency. In *Proceedings of the 16th International Conference on Neural Information Processing Systems*, pages 321–328.
- Zhu, X., Ghahramani, Z., and Lafferty, J. D. (2003). Semi-supervised learning using gaussian fields and harmonic functions. In *Proceedings of the 20th International Conference on Machine Learning*, pages 912–919.